DOCUMENT RESUME

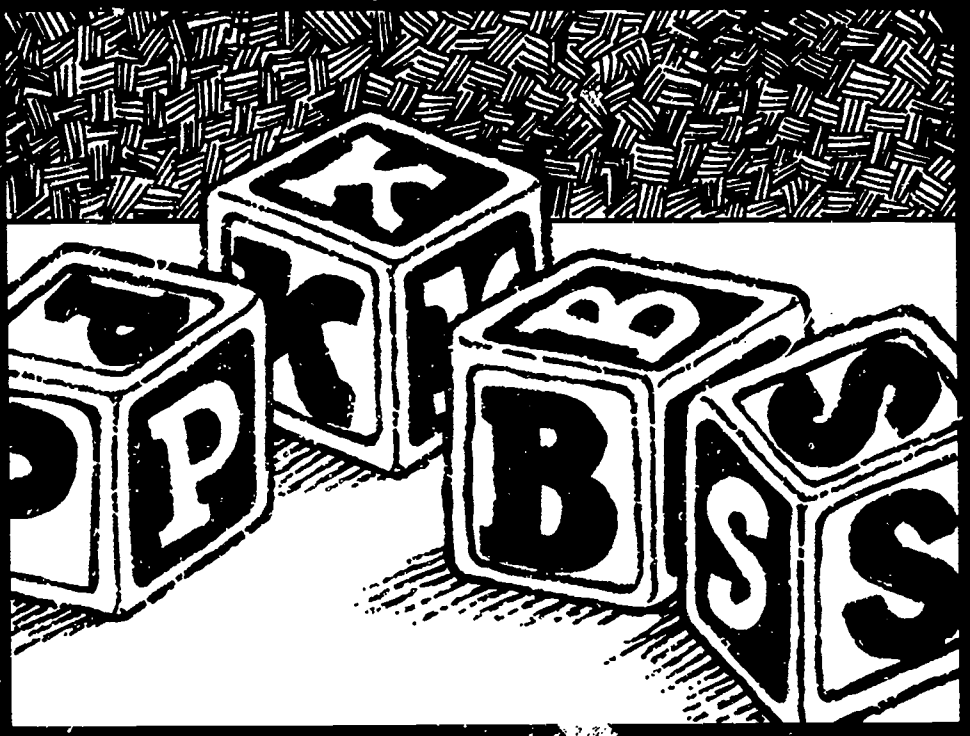
ED 378 210 TM 022 556

AUTHOR Merrell, Kenneth W.
TITLE Preschool and Kindergarten Behavior Scales. Test Manual.
REPORT NO ISBN-0-88422-152-0
PUB DATE 94
NOTE 72p.
AVAILABLE FROM Clinical Psychology Publishing Company, Inc., 4 Conant Square, Brandon, VT 05733.
PUB TYPE Guides - Non-Classroom Use (055) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Behavior Patterns; Behavior Problems; *Behavior Rating Scales; *Educational Assessment; *Interpersonal Competence; Intervention; Kindergarten; *Kindergarten Children; Norm Referenced Tests; Preschool Children; Preschool Education; Primary Education; Scoring; Standardized Tests; Testing; Test Interpretation; *Test Use; User Needs (Information)
IDENTIFIERS *Preschool and Kindergarten Behavior Scales

ABSTRACT

The Preschool and Kindergarten Behavior Scales (PKBS) is a behavioral rating instrument for use in evaluating social skills and problem behavior patterns of preschool- and kindergarten-aged children, aged three through six. It is a norm-referenced, standardized instrument developed specifically for use in assessing young children in a variety of settings and by a variety of behavioral informants. The PKBS includes two major scales: social skills and problem behavior. In addition to an overview of the test, this manual contains chapters related to administering, scoring, and interpreting and PKBS; the technical properties, including reliability and validity, of the instrument; linking the assessment process to behavioral interventions; and information on its development and standardization. (Contains 58 references.) (SLD)



Preschool and Kindergarten Behavior Scales

PRESCHOOL AND KINDERGARTEN BEHAVIOR SCALES

Test Manual

Kenneth W. Merrell
Department of Psychology
Utah State University

CPPC
Clinical Psychology Publishing Company, Inc.
4 Conant Square, Brandon, VT 05733

Library of Congress Cataloging-in-Publication Data

Merrell, Kenneth W.
Preschool and kindergarten behavior scales / Kenneth W. Merrell.
p. cm.
Includes bibliographical references.
ISBN 0-88422-152-0
1. Preschool and Kindergarten Behavior Scales. I. Title.
[DNLM: 1. Child Behavior Disorders—diagnosis. 2. Child Development. 3. Psychiatric Status Rating Scales—in infancy & childhood. 4. Psychological Tests—in infancy & childhood WS 350.6 M568p 1994]
RJ503.7.P73M47 1994
618.92'89075—dc20
DNLM/DLC
for Library of Congress 94-12893
CIP

Library of Congress Catalog Card Number: 94-12893

ISBN: 0-88422-152-0

CPPC
4 Conant Square
Brandon, Vermont 05733

Cover design: Susan Thomas

Printed in the United States of America.

ACKNOWLEDGMENTS

The development of the Preschool and Kindergarten Behavior Scales (PKBS) began in 1991 and has proceeded to the point of this first edition with the assistance of many individuals. I have been particularly fortunate in this venture to have had the help of an excellent research team of graduate and undergraduate students at Utah State University. Camille Odell provided invaluable assistance in her role as research team coordinator, in working closely with many preschool teachers during the reliability and validity studies, and in doing a great deal of background research during the item development phase. Derrick Bootle, Anna Marie Gay, Jennifer Stepan, and Julie Zollinger each spent countless hours scoring and rescoring PKBS protocols and meticulously coding an immense amount of data for later analysis. Clarice Jentzsch's work on the design and layout of the PKBS rating form was beneficial in developing a final test protocol that is both easy to use and professional in appearance.

Without the assistance of the many teachers and parents who provided ratings of the children in the standardization sample, the PKBS research could not have been completed. A number of people, too numerous to mention by name, volunteered to act as data coordinators at the various standardization sites across the nation, and I am indebted to these individuals for promoting the project and actually getting the research instruments into the hands of the parents and teachers who completed them. Of the many individuals who provided assistance in this regard, I am particularly grateful to Debbie Gugino and Judy Miller of the Clark County Public School District in Las Vegas, Nevada, and to Ann Miller and Jed Waddoups of the Weber County Public School District in Ogden, Utah, all of whom were instrumental in helping me gain access to very large and diverse populations that became part of the PKBS standardization sample.

Finally, I am grateful to Dr. Gerald B. Fuller and the staff of Clinical Psychology Publishing Company, Inc., for the confidence and support they placed in this project during the early stages of development, and for the professional manner in which the final production of the scales and test manual were facilitated.

As I have stated at the end of Chapter 4, this first edition of the PKBS manual is considered to be a preliminary effort, with ongoing research and subsequent revisions to the manual planned. I would be interested in receiving copies of research reports and conference presentations in which the PKBS was used, so that this information could be referenced in future editions of the test manual. Please send copies of research reports to Kenneth W. Merrell, Ph.D., Department of Psychology, Utah State University, Logan, UT 84322-2810.

TABLE OF CONTENTS

1

INTRODUCTION AND OVERVIEW

The *Preschool and Kindergarten Behavior Scales* (PKBS) is a behavioral rating instrument for use in evaluating social skills and problem behavior patterns of preschool and kindergarten-aged children (ages 3, 4, 5, and 6). It is a norm-referenced, standardized instrument developed specifically for use in assessing young children in a variety of settings and by a variety of behavioral informants. The PKBS includes two separate major scales: *Social Skills* (34 items) and *Problem Behavior* (42 items). Each of these scales includes a number of empirically derived subscales that are useful for identifying specific clusters or subdomains of social skills and problem behaviors. The Social Skills scale includes items that describe positive social skills characteristic of well-adjusted children in the 3 to 6 age range. The Problem Behavior scale includes items that describe various problem behaviors commonly seen in 3- to 6-year-old children who are experiencing adjustment problems. The PKBS differs from most other behavioral rating scales that may be used with the preschool and kindergarten population in that (a) it was developed and designed *specifically* for use with this age group, and (b) it contains two separate scales (Social Skills and Problem Behavior) that were normed with the same standardization population.

The PKBS was developed to be used for the following purposes:

1. As a screening tool for identifying preschool and kindergarten-aged children who are at risk for developing serious behavioral, social, and emotional problems.
2. As part of a multi-axial assessment battery for formally identifying and classifying children with severe behavioral and emotional problems.
3. As a tool for assessing social skills deficiencies and behavioral problems for the purpose of developing appropriate interventions.
4. As a research instrument for studying the developing social-behavioral patterns of young children.

In addition to this introductory chapter, this manual contains chapters pertaining to administering, scoring, and interpreting the PKBS; and information on the development and standardization of the PKBS, the technical properties of the instrument, and linking the assessment process to behavioral interventions. This introductory chapter includes additional information on the need for early assessment and intervention, dimensions of early childhood behavioral development, and an overview of the content and organization of the PKBS.

The Need for Early Assessment and Intervention

Two recent forces have had the effect of strongly increasing the need for early assessment and intervention (i.e., educational and psychological services) with young children and their families. The first force is the impact of federal legislation that has been enacted since the 1970s. Public Law 94-142,

the Education of All Handicapped Children Act of 1975, ensured that children and adolescents ages 3 through 21 with disabilities would be entitled to a free and appropriate public education and related services. Public Law 99-457, known as the Education of the Handicapped Act Amendments of 1986, extended the provision of special education–related services to the age range of birth to 3. Both of these laws have since been amended by P.L. 101-476, the Individuals with Disabilities Education Act (IDEA) of 1990. IDEA reaffirmed, expanded, and redefined the importance and necessity of special education and related services to children with disabilities, including the early childhood/preschool-age population and their families. The second force that is currently increasing the need for the provision of psychological services to at-risk young children and their families is the complex combination of demographic, economic, and social influences within the United States. In its most recent report on the state of America's children, the Children's Defense Fund (1992) painted an alarming picture of a nation in which an increasing percentage of young children are living in poverty, lacking basic health care services, becoming increasingly vulnerable to violence and family crises, wanting for safe child care, and generally becoming increasingly at risk for a wide array of developmental problems. When commenting on the state of mental health services for children, Tuma (1989) stated that children continue to be one of the most neglected groups in mental health, and that large proportions of at-risk and disabled children do not receive adequate services. Thus, the necessity of providing educational and psychological services to young children in the United States is not only codified within federal law, but is becoming of paramount importance as more of these children become vulnerable to a variety of developmental problems.

Although there is a strong demonstrated need for broad-based behavioral assessment and screening instruments such as the PKBS, relatively few are available for use with the early childhood/preschool population. In comparison with the large number of assessment instruments developed for use with the school-age population, much less attention has been given to instruments for use with the early childhood/preschool population. Of the instruments that are currently available for use with this age group, many have inadequate standardization samples and psychometric properties. Some behavioral assessment and screening instruments currently available for use with the early childhood/preschool population may have adequate psychometric properties, but are merely downward extensions of instruments developed for use with school-age children and were not conceptualized and designed specifically for use with this younger age group. Thus the PKBS represents a much needed and major effort at improving the utility of behavioral assessment and screening with children in the 3-6 age range.

Dimensions of Behavioral Development During Early Childhood

As the research base on children's social competence and psychopathology has grown over the past two decades, so has the accuracy of classifying these two forms of behavior into more specific and concise categories. This practice, known as the *behavioral dimensions* approach to classification (Merrell, 1994), shows considerable promise as an empirically sound way of classifying children's behavioral, social, and emotional problems. The behavioral dimensions approach is based on the use of complex statistical procedures such as factor analysis, cluster analysis, and path analysis, which provide a methodology for identifying intercorrelated behavioral syndromes or clusters.

Social competence is considered to be a highly complex and multidimensional construct comprising interrelated behaviors that are prosocial, adaptive, and positive. Gresham (1986) conceptualized the broad domain of social competence as being composed of the following three subdomains: (a) adaptive behavior, (b) social skills, and (c) peer acceptance. Because the PKBS was not intended to measure the full range of preschool-age adaptive behaviors (e.g., independence skills, communication skills, self-care skills) but to focus more on the social skills and peer relations aspects of social competence, the latter two domains are emphasized in the test and in this manual.

When initiated, social skills lead to desirable social outcomes. From a behavioral standpoint, initiation of social skills increases the probability of reinforcement and decreases the probability of punishment

or extinction based upon one's social behavior (Gresham & Reschly, 1987). For children, examples of behavioral classes representing social skills include academic and task-related competence, cooperation with peers, reinforcement of peers' behavior, and social initiation behaviors. Although peer relations has been considered to be the third overall component or domain of social competence, it is often thought of as a *result or product* of one's social skills. This view is reasonable in that social reputation and the quality of one's social relations are often a result of how effectively one interacts socially with peers (Landau & Milich, 1990; Oden & Asher, 1977). Positive peer relations are associated with peer acceptance, whereas negative peer relations are linked with peer rejection.

Since about the early 1980s, a few researchers have attempted to develop dimensional classification taxonomies for social skills (e.g., Gresham & Elliott, 1990; Merrell, 1993a; Walker & McConnell, 1988). For example, Walker, McConnell, and Clarke (1985) identified two major types of social adjustment that children must make upon entering school, namely *peer related* and *teacher related*. Given that there is more variation in types of formal education and other structured experiences for preschool-age than school-age children, the second behavioral domain is perhaps better thought of and referred to as *adult related* rather than *teacher related*. Peer-related social adjustment involves the social and behavioral dynamics that occur between children in unstructured and free-play settings (e.g., engaging in play activities with other children and sharing toys), whereas adult-related social adjustment involves meeting the social and behavioral expectations and demands of adult caregivers such as parents, daycare providers, and preschool teachers (e.g., following rules and cooperating).

In contrast to the relatively recent and fairly sparse literature on dimensions of child social competence, a large body of research evidence exists for the dimensional approach to classifying behavioral and emotional problems. Experts in the field of child psychopathology are in general agreement that child behavioral and emotional disorders can be divided into two broad classes or dimensions, namely *internalizing* and *externalizing* problems (Cicchetti & Toth, 1991). Sometimes these two behavioral dimensions have been referred to as *overcontrolled* and *undercontrolled*, but the types of problems seen within each of the two dimensions are similar.

The internalizing dimension of behavioral and emotional problems includes symptoms of depression, social withdrawal, anxious and inhibited reactions, and the development of somatic (physical) problems that appear to be related to inner emotional distress. Although these different internalizing symptoms superficially appear to be distinct from each other, there is strong evidence of a *comorbid* relationship among them, meaning that a child who exhibits symptoms of depressive withdrawal or excessive sadness is highly likely to also exhibit some related internalizing symptoms such as somatic problems, fears, and anxiety (Masser & Cloninger, 1990).

The externalizing dimension, on the other hand, consists of aggressive, defiant, acting-out, disruptive, oppositional, and hyperactive behaviors. Like the internalizing dimension, behavioral problems within the externalizing dimension are thought to exist in a comorbid relationship, in that there is usually considerable relatedness or congruence among externalizing symptoms. For example, it is often quite difficult to distinguish Attention-Deficit Hyperactivity Disorder from Conduct Disorders, because so many of the symptoms overlap in individual cases (Campbell & Werry, 1986). Even during the early stage of behavioral development of the early childhood/preschool population, highly active and aggressive behaviors are often found to coexist (Campbell, 1991).

In sum, children's behavioral development can be conceptualized along two general dimensions, namely social competence and problem behaviors. Within each of these general dimensions, there is evidence that behaviors tend to aggregate in related clusters.

Content and Organization of the Preschool and Kindergarten Behavior Scales

The content and organization of the PKBS was developed to approximate closely the dimensions of early childhood behavioral development that were discussed in the previous section. Therefore the items in the Social Skills scale (Scale A) were designed to reflect both peer-related and adult-related

forms of social adjustment, whereas the items in the Problem Behavior scale (Scale B) were designed to reflect both internalizing and externalizing forms of problem behavior.

Scale A: Social Skills

The Social Skills scale includes 34 items that describe adaptive or positive behaviors that are likely to lead to positive personal and social outcomes. These items are rated on a 4-point scale where the anchor points are as follows:

0 = Never
1 = Rarely
2 = Sometimes
3 = Often

The items within Scale A are divided into three empirically derived subscales:

Subscale A1, Social Cooperation, includes 12 items that reflect behaviors and characteristics deemed important in following instructions from adults, cooperating and compromising with peers, and showing appropriate self-restraint. The items in this subscale appear to be linked to both the peer-related and adult-related forms of social adjustment, but are probably more strongly connected to adult-related social adjustment, as most of the items involve appropriate compliance with the types of structure and regulation typically imposed by parents, preschool teachers, and daycare providers. Examples of items in this subscale include *Follows instructions from adults, Uses free time in an acceptable way, Takes turns with toys or other objects,* and *Responds appropriately when corrected.*

Subscale A2, Social Interaction, includes 11 items that reflect behaviors and characteristics deemed important in gaining and maintaining acceptance and friendship from others. The items in this subscale appear to be primarily connected to the peer-related form of social adjustment, though a few of these items do involve appropriate social interaction with adults. Examples of items in this subscale include *Stands up for other children's rights, Has skills or abilities that are admired by peers, Comforts other children who are upset,* and *Invites other children to play.*

Subscale A3, Social Independence, includes 11 items that reflect behaviors and characteristics deemed important in achieving social independence within the domain of the peer group. Some of these social skills involve being able to separate appropriately from adult caregivers, whereas others involve showing appropriate confidence and positive assertiveness in interactions with peers. The skills measured within this subscale appear to be primarily linked to the peer-related form of social adjustment. Examples of items in this subscale include *Is confident in social situations, Plays with several different children, Makes friends easily,* and *Is able to separate from parent without extreme distress.* The items in Scale A are shown, divided by subscale area, in Table 1.1.

Scale B: Problem Behavior

The Problem Behavior scale includes 42 items that describe various problem behaviors commonly seen in the early childhood/preschool population. These items are rated on the same 4-point scale previously described for Scale A. The items within Scale B are divided into five empirically derived narrow-band subscales, and two empirically derived broad-band scales. The broad-band scales are based on a division into the externalizing and internalizing behavioral dimensions, which were discussed in the previous section of this chapter. Subscales B1, B2, and B3 are considered to be externalizing problem scales, whereas subscales B4 and B5 are considered to be internalizing problem scales. More information on the relationship between the narrow-band and broad-band subscales and how they were derived is presented in later chapters of this manual. The five narrow-band subscales, which are presented in Table 1.2, are described as follows:

Subscale B1, Self-Centered/Explosive, includes 11 items indicative of volatile, inconsiderate, and unpredictable behaviors. Children who receive high scores on this subscale might be described as oppositional, defiant, emotionally and behaviorally labile, and unpredictably temperamental. This cluster of behaviors aligns within the externalizing problem behavior domain. Examples of items in this subscale include *Has temper outbursts or tantrums, Yells or screams when angry, Defies parent or caregiver,* and *Is temperamental or irritable.*

Subscale B2, Attention Problems/Overactive, is an externalizing problem subscale that includes eight items indicative of an impulsive, restless, fidgety, noisy, and overactive behavioral style. Children who are identified as having Attention-Deficit Hyperactivity Disorder will likely receive high scores on this subscale. Examples of items within this subscale include *Acts impulsively without thinking, Makes noises that annoy others, Has difficulty concentrating or staying on-task,* and *Disrupts ongoing activities.*

Subscale B3, the third externalizing problems subscale, includes eight items involving intimidation and harm to others through coercive means. Children who receive high scores on this subscale are likely to be disliked and avoided by other children because of their hurtful and coercive style of interacting with peers, particularly if they have concurrent deficits in social skills. Examples of items within this subscale include *Teases or makes fun of other children, Is physically aggressive, Calls people names,* and *Bullies or intimidates other children.*

Subscale B4, Social Withdrawal, includes seven items that describe a pattern of avoidance and withdrawal from other children that may be accompanied by excessive unhappiness. Consistent with the research literature, this subscale falls within the internalizing problems domain. Examples of items within this subscale include *Avoids playing with other children, Has problems making friends, Withdraws from the company of others,* and *Seems unhappy or depressed.*

Subscale B5, Anxiety/Somatic Problems, is an internalizing problems subscale that consists of eight items. These items are indicative of a fearful, tense, and anxious pattern of behavior that may be accompanied by various somatic or physical problems, such as stomachaches, pain, and sickness, for which no physical cause is known. Children who are rated high on this subscale often have difficulty separating from parents and caregivers, are fearful, and may internalize these characteristics into physical problems. Examples of items from this subscale include *Becomes sick when upset or afraid, Clings to parent or caregiver, Is afraid or fearful,* and *Complains of aches, pains, or sickness.*

Procedures for obtaining scale and subscale scores, as well as technical information on the structure of the scales, are presented in subsequent chapters of this manual.

Table 1.1

Items in PKBS Scale A, Social Skills, Listed by Subscale

Number	Item wording
	Subscale A1, Social Cooperation
2	Is cooperative
7	Follows instructions from adults
10	Shows self-control
12	Uses free time in an acceptable way
16	Sits and listens when stories are being read
22	Cleans up his/her messes when asked
23	Follows rules
25	Shares toys and other belongings
28	Gives in or compromises with peers when appropriate
29	Accepts decisions made by adults
30	Takes turns with toys and other objects
32	Responds appropriately when corrected
	Subscale A2, Social Interaction
5	Tries to understand another child's behavior ("Why are you crying?")
14	Participates in classroom or family discussions
15	Asks for help from adults when needed
17	Stands up for other children's rights ("That's his!")
19	Has skills or abilities that are admired by peers
20	Comforts other children who are upset
21	Invites other children to play
24	Seeks comfort from an adult when hurt
27	Apologizes for accidental behavior that may upset others
33	Is sensitive to adult problems ("Are you sad?")
34	Shows affection for other children
	Subscale A3, Social Independence
1	Works or plays independently
3	Smiles and laughs with other children
4	Plays with several different children
6	Is accepted and liked by other children
8	Attempts new tasks before asking for help
9	Makes friends easily
11	Is invited by other children to play
13	Is able to separate from parent without extreme distress
18	Adapts well to different environments
26	Stands up for his/her rights
31	Is confident in social situations

Table 1.2

Items in PKBS Scale B, Problem Behavior, Listed by Subscale

Number	Item wording
	Subscale B1, Self-Centered/Explosive
7	Has temper outbursts or tantrums
8	Wants all the attention
10	Will not share
13	Yells or screams when angry
19	Must have his/her own way
22	Defies parent, teacher, or caregiver
31	Has unpredictable behavior
32	Is jealous of other children
33	Is moody or temperamental
37	Whines or complains
41	Is easily provoked; has a "short fuse"
	Subscale B2, Attention Problems/Overactive
1	Acts impulsively without thinking
6	Makes noises that annoy others
14	Takes things away from other children
15	Has difficulty concentrating or staying on-task
16	Disobeys rules
20	Is overly active; unable to sit still
25	Is restless and "fidgety"
39	Disrupts ongoing activities
	Subscale B3, Antisocial/Aggressive
3	Teases or makes fun of other children
11	Is physically aggressive (hits, kicks, pushes)
21	Seeks revenge against others
26	Calls people names
29	Bullies or intimidates other children
34	Destroys things that belong to others
40	Tells lies
42	Bothers and annoys other children

(continued)

Table 1.2 (continued)

	Subscale B4, Social Withdrawal
4	Does not respond to affection from others
12	Avoids playing with other children
17	Has problems making friends
27	Is difficult to comfort when upset
28	Withdraws from the company of others
30	Seems unhappy or depressed
33	Acts younger than his/her age
	Subscale B5, Anxiety/Somatic Problems
2	Becomes sick when upset or afraid
5	Clings to parent or caregiver
9	Is anxious and tense
18	Is afraid or fearful
23	Complains of aches, pain, or sickness
24	Resists going to preschool or day care
36	Is overly sensitive to criticism or scolding
38	Gets taken advantage of by other children

2

ADMINISTRATION, SCORING, AND INTERPRETATION

Within this chapter, recommended procedures for administering the PKBS and obtaining and using scores and score levels are overviewed. First, recommended test administration procedures are discussed. Second, procedures for obtaining raw scores and converting these scores into standard scores, percentile scores, and functional levels are presented. Finally, some practical guidelines for interpreting test scores and linking assessment to intervention are provided.

Test Administration

The PKBS is a broad-based behavior rating scale for use with children between the ages of 3 and 6. It is designed primarily to be completed by parents and teachers, but may also be completed by other persons who know the child well enough to make an informed rating. Examples of individuals other than parents or teachers who might rate a child using the PKBS include daycare providers, grandparents, foster parents or guardians, social services caseworkers, and other education professionals (e.g., administrators, psychologists, speech-language pathologists, counselors). The main criterion for determining whether an individual should be able to use the PKBS to rate a child should be knowledge of the child rather than the individual's specific relationship to the child. It is recommended that individuals who rate the child who are not the child's parents should have had the opportunity to observe and interact with the child for a minimum of 3 months before rating him or her using the PKBS. Regardless of who completes the rating form, ratings of child behavior should be based on the informants' observations of the child's behavior during the 3-month period prior to the completion of the form.

Completion of the PKBS will take the majority of raters between about 8 and 12 minutes. First, the *Child Information* section on page 1 of the PKBS rating form should be completed as fully as possible. The *Rater Information* section on page 1 of the form should also be completed carefully so that persons other than the rater who use the test results will have relevant information about the person who completed the ratings. After the instructions on page 1 have been read carefully, the rater then provides a rating for all items on Scale A (Social Skills) and Scale B (Problem Behavior), which are found on pages 2 and 3 of the rating form. It is important to complete every item on each of the two scales, even if the rater is not exactly sure how to rate a specific behavior. The best practice when completing each item is to make the most accurate possible estimate for specific items in question, rather than leaving them blank. Failure to complete items makes a meaningful interpretation of the test results difficult, and in some cases impossible. After completion of the two PKBS scales, raters may then provide additional information about the child that they believe might be useful in the *Additional Information* section on the top of page 4. This information is optional, but is recommended, as it may provide significant information to persons who will interpret the scale, based on the rater's unique association with and

observation of the child, as well as any extenuating circumstances that might affect the ratings. After completion of the PKBS rating form, the record form should be given to the individual who will be responsible for scoring and interpretation.

User Qualifications

Although the PKBS can be completed by virtually any rater who knows the target child well and has had ample opportunities to observe the child's behavior, scoring and interpretation of the PKBS, as well as using PKBS scores to make diagnostic or classification decisions, should be done only by individuals with the prerequisite knowledge and training.

It is recommended that users of behavior rating scales such as the PKBS should have a basic understanding of the principles of educational and psychological testing, including the standards for testing that have been developed jointly by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (AERA, APA, & NCME, 1985). Specific training in understanding and assessing child behavioral and emotional problems is also recommended for optimal use of the PKBS. Within this test manual, specific information regarding the psychometric properties of the PKBS (e.g., reliability and validity evidence) is presented, and this information may be very helpful towards understanding and using the test correctly. However, ultimate responsibility for correct use of the PKBS rests with the user of the test, and this instrument should not be utilized for purposes for which it is not purported to be useful.

Examiners who use PKBS data when developing Individualized Education Plans or other types of interventions should have a sound understanding of normal and abnormal social-behavioral development during early childhood. Knowledge of behavioral intervention techniques is especially important when developing intervention for early childhood/preschool populations, as children in this age group typically do not have the cognitive maturity, insight, and verbal mediation skills that are necessary for counseling-based interventions.

Scoring the PKBS

Scoring the PKBS is relatively simple and straightforward and involves two steps:

1. Calculating raw scores for the subscale and total scores.
2. Converting raw scores to standard scores, percentile ranks, and Functional Levels using the raw score conversion tables provided in this chapter of the manual.

How to Treat Missing Data

After the PKBS has been completed by the rater, the individual responsible for scoring and interpreting the test should check the rating form to ensure that all Social Skills and Problem Behavior items have been completed. Because the instructions on the rating form advise the rater to complete all items, and because of the manner in which the items are presented, there will seldom be a problem with missing data. However, in cases where the ratings are incomplete, it is recommended that the following steps be taken in treating missing data. First, the person who completed the PKBS rating form should be contacted, and an effort should be made to have him or her complete the missing ratings. If this first step does not result in full completion of the rating form, then a second step should be taken. The recommended second step is to locate the items on which ratings are missing, and determine the subscales in which the missing items constitute a part. If there are no more than two items missing from any subscale, it is recommended that the test scorer conduct a *mean item substitution*, by calculating the mean rating of the items that have been completed in that subscale, rounding the mean value to the nearest whole number, and using the obtained value for the missing item(s). For example, if items 22 and 32 on subscale A1 (Social Cooperation) were missing, and the mean value of the

remaining 10 items on this subscale was 2.1, a rating of 2 could be substituted for the missing items. If this mean substitution procedure is used, it is recommended that the examiner make note of the missing data and how they were treated in any verbal or written interpretive reports that are made. If more than two items are missing from any subscale, it is recommended that the Functional Level for that subscale, as well as the total score for that scale, should not be obtained, and the ratings should be looked at in a qualitative rather than quantitative manner. In the rare case where many items across subscales are not completed, the examiner is advised simply to not use the obtained PKBS data.

Calculating PKBS Subscale Raw Scores

Raw scores for the PKBS subscales are calculated by entering the circled value for each item into the nonshaded scoring key column with which it corresponds and then summing entered values for each of the vertical columns. The vertical columns (A1, A2, etc.) represent PKBS subscales. There are three vertical subscale columns for Scale A, whereas Scale B has five of these columns. The sum of each vertical column is entered at the bottom of the scoring key, in the blank boxes that are located to the right of the word "Totals," just above the subscale numbers. These subscale raw scores are then entered into the Raw Score section of the *PKBS Score Grid* on page 4 of the rating form. Examples of how to use the scoring keys for Scale A and Scale B, respectively, are provided in Figure 2.1 and Figure 2.2.

Calculating PKBS Total and Area Raw Scores

For Scale A, the total raw score is obtained by summing the three subscale raw scores (A1, A2, A3). Likewise, for Scale B, the total raw score is obtained by summing the five subscale raw scores (B1 through B5). However, unlike Scale A, Scale B also includes two *area scores*, namely *Externalizing Problems* and *Internalizing Problems*. The Externalizing Problems raw score on Scale B consists of the sum of raw scores for subscales B1, B2, and B3. The Internalizing Problems raw score on Scale B consists of the sum of raw scores for subscales B4 and B5. Thus, these two area scores for Scale B are obtained by summing the appropriate subscale raw scores. An example of how to make these score calculations is provided in Figure 2.3, which shows a completed *PKBS Score Grid* based on the raw scores provided in the examples in Figures 2.1 and 2.2.

Converting Raw Scores to Standard Scores, Percentile Ranks, and Functional Levels

Once the raw scores for the PKBS subscale, area, and total scores have been calculated, the next step is to convert these raw scores to appropriate standard scores, percentile ranks, and Functional Levels. These converted scores and levels are entered, along with the obtained raw scores, on the *PKBS Score Grid* on page 4 of the test form. These score conversions are accomplished by taking the following steps:

1. Convert the total scores for both Scale A and Scale B to standard scores, percentile scores, and Functional Levels by using the age-appropriate (3-4 or 5-6) conversion table (Tables 2.1, 2.2, 2.3, and 2.4), located in this chapter.

2. Convert the Externalizing Problem and Internalizing Problem area raw scores to percentile scores by using Table 2.5 and Table 2.6, located in this chapter.

3. Convert the subscale raw scores for Scale A to Functional Levels by using Table 2.7, and convert the subscale and area raw scores for Scale B to Functional Levels by using Table 2.8. Both of these tables are located in this chapter.

Additional Score Comparisons

Additional comparisons of children's PKBS scores can be made by contrasting their raw scores to the appropriate gender and grade level descriptive statistics (e.g., means and standard deviations) from

PKBS

Scale A
Social Skills

	Never	Rarely	Sometimes	Often	Scoring Key		
1. Works or plays independently	0	1	2	(3)			3
2. Is cooperative	0	1	(2)	3	2		
3. Smiles and laughs with other children	0	(1)	2	3			1
4. Plays with several different children	(0)	1	2	3			0
5. Tries to understand another child's behavior ("Why are you crying?")	0	(1)	2	3		1	
6. Is accepted and liked by other children	0	(1)	2	3			1
7. Follows instructions from adults	0	1	(2)	3	2		
8. Attempts new tasks before asking for help	0	(1)	2	3			1
9. Makes friends easily	(0)	1	2	3			0
10. Shows self-control	0	1	(2)	3	2		
11. Is invited by other children to play	0	(1)	2	3			1
12. Uses free time in an acceptable way	0	1	(2)	3	2		
13. Is able to separate from parent without extreme distress	(0)	1	2	3			0
14. Participates in family or classroom discussions	0	(1)	2	3		1	
15. Asks for help from adults when needed	0	1	(2)	3		2	
16. Sits and listens when stories are being read	0	1	2	(3)	3		
17. Stands up for other children's rights ("That's his!")	(0)	1	2	3		0	
18. Adapts well to different environments	(0)	1	2	3			0
19. Has skills or abilities that are admired by peers	0	(1)	2	3		1	
20. Comforts other children who are upset	0	(1)	2	3		1	
21. Invites other children to play	0	(1)	2	3		1	
22. Cleans up his/her messes when asked	0	1	(2)	3	2		
23. Follows rules	0	1	2	(3)	3		
24. Seeks comfort from an adult when hurt	0	1	2	(3)		3	
25. Shares toys and other belongings	0	1	(2)	3	2		
26. Stands up for his/her rights	0	(1)	2	3			1
27. Apologizes for accidental behavior that may upset others	0	1	(2)	3		2	
28. Gives in or compromises with peers when appropriate	0	1	(2)	3	2		
29. Accepts decisions made by adults	0	1	2	(3)	3		
30. Takes turns with toys and other objects	0	1	2	(3)	3		
31. Is confident in social situations	(0)	1	2	3			0
32. Responds appropriately when corrected	0	1	(2)	3	2		
33. Is sensitive to adult problems ("Are you sad?")	0	1	(2)	3		2	
34. Shows affection for other children	0	(1)	2	3		1	
				Totals	28	15	8
					A1	A2	A3

Figure 2.1

An example of converting PKBS item ratings of a 5-year-old child to subscale raw scores using the scoring key for Scale A, Social Skills.

PKBS

Scale B Problem Behavior	Never	Rarely	Sometimes	Often	Scoring Key				
1. Acts impulsively without thinking	0	(1)	2	3		1			
2. Becomes sick when upset or afraid	0	1	2	(3)					3
3. Teases or makes fun of other children	(0)	1	2	3			0		
4. Does not respond to affection from others	0	1	(2)	3				2	
5. Clings to parent or caregiver	0	1	2	(3)					3
6. Makes noises that annoy others	(0)	1	2	3		0			
7. Has temper outbursts or tantrums	(0)	1	2	3	0				
8. Wants all the attention	(0)	1	2	3	0				
9. Is anxious or tense	0	1	2	(3)					3
10. Will not share	0	(1)	2	3	1				
11. Is physically aggressive (hits, kicks, pushes)	(0)	1	2	3			0		
12. Avoids playing with other children	0	1	2	(3)				3	
13. Yells or screams when angry	0	(1)	2	3	1				
14. Takes things away from other children	(0)	1	2	3		0			
15. Has difficulty concentrating or staying on task	0	1	(2)	3		2			
16. Disobeys rules	0	(1)	2	3		1			
17. Has problems making friends	0	1	2	(3)				3	
18. Is afraid or fearful	0	1	2	(3)					3
19. Must have his/her own way	0	(1)	2	3	1				
20. Is overly active; unable to sit still	(0)	1	2	3		0			
21. Seeks revenge against others	(0)	1	2	3			0		
22. Defies parent, teacher, or caregiver	0	(1)	2	3	1				
23. Complains of aches, pain, or sickness	0	1	2	(3)					3
24. Resists going to preschool or day care	0	1	2	(3)					3
25. Is restless and "fidgety"	0	(1)	2	3		1			
26. Calls people names	0	(1)	2	3			1		
27. Is difficult to comfort when upset	0	1	(2)	3				2	
28. Withdraws from the company of others	0	1	2	(3)				3	
29. Bullies or intimidates other children	(0)	1	2	3			0		
30. Seems unhappy or depressed	0	1	2	(3)				3	
31. Has unpredictable behavior	0	(1)	2	3	1				
32. Is jealous of other children	0	(1)	2	3	1				
33. Acts younger than his/her age	0	1	(2)	3				2	
34. Destroys things that belong to others	(0)	1	2	3			0		
35. Is moody or temperamental	0	1	(2)	3	2				
36. Is overly sensitive to criticism or scolding	0	1	(2)	3					2
37. Whines or complains	0	1	(2)	3	2				
38. Gets taken advantage of by other children	0	1	2	(3)					3
39. Disrupts ongoing activities	(0)	1	2	3		0			
40. Tells lies	0	(1)	2	3			1		
41. Is easily provoked; has a "short fuse"	0	(1)	2	3	1				
42. Bothers and annoys other children	0	(1)	2	3			1		
				Totals	11	5	3	18	23
					B1	B2	B3	B4	B5

Figure 2.2

An example of converting PKBS item ratings of a 5-year-old child to subscale raw scores using the scoring key for Scale B, Problem Behavior.

PKBS

Additional Information

Please use the following lines to provide any additional information about this child that you believe would be useful.

Jason has a very difficult time when his mother drops him off at preschool. He often clings to her and cries. At preschool, he often seems lonely and lost, and he has few friends. Jason often complains of stomachaches and seems afraid or nervous. He is a gentle child who will respond to adults once he is comfortable around them.

PKBS Score Grid

PKBS Score	Raw Score	Standard Score	Percentile Rank	Functional Level
SOCIAL SKILLS				
A1 Social Cooperation	28			Average
A2 Social Interaction	15			Moderate deficit
A3 Social Independence	8			Significant deficit
AT **SOCIAL SKILLS TOTAL**	51	63	3	Significant deficit
PROBLEM BEHAVIOR				
B1 Self-Centered/Explosive	11			Average
B2 Attention Problems/Overactive	5			Average
B3 Antisocial/Aggressive	3			Average
Externalizing Problems	19		40	Average
B4 Social Withdrawal	18			Significant Problem
B5 Anxiety/Somatic Problems	23			Significant Problem
Internalizing Problems	41		>99	Significant Problem
BT **PROBLEM BEHAVIOR TOTAL**	60	116	86	Moderate Problem

Note: For Scale A, higher scores indicate greater levels of social skills.
For Scale B, higher scores indicate greater levels of problem behaviors.

Figure 2.3 *An example of a completed PKBS Score Grid, using the raw score data for a 5-year-old child provided in Figure 2.1 and Figure 2.2.*

Table 2.1

Social Skills (Scale A) Total Score Converstion Table for Ages 3-4: Standard Scores, Percentile Ranks, and Functional Levels that Correspond to Raw Scores

Raw score	Standard score	Percentile rank		Raw score	Standard score	Percentile rank
0-29	< 52	< 1		65	87	18
30	52	1		66	88	19
31	53	1		67	89	20
32	54	1		68	90	21
33	55	1		69	91	23
34	56	1		70	92	25
35	57	2		71	93	27
36	58	2		72	94	28
37	59	2		73	95	30
38	60	2		74	96	32
39	61	2		75	97	35
40	62	2		76	98	37
41	63	3		77	99	40
42	64	3		78	100	44
43	65	3		79	101	47
44	66	3		80	102	49
45	67	3		81-82	103	51
46	68	4		83	104	57
47	69	5		84	105	61
48	70	5		85	106	64
49	71	5		86	107	67
50	72	6		87	108	70
51	73	6		88	109	72
52	74	7		89	110	76
53	75	7		90	111	79
54	76	8		91	112	81
55	77	8		92	113	84
56	78	9		93	114	86
57	79	10		94	115	88
58	80	11		95	116	89
59	81	12		96	117	91
60	82	12		97	118	94
61	83	13		98	119	95
62	84	14		99	120	97
63	85	15		100	121	97
64	86	16		101	122	98
				102	123	> 99

standard scores are based on an average score of 100 with a standard deviation of 15

Functional level
High functioning: 91-102
Average: 68-90
Moderate deficit: 50-67
Significant deficit: 0-49

Table 2.2

Problem Behavior (Scale B) Total Score Conversion Table for Ages 3-4: Standard Scores, Percentile Ranks, and Functional Levels that Correspond to Raw Scores

Raw score	Standard score	Percentile rank		Raw score	Standard score	Percentile rank
0	74	1		59-60	113	81
1	75	2		61	114	82
2-3	76	4		62-63	115	84
4	77	6		64	116	86
5-6	78	7		65-66	117	87
7-8	79	9		67	118	89
9	80	11		68-69	119	90
10-11	81	12		70	120	90
12	82	14		71-72	121	91
13-14	83	16		73	122	92
15	84	18		74-75	123	93
16-17	85	19		76-77	124	94
18	86	21		78	125	95
19-20	87	22		79-80	126	95
21	88	23		81	127	96
22-23	89	25		82-83	128	97
24	90	28		84	129	97
25-26	91	29		85-86	130	97
27	92	33		87	131	98
28-29	93	35		88-89	132	98
30-31	94	38		90	133	98
32	95	40		91-92	134	98
33-34	96	42		93	135	98
35	97	45		94-95	136	99
36	98	46		96	137	99
37-38	99	47		97-98	138	99
39-40	100	50		99	139	99
41	101	53		100-101	140	99
42-43	102	55		102-103	141	>99
44	103	59		104	142	>99
45-46	104	61		105-106	143	>99
47	105	64		107	144	>99
48-49	106	66		108	145	>99
50	107	69		109-110	146	>99
51-52	108	71		111-112	147	>99
53	109	73		113	148	>99
55	110	76		114-115	149	>99
56-57	111	78		116	150	>99
58	112	80		117-126	>150	>99

standard scores are based on an average score of 100 with a standard deviation of 15

Functional level
High functioning: 0-17
Average: 18-58
Moderate problem: 59-79
Significant problem: 80-126

Table 2.3

Social Skills (Scale A) Total Score Conversion Table for Ages 5-6: Standard Scores, Percentile Ranks, and Functional Levels that Correspond to Raw Scores

Raw score	Standard score	Percentile rank		Raw score	Standard score	Percentile rank
0-41	<53	<1		72	86	15
42	53	1		73	87	16
43	54	1		74	88	18
44	55	2		75	90	19
45	56	2		76	91	21
46	57	2		77	92	23
47	58	2		78	93	26
48	60	2		79	94	28
49	61	2		80	95	30
50	62	3		81	96	33
51	63	3		82	97	36
52	64	3		83	99	39
53	65	3		84	100	43
54	66	4		85	101	46
55	67	4		86	102	50
56	68	4		87	103	53
57	70	5		88	104	56
58	71	5		89	105	60
59	72	6		90	106	63
60	73	6		91	108	66
61	74	7		92	109	70
62	75	8		93	110	73
63	76	8		94	111	77
64	77	9		95	112	81
65	78	9		96	113	83
66	79	10		97	114	87
67	81	10		98	115	90
68	82	11		99	116	92
69	83	12		100	118	95
70	84	13		101	119	96
71	85	14		102	120	>99

standard scores are based on an average score of 100 with a standard deviation of 15

Functional Level
High functioning: 95-102
Average: 76-94
Moderate deficit: 59-75
Significant deficit: 0-58

Table 2.4

Problem Behavior (Scale B) Total Score Conversion Table for Ages 5-6: Standard Scores, Percentile Ranks, and Functional Levels that Correspond to Raw Scores

Raw score	Standard score	Percentile rank	Raw score	Standard score	Percentile rank
0	75	≤2	57	114	84
1-2	76	4	58-59	115	85
3-4	77	6	60	116	86
5	78	7	61-62	117	87
6	79	8	63	118	88
7-8	80	10	64	119	89
9	81	12	65-66	120	90
10	82	14	67	121	91
11-12	83	15	68-69	122	92
13	84	17	70	123	92
14-15	85	19	71-72	124	93
16	86	21	73	125	94
17-18	87	23	74-75	126	95
19	88	25	76	127	95
20-21	89	27	77	128	96
22	90	29	78-79	129	96
23-24	91	31	80	130	97
25	92	33	81-82	131	97
26	93	35	83	132	97
27-28	94	38	84-85	133	98
29	95	40	86	134	98
30-31	96	42	87-88	135	98
32	97	44	89	136	99
33-34	98	47	90-91	137	99
35	99	50	92	138	99
36-37	100	52	93-94	139	99
38	101	55	95	140	99
39-40	102	57	96	141	99
41	103	60	97-98	142	99
42	104	63	99	143	>99
43-44	105	65	100-101	144	>99
45-46	106	67	102	145	>99
47	107	70	103	146	>99
48	108	72	104-105	147	>99
49-50	109	74	106-107	148	>99
51	110	76	108	149	>99
52-53	111	78	109-110	150	>99
54	112	80	111-126	>150	>99
55-56	113	82			

standard scores are based on an average score of 100 with a standard deviation of 15

Functional Level
No problem. 0-15
Average: 16-54
Moderate problem: 55-75
Significant problem: 76-126

Table 2.5

Percentile Scores that Correspond with Externalizing Problems Raw Scores (Subscales B1, B2, and B3 Combined) on Scale B

Ages 3-4				Ages 5-6			
Raw score	Percent. rank	Raw score	Percent. rank	Raw score	Percent. rank	Raw score	Percent. rank
0	4	30	60	0	4	30	65
1	6	31	63	1	6	31	67
2	8	32	64	2	8	32	70
3	10	33	66	3	10	33	72
4	11	34	68	4	11	34	74
5	13	35	69	5	13	35	76
6	14	36	71	6	14	36	78
7	16	37	73	7	16	37	79
8	18	38	75	8	18	38	80
9	19	39	76	9	20	39	81
10	21	40	78	10	22	40	83
11	23	41	79	11	24	41	84
12	24	42	81	12	26	42	85
13	25	43	83	13	28	43	86
14	27	44	85	14	30	44	87
15	28	45	86	15	32	45	88
16	30	46	87	16	33	46	89
17	31	47	88	17	35	47	90
18	33	48	90	18	37	48	91
19	35	49	91	19	40	49	92
20	38	50-51	92	20	42	50	93
21	40	52-53	93	21	44	51-52	94
22	42	54-55	94	22	46	53-54	95
23	44	56-58	95	23	49	55-57	96
24	46	59-60	96	24	51	58-60	97
25	48	61-62	97	25	54	61-64	98
26	51	63-66	98	26	56	65-70	99
27	53	67-71	99	27	59	71-81	>99
28	56	74-81	>99	28	61		
29	58			29	63		

Table 2.6

Percentile Scores that Correspond with Internalizing Problems Raw Scores (Subscales B4 and B5 Combined) on Scale B

Ages 3-4		Ages 5-6	
Raw score	Percentile rank	Raw score	Percentile rank
0	4	0	6
1	7	1	9
2	11	2	13
3	14	3	16
4	18	4	20
5	22	5	25
6	24	6	29
7	28	7	34
8	32	8	38
9	36	9	42
10	39	10	47
11	45	11	52
12	49	12	57
13	53	13	61
14	59	14	67
15	65	15	72
16	70	16	76
17	74	17	81
18	80	18	83
19	83	19	86
20	87	20	89
21	89	21	91
22	91	22	92
23	92	23	94
24	94	24	95
25	95	25	96
26	95	26	97
27	96	27-28	98
28	97	29-32	99
29	97	33-45	>99
30	98		
31	98		
32-34	99		
35-45	>99		

Table 2.7

Functional Levels that Correspond to Subscale Raw Scores for PKBS Scale A, Social Skills

PKBS Score	Age	Functional level			
		High functioning	Average	Moderate deficit	Significant deficit
A1 Social Cooperation	3-4	34-36	24-33	16-23	0-15
	5-6	35-36	25-34	20-24	0-19
A2 Social Interaction	3-4	29-33	20-28	11-19	0-10
	5-6	31-33	23-30	15-22	0-14
A3 Social Independence	3-4	31-33	23-30	18-22	0-17
	5-6	32-33	26-31	20-25	0-19

Table 2.8

Functional Levels that Correspond to Subscale and Area Raw Scores for PKBS Scale B, Problem Behavior

PKBS Score	Age	Functional level			
		No problem	Average	Moderate problem	Significant problem
B1 Self-Centered/Explosive	3-4	0-3	4-17	18-24	25-33
	5-6	0-3	4-16	17-22	23-33
B2 Attention Problems/Overactive	3-4	0-3	4-14	15-18	19-24
	5-6	0-3	4-12	13-17	18-24
B3 Antisocial/Aggressive	3-4	0	1-9	10-14	15-24
	5-6	0	1-9	10-14	15-24
Externalizing Problems	3-4	0-9	10-41	42-55	56-81
	5-6	0-9	10-37	38-54	55-81
B4 Social Withdrawal	3-4	0-1	2-8	9-12	13-21
	5-6	0	1-7	8-11	12-21
B5 Anxiety/Somatic Problems	3-4	0-2	3-9	10-14	15-24
	5-6	0-2	3-9	10-13	14-24
Internalizing Problems	3-4	0-4	5-18	19-25	26-45
	5-6	0-3	4-16	17-24	25-45

the PKBS normative group, located in Tables 3.3 and 3.4 in Chapter 4. Caution is advised when using the age- and gender-specific comparative information, due to the relatively small numbers of subjects in each gender/age breakdown cell.

Interpreting PKBS Scores

Understanding and Using Functional Levels

All PKBS scores can be interpreted with the assistance of the corresponding *Functional Level* to help understand the relative meaning of a child's behavior in each score area. The Functional Levels were developed to indicate the general level of social-behavioral adjustment that is indicated by a normative comparison of the scores. Four Functional Levels are used for each of the two PKBS scales. Although the decision rules used to develop these functional levels were the same for both instruments, their wording and interpretation are somewhat different between Scale A and Scale B in order to account for conceptual and statistical differences between social skills and problem behavior. The idea behind the use of Functional Levels is that general categories reflecting different levels of social-behavioral adjustment may be useful in understanding how well children are actually able to function in terms of meeting the day-to-day demands of peer relationships and interactions with significant adults.

Functional Levels for Scale A

The four Functional Levels used for interpreting Scale A (Social Skills) scores include *High Functioning, Average, Moderate Deficit,* and *Significant Deficit*. These four Functional Levels for Scale A are explained as follows:

The High Functioning level is used to indicate Social Skills scores that are higher than 80% of the norm group scores. Children whose PKBS scores are in the High Functioning Level have developed excellent social skills and tend to be well-liked by both peers and adults.

The Average level includes Social Skills scores that range from approximately the 80th to 20th percentile levels of the norm group. Scores in this range reflect the behavioral adjustment of typical children; they have probably developed adequate social skills and use them effectively in most situations, but may occasionally experience peer adjustment difficulties that are not severe in nature.

The Moderate Deficit level includes Social Skills scores that range from approximately the 20th to 5th percentile levels. Children with scores at this level are typically close to or somewhat more than one standard deviation below the normative mean in their observed social skills. They are usually good candidates for more comprehensive assessment of their social skills, and may benefit from a specially designed, deficit-matched social skills intervention.

The Significant Deficit level includes Social Skills scores that are in the lowest 5% of the normative group. Children whose scores are in this range may be likely to exhibit social skills deficits and peer relationship problems that are considered to be severe, especially when their Social Skills total score or more than one Social Skills subscale score are at this level. Children whose Social Skills scores are consistently at the Significant Deficit level are generally in need of extensive social skills training that is matched specifically to their personal areas of deficit. They may need careful monitoring and prompting in their peer interactions in order to facilitate friendship making and to avoid peer rejection. The lowest 5% level was selected for inclusion in the Significant Deficit level because recent conservative prevalence estimates indicate that approximately 3% to 6% of children exhibit social skills deficits to a great enough extent that special education or other specific behavioral interventions are desirable for remediation of these deficits and to prevent the occurrence of more serious problems (Cullinan, Epstein, & Kauffman, 1984; Kauffman, 1992).

Functional Levels for Scale B

The four Functional Levels used for interpreting Scale B (Problem Behavior) are *No Problem*, *Average*, *Moderate Problem*, and *Significant Problem*. These four levels are explained as follows:

The No Problem level is used to indicate problem behavior scores that are at or below the 20th percentile level of the norm group. Children whose PKBS Problem Behavior scores are rated at this level exhibit significantly fewer problem behaviors than most other children. They are unlikely to get into trouble because of their behavior. Although many, if not most, children whose Problem Behavior scores are at the No Problem level are likely to have also developed good to excellent social skills, such a connection is not always evident. Some children who exhibit very few problem behaviors may also be shy, withdrawn, and somewhat socially isolated.

The Average level includes Problem Behavior scores that range from approximately the 20th to 80th percentile levels. Children whose PKBS Problem Behaviors are rated in this range might be considered typical. They may occasionally exhibit either internalizing or externalizing problem behaviors, but seldom to the point where they are likely to cause distress to themselves or others.

The Moderate Problem level includes Problem Behavior scores that range from the 80th to 95th percentile levels when compared to the same-age norm group. Students whose problem behaviors are rated at this level are close to or somewhat higher than one standard deviation above the normative mean in terms of their problem behaviors. They may be in trouble on more than an occasional basis due to engaging in antisocial behaviors or, alternatively, they may be likely to cause a fair amount of concern to their parents, teachers, or caregivers due to their internalizing symptoms. Children whose PKBS Problem Behavior scores are at this level may be good candidates for more comprehensive assessments, and may also benefit from a problem-matched behavioral-emotional treatment program.

The Significant Problem level includes scores that are in the highest 5% of Problem Behavior scores in comparison to the norm group (95th percentile and higher). Children whose ratings are consistently at this level are likely to exhibit severe problem behavior symptoms, especially when the Problem Behavior Total score or more than two of the five subscale scores are rated at this level. These children tend to be frequently in trouble due to acting-out behavior, or they may cause considerable distress to adults due to socially withdrawn, depressive, or internalizing features. Alternatively, some children whose Problem Behavior ratings are at the Significant Problem level may exhibit a mixed variety of externalizing and internalizing symptoms. Children whose problem behaviors are at this level will undoubtedly be the focus of serious concern and are likely to require additional assessment and a carefully designed, problem-matched intervention. The highest 5% of scores were selected for inclusion in the Significant Problem level because of prevalence estimates indicating that approximately 3% to 6% of children exhibit behavioral problems to a great enough extent that special education or other specific behavioral interventions are desirable for remediation of these deficits and to prevent the occurrence of more serious problem behaviors or behavioral disorders (Cullinan et al., 1984; Kauffman, 1992).

Differences Between Subscale, Area, and Total Scores

Functional Levels are the only interpretive aid provided for subscale scores, other than consulting gender by age raw score descriptive statistics from the standardization sample. Standard scores and percentile ranks are not included for subscales, as the distribution of subscale scores for the standardization sample was more skewed (negatively skewed for Social Skills scores and positively skewed for Problem Behavior scores) than the distribution of total scores. As a result, conversion of subscale scores into standard scores could easily lead to faulty conclusions about the meaning of the score. The two area scores on Scale B (Externalizing Problems and Internalizing Problems) are converted to percentile ranks as well as Functional Levels, but as in the case of the subscale scores, standard score conversions are not made due to the problem of a restricted range of scores and relatively skewed score distributions. The four Functional Levels for each of the scales serve the purpose of providing a general framework around a score whereby hypotheses can be generated and interpretation assisted.

The total scores for Scale A and Scale B not only are converted into Functional Levels, but are also transformed into standard scores (with a normative mean of 100 and standard deviation of 15) and corresponding percentile ranks. Professionals working in developmental preschool settings are often familiar with this standard score system, as it is widely used on standardized, norm-referenced intellectual ability and developmental tests, as well as some other behavior rating scales. When using these standard scores for interpretive purposes, it should always be remembered that directionality of the scores on Scale A and Scale B mean different things. Higher scores on Scale A indicate greater levels of social skills, whereas higher scores on Scale B indicate perceptions of greater levels of problem behavior.

Because the total scores for Scale A and Scale B are aggregated measures consisting of the sum of their respective subscales, more importance should be placed on the total scores than on individual subscale scores in interpretation. The subscale scores consist of between 7 and 12 items each, and although they can provide useful information about specific aspects of social skills or certain clusters of behavior problems, they should be used more cautiously than the total scores. The two area scores (Externalizing Problems and Internalizing Problems) on Scale B serve an intermediate purpose in this regard; although they are not converted into standard scores, they are aggregated measures of specific behavioral or emotional problem clusters, and thus have greater reliability and discriminating power than do the subscales from which they are comprised. The Functional Levels and percentile scores of these two Problem Behavior area scores can be used to generate hypotheses regarding diagnosis, classification, and intervention planning.

Use of PKBS Scores in Screening and Assessment

As was inferred in the discussion of Functional Levels, it is recommended that certain score levels be used as benchmarks for screening and assessment purposes. For Scale A and Scale B total scores, it is recommended that standard scores of approximately one to one and one-half standard deviation in the least desired direction from the normative mean be used as the cutoff point for screening for additional assessment or social-behavioral interventions. For Scale A, this cutoff level would refer to scores one to one and one-half standard deviations or more below the normative mean (85 to 78 or less), whereas for Scale B, this level would refer to scores approximately one to one and one-half standard deviations above the mean (115 to 122 or greater). Scores in this range will generally correspond with either the Moderate Problem/Moderate Deficit or Significant Problem/Significant Deficit Functional Levels. The use of a one to one and one-half standard deviation screening cutoff will result in some "false-positive" errors (identification of some children who do not have significant social-emotional behavior problems) that can be easily eliminated with further assessment, but will result in very few or no "false-negative" errors (failure to identify children who have significant social-emotional behavior problems). For this same purpose, one should also consider including children whose PKBS total scores do not quite reach the one to one and one-half standard deviation criteria, but whose subscale scores are at the moderate to significant problem or deficit Functional Levels on two or more subscales on Scale A, or three or more subscales on Scale B.

For purposes of formal assessment, it is recommended that children whose total scores on *both* Scale A and Scale B are at the *Significant Problem* Functional Level be seriously considered for placement in special education programs or other means of delivering specially designed social-behavioral interventions. However, this recommendation should be followed *only when PKBS scores at this level are corroborated by other forms of behavioral assessment, including direct observational data*. The preferred method for conducting behavioral assessments is to use a multimethod, multisource, multirater design, of which PKBS scores would be just one part. Failure to use an assessment approach of sufficient breadth may result in faulty findings related to source, instrument, and setting variance (Martin, Hooper, & Snow, 1986).

Issues in Using Behavior Rating Scales

General Characteristics of Rating Scales

Behavior rating scales such as the PKBS provide a standardized format for the development of summative judgments about child behavioral characteristics, by an informant who knows the child well. The informant is usually a parent or teacher, but in the case of the PKBS, a daycare provider, preschool aide, grandparent, or social services caseworker might also be the informant.

As an assessment methodology, behavior rating scales are less direct than either direct behavioral observation or structured behavioral interviewing (McMahon, 1984), in that they measure *perceptions* of specified behaviors rather than providing a firsthand measure of the existence of the behavior. However, rating scales are an objective method, and yield more reliable data than either unstructured clinical interviewing or projective techniques (Martin et al., 1986). As behavior rating scales became more widely used during the 1970's, they were typically viewed with suspicion and used as a "last resort" by behaviorally oriented clinicians (Cone & Hawkins, 1977), but as the research base and technological refinements in rating scales have become more advanced, there appears to be a more broad acceptance of their use. Conners and Werry (1979) defined rating scales as an "algebraic summation, over variable periods of time and numbers of social situations, of many discrete observations" (p. 341). The term *algebraic* indicates that for each rating scale item, various rating choices are available, each of which symbolizes a particular level of behavior. An additive checklist, on the other hand, is a list of symptoms or characteristics that the rater checks if present, and which yields a total score for the number of items checked. In general, the algebraic format provided by rating scales is preferred to the simple additive format provided by checklists, because it allows for more precise measurement of behavioral frequency or intensity (Merrell, 1993a).

Advantages of Using Behavior Rating Scales

The widespread popularity of using behavior rating scales is not incidental—they offer many advantages for clinicians and researchers conducting assessments. The main advantages of behavior rating scales are summarized in the following six points:

1. When compared with direct behavioral observation, behavior rating scales are less expensive in terms of professional time involved, and amount of training required to utilize the assessment system (Merrell, 1993a).

2. Behavior rating scales are capable of providing data on low frequency but important behaviors that might not be seen in a limited number of direct observation sessions (Sattler, 1988). An example that serves to illustrate this point is a physical attack by one child toward another. In most cases, this type of behavior does not occur on a constant or consistent schedule, and might be missed within the constraints of conducting two brief observations. Nonetheless, it is extremely important to know about the occurrence of such behaviors.

3. As mentioned earlier in this chapter, behavior rating scales are an objective assessment method that provide more reliable data than do unstructured interviews or projective techniques (Martin et al., 1986).

4. Behavior rating scales can be used to assess subjects who cannot readily provide information about themselves (Merrell, 1993a). With preschool-age children, this advantage is particularly important; most preschool-age children lack the insight, cognitive maturity, and verbal mediation skills to observe and describe their own behavior accurately. Thus, the evaluator must rely on objective descriptions by others.

5. Rating scales capitalize on observations over a period of time in a child or adolescent's "natural" environment (i.e., preschool or home settings) (McMahon, 1984).

6. Rating scales capitalize on the judgments and observations of persons who are highly familiar with the child or adolescent's behavior, such as parents or teachers, and who are thus considered to be "expert" informants (Martin et al., 1986).

It is easy to see from these six illustrations of the advantages of using rating scales why these scales are widely used—they get at the "big picture" of the assessment problem in a short amount of time, at moderate cost, and with a good deal of face and clinical validity.

Problems Associated with Using Behavior Rating Scales

Despite the many advantages offered by rating scales, there are some catches, as one might suspect. The most sophisticated rating scales available can help provide objective, reliable, and socially valid information on both broad and narrow dimensions of behavioral, social, and emotional problems, but the nature of rating scale technology contains several potential flaws, which are important to understand. At the onset of discussing problems associated with behavior rating scales, it is useful to remember that by their nature (i.e., assessing *perceptions* of problems), rating scales provide *idiographic* rather than *nomothetic* information—they are capable of providing a portrait of a general idea or conception of behavior, but do not provide actual observational data.

Martin et al. (1986) categorized the measurement problems of behavior rating scales into two classes: *bias of response* and *error variance*. Bias of response refers to the way that informants completing the rating scales may potentially create additional error by the way they use the scales. There are three specific types of response bias, including (a) *halo effects* (rating a child in a positive or negative manner because he or she possesses some other positive or negative characteristic not pertinent to the rated item), (b) *leniency or severity* (the tendency of some raters to have an overly generous or overly critical response set when rating all subjects), and (c) *central tendency effects* (the tendency of raters to select midpoint ratings and to avoid endpoints of the scale such as "never" or "always").

Error variance is closely related to and often overlaps with response bias but provides a more general representation of some of the problems encountered with this form of assessment. According to Martin et al. (1986), there are four different types of variance that may create error in the obtained results of a rating scale assessment, which are overviewed as follows. *Source variance* refers to the subjectivity of the rater, and any of the idiosyncratic ways that he or she completes the rating scales. *Setting variance* occurs as a result of the situational specificity of behavior (Kazdin, 1979), in that humans tend to behave differently in different environments due to the different eliciting and reinforcing properties present. *Temporal variance* refers to the tendency of behavior ratings to be only moderately consistent over time—partly due to changes in the observed behavior over time, and partly due to changes in the rater's approach to the rating task over time. Finally, *instrument variance* refers to the fact that different rating scales measure often related but slightly differing hypothetical constructs (e.g., aggressive behavior vs. delinquent behavior), and a severe problem behavior score on one scale may be compared with only a moderate problem behavior score on a differing rating scale for the same person. Another problem that creates instrument variance is the fact that each rating scale utilizes different normative populations with which to make score comparisons, and if the norm populations are not randomly selected and representative of the population as a whole, similar scores on two different rating scales may not mean the same thing.

Best Practices in Using Behavior Rating Scales

Although there are several types of problems inherent in the use of behavior rating scales, there are also effective ways of minimizing those problems, several of which will now be described. In order to minimize some of the problems associated with the use of rating scales such as the PKBS, three specific "best practices" suggestions are offered.

The first suggestion is to *use rating scales routinely for early screening*. Effective screening practices involve being able to identify systematically and with a high degree of accuracy children who are in the early stages of developing behavioral, social, or emotional problems. The identified subjects are then evaluated more carefully to determine whether their social-behavioral problems warrant special program eligibility and intervention services. The purpose of screening for social-behavioral problems is usually for *secondary prevention*, which is the prevention of the existing problem becoming worse (Kauffman, 1989). Screening for early intervention is one of the best uses of behavior rating scales, as they cover a wide variety of important behaviors and take very little time to administer and score. Given that the PKBS is designed to be used with the early childhood/preschool population, the screening for early identification use of rating scales is particularly important.

The second suggestion offered is to *use the "aggregation principle."* This principle involves obtaining ratings from a variety of sources, each of which might present a slightly differing picture. When using rating scales for purposes other than routine screening, obtaining aggregated rating scale data is recommended in order to reduce bias of response and variance problems in the assessment. In practice, using aggregated measures means to obtain rating evaluations from different raters in different settings, and to use more than one type of rating scale to accomplish this (Martin et al., 1986).

A final suggestion is to *use behavior rating scales to assess progress during and after interventions*. It has been demonstrated that continuous assessment and monitoring of student progress following the initial assessment and intervention is very important in successful implementation of behavioral interventions (Kerr & Nelson, 1989). Progress toward behavioral intervention goals can be assessed easily on a weekly or bi-weekly schedule using appropriate rating scales. Though rating scales may not be the best measurement choice for daily assessment data, there are a number of other simple ways of assessing progress daily, such as using performance records or brief observational data. Additional assessment following the intervention can also be a useful process. The main reason for follow-up assessment is to determine how well the intervention effects have been maintained over time (e.g., after 3 months), and how well the behavioral changes have generalized to other settings (e.g., the home setting and other classrooms). In actual practice, a follow-up assessment might involve having teacher(s) and parent(s) complete behavior rating scales on a child after a specified time period has elapsed following the child's participation in a social skills training program. The data obtained from this follow-up assessment can be used to determine whether or not follow-up interventions seem appropriate and may be useful in developing future intervention programs if it is determined that social-behavioral gains are not being maintained over time or generalized across specific settings.

Linking Assessment to Intervention

One of the primary purposes of conducting psychological and educational assessments is to obtain information that will be useful for developing interventions to ameliorate problems that are identified through the assessment process. With that end in mind, the purpose of this section of the chapter is to provide some additional information and suggestions that may be helpful in linking the PKBS to the process of intervention planning. This section is not designed to be a comprehensive intervention manual—such depth is beyond the scope of this test manual, and test users are referred to excellent books by Barnett and Casey (1992) and Odom, McConnell, and McEvoy (1992) for in-depth discussions of social skills and problem behavior interventions for young children. What is provided in this section of the chapter is information relating to general intervention considerations and practices with the PKBS.

There are three general ways in which the PKBS may be useful as a link to intervention planning. The first way simply involves selecting children whose PKBS scores show moderate to significant deficits or problems, and who are thus prime candidates for intervention. The PKBS Functional Level data and the screening procedures that were recommended in Chapter 2 both will be useful in this regard.

The second way that PKBS scores may serve as a useful link in intervention planning is in identifying *specific clusters* of behavioral problems of skills deficits that need attention. Rather than simply identifying children whose global PKBS scores are indicative of moderate to severe problems or deficits, and then prescribing a global social skills training or behavioral problem reduction intervention, a more desirable approach is to make the intervention *problem specific* (Peacock Hill Working Group, 1991). In other words, the intervention should closely match the problem rather than be generic in nature. In many cases, the PKBS may be useful for developing problem-matched interventions by analyzing the subscale and area score levels. For example, the PKBS ratings for a given child may suggest an overall high level of problem behaviors, but an analysis of the subscale and area scores within Scale B may indicate that the problems do not occur "across the board," but appear in related clusters. Given the conceptual differences between internalizing and externalizing dimensions of problem behavior, it would not be unusual for a highly aggressive, acting-out child to have significant score elevations on subscales B1, B2, and B3, but to be rated within normal limits on the internalizing subscales B4 and B5. Thus, a careful analysis of the subscale and area scores on Scale B could be used to identify specific rather than general targets for an intervention plan. Likewise, the social skills ratings from PKBS Scale A may be used in the same problem-specific manner. Although a number of children may exhibit global social skills deficits, it is quite likely that some children will be rated as having adequate competence in one social skills area, and having a deficit in another area. For example, a child may be quite competent at being socially independent (as evidenced by a high score on subscale A3), but have continual problems cooperating with peers (as evidenced by a low score on subscale A1). Thus, a careful inspection of PKBS score profile patterns may yield information that is useful for creating specific problem-matched interventions, which are typically a better practice than using generic or global intervention strategies.

The third way in which the PKBS may be an effective link to intervention is by using the specific wording of scale items to develop Individual Education Plan (I.E.P.) goal statements for children who have been assessed using the PKBS and identified as requiring special services. The items from Scale A may be incorporated directly into an I.E.P. For example, if inspection of the PKBS ratings of a given student reveals that he/she consistently receives ratings of "0" (never) on item A25 "Shares toys and other belongings," the I.E.P. goal statement for a hypothetical student named Anne might read "Anne will share toys with other students when appropriate," followed by the description of an appropriate level for the goal and an observation method. For the items on Scale B (Problem Behavior), the wording can be rephrased slightly to reflect appropriate goal statements. For example, if a hypothetical student named Jaime consistently receives ratings of "3" (often) on item B37, "Has difficulty concentrating or staying on task," the I.E.P. goal statement might read "Jaime will stay on task during structured preschool activities," with an appropriate percentage of on task behavior and an observation method specified.

In sum, one of the most important potential uses of the PKBS is for developing intervention goals and plans for children who have been identified as having problem behavior excesses and/or social skills deficits. Three practices for linking PKBS assessment data to intervention have been identified, and test users are encouraged to employ these practices in conjunction with effective intervention strategies for teaching social skills and reducing behavior problems.

3

DEVELOPMENT AND STANDARDIZATION

Within this chapter information on the development of the PKBS is presented. This information includes instrument development and refinement, data collection procedures, and various characteristics of the normative sample (e.g., geographic locale, race/ethnicity, socioeconomic status, setting type, age/gender).

Item and Instrument Development

Items were initially developed for inclusion in the PKBS using what Lanyon and Goodstein (1982) have referred to as a *rational-theoretical* approach to test construction. Within this general approach, both *intuitive* and *content validation* methods were incorporated (Martin, 1988). The first step involved in the item development process was to select specific behavioral domains for which scale items would be representative. To this end, two different behavioral domains were selected for each scale. For the Social Skills scale (Scale A), a decision was made to include items that would represent the domains of *peer-related* and *adult-related* social adjustment. This breakdown of social adjustment into two separate but related domains was loosely based on the previously discussed work by Walker et al. (1985), who identified two major types of social adjustment that children must make upon entering school, namely *peer related* and *teacher related*. This dichotomous breakdown of types of social adjustment was also used previously by the author of the PKBS in the development of a behavior rating scale for use with elementary- and secondary-aged students, the School Social Behavior Scales (Merrell, 1993b). Given that there is more variation in types of formal education and other structured experiences for preschool-age than school-age children, the second behavioral domain is perhaps better thought of and referred to as *adult related* rather than *teacher related*. Peer-related social adjustment involves the social and behavioral dynamics that occur between children in unstructured and free-play settings (e.g., engaging in play activities with other children and sharing toys), whereas adult-related social adjustment involves meeting the social and behavioral expectations and demands of adult caregivers such as parents, daycare providers, and preschool teachers (e.g., following rules and cooperating).

For the Problem Behavior scale (Scale B), a decision was made to include the domains of *internalizing* and *externalizing* problems. As has already been discussed, this breakdown is grounded in the behavioral dimensions approach to classifying problem behaviors (Kauffman, 1989), which is based on the pioneering work of Achenbach (1985) and Quay (1986). Internalizing problems include *overcontrolled* or *self-directed* syndromes such as anxiety, depression, social withdrawal, and somatic complaints. Externalizing problems include *undercontrolled* or *other-directed* behavioral syndromes such as aggression, delinquent behavior, hyperactivity, antisocial behavior, and other forms of disruptive behavior problems.

To select potential items representing the appropriate behavioral domains for the two scales, a list of behavioral descriptors was compiled following an extensive search of the literature of normal and abnormal child development, paying particular attention to the 3- through 6-year age ranges. The literature reviewed for the selection of behavioral descriptions included developmental psychology textbooks, edited books on child psychopathology and social competence, and numerous research articles from child development and abnormal child psychology journals. At the conclusion of this literature search, 131 behavioral descriptions were selected for potential use in the Social Skills scale, including 75 descriptions of peer-related behavior, and 56 descriptions of adult-related behavior. For the Problem Behavior scale, 110 behavioral descriptions were included initially, with 59 descriptions of internalizing problems and 51 descriptions of externalizing problems. These lists were then narrowed down by the author to 50 descriptions for the Social Skills scale (25 each in the peer-related and adult-related domains) and 53 descriptions for the Problem Behavior scale (27 internalizing and 26 externalizing).

These remaining 103 items were then developed into scale item format and incorporated into a questionnaire for evaluation by early childhood professionals in a formal content validation procedure. The items were evaluated by a panel of 16 different early childhood professionals, including preschool teachers and directors, daycare operators, a pediatric nurse, and a pediatrician. These individuals were asked to rate each item on how well the item described characteristic early childhood behaviors. The professionals were also encouraged to make comments on the structure of items. Panel members had an average of 10.3 years' experience working directly with preschool-age children in a professional capacity, and an average of 5.03 years' of post-high school education in child development, early childhood education, and related fields. The items were rated for how well each item described early childhood behavior using a 5-point scale ("very poor" to "excellent").

After the item rating data were compiled from the expert content validation procedure, the author of the PKBS developed a final item format by deleting items that were rated as less than "fair" by the panelists, or being otherwise undesirable for inclusion in the scale (e.g., lack of item specificity, awkward wording, or confusing content). Some items were also reworded slightly after the professional review. The item tryout version of the scale included 35 items on the Social Competence scale (18 representing peer-related and 17 representing adult-related social adjustment) and 45 items on the Problem Behavior scale (23 representing internalizing problems and 22 representing externalizing problems).

The final rating format that was developed for the PKBS was a 4-point rating scale, which is described as follows:

0 = *Never* true
1 = *Rarely* true
2 = *Sometimes* true
3 = *Often* true

Although this particular rating format was developed specifically for the PKBS, it is quite similar to the rating scale formats utilized in some other widely used and researched child behavior rating scales, such as the Conners Parent and Teacher Rating Scales (Conners, 1990), the Walker-McConnell Scale of Social Competence and School Adjustment (Walker & McConnell, 1988), and the Social Skills Rating System (Gresham & Elliott, 1990). Several other slightly modified rating formats were experimented with during the initial development of the PKBS, but were rejected for several reasons, including a lack of specificity, restricted range, or overly complicated rating procedures.

Following the collection of normative data and after the factor analytic work (which is discussed in Chapter 4) was complete, one item was dropped from Scale A (leaving a final total of 34 items), and three items were dropped from Scale B (leaving a final total of 42 items). These four items were dropped from the final version of the test because the factor analysis procedures produced evidence suggesting that they did not fit statistically with the other items, in that they lacked specificity and had weak factor loadings. The final subscale structure was based on the results of the factor analyses

that were conducted, which are described in Chapter 4. The items in the final version of the PKBS are presented by subscale area in Tables 1.1 and 1.2 in Chapter 1.

Data Collection Procedures

Normative PKBS ratings were obtained for 2,855 preschool- and kindergarten-aged children during 1992 and 1993. The entire standardization sample was used in developing the PKBS subscale structure through factor analysis procedures, but a small number of ratings in the sample were not included when developing the tables for PKBS score norms and PKBS score conversions, due to the fact that either age or gender information was incomplete or out of range. More information on the final normative sample used is presented later in this chapter.

Participants for obtaining the PKBS norm sample were recruited through contacting (a) research and early childhood education personnel from a number of public school districts, (b) proprietors of private preschools, (c) administrators or coordinators of public preschools (e.g., Head Start programs), and (d) administrators of pediatric medical clinics. Contacts with potential research sites were made through a combination of direct in-person communications and mailed invitations to participate accompanied by a research prospectus.

Agencies and individuals who indicated an interest in participating in the PKBS research project were provided with data collection packets, including specified numbers of research protocols, detailed instructions for data collection coordinators, instructions for individuals completing the rating forms, and parent consent forms. Because one of the major research goals was to obtain PKBS ratings provided by both early childhood professionals and parents, data collection coordinators from participating sites were provided with specific instructions for obtaining parent involvement, details of how to match parent and professional ratings of the same child, and information on coding the data collection forms in a manner that would ensure that rated children and their parents would not be personally identifiable.

Characteristics of Communities Represented in the Norm Sample

The norm sample consisted of subjects from 24 different U.S. communities, which included representation from each of the four U.S. geographical regions and a total of 16 states. Table 3.1 includes the names of specific communities and states represented in the norm sample, divided by geographic region. As the information in this table indicates, the communities represented in the PKBS norm sample are from diverse locales and diverse circumstances, with a substantial mix of urban, suburban, and rural areas included.

Table 3.1

Communities Represented in the PKBS Standardization Sample, Listed by U.S. Geographic Regions

West	North Central	Northeast	South
Covina, CA	Adams County, OH	Buffalo, NY	Bay St. Louis, MS
Las Vegas, NV	Belleville, IL	Hohnsdale, PA	Caldwell County, NC
Logan, UT	Cedar Rapids, IA	Newark, NJ	Collinsville, VA
Ogden, UT	Chippewa Falls, WI	Shelocta, PA	Georgetown, TX
	Coralville, IA	York, PA	Poplarville, MS
	Duluth, MN		Raleigh, NC
	Evanston, IL		Waveland, MS
	Terre Haute, IN		

Gender

Of the 2,855 children whose ratings comprised the PKBS norm sample, 1,484 were male, 1,323 were female, and 48 were not identified by gender on the research forms. In order to assess the overall influence of child gender on PKBS scores, a series of point-biserial correlations were computed using gender as a dichotomous independent variable, and PKBS subscale, area, and total raw scores as dependent variables. The overall range of obtained correlation coefficients (which are presented in Table 3.2) was relatively weak, though in many cases statistically significant at the $p < .001$ level. The correlation between gender and the Scale A (Social Skills) total score was .13, whereas the correlation between gender and the Scale B (Problem Behavior) total score was −.14. The general direction and magnitude of these correlations suggests that female gender is weakly associated with higher Social Skills scores, whereas male gender is weakly associated with higher Problem Behavior scores. It is interesting to note that the PKBS score with the strongest correlation to gender was Problem Behavior subscale B2 (Attention Problems/Overactive), which appears to be consistent with a significant body of literature suggesting that the symptoms of Attention-Deficit Hyperactivity Disorder tend to be more strongly associated with male gender.

Based on the overall results of these correlations, a decision was made to not use separate PKBS score norms based on gender as a general rule, given the relatively weak pattern of correlations. The decision to not develop separate PKBS score norms for boys and girls was also made for a more practical reason. Although gender may subtly influence the development and rating of social behavior, expectations for appropriate behavior in early childhood education settings are presumably based on how compatible behaviors are with the overall structure and rules of a setting, rather than how compatible these behaviors might be with the "norm" for boys or girls. Although the PKBS score conversion tables are not based on a male-female breakdown, separate age-by-gender PKBS descriptive statistics may be consulted for comparison of individual cases. These tables include Tables 3.3 and Table 3.4 in this chapter.

Age

The targeted age group for the PKBS was the early childhood/preschool and kindergarten age population. Thus, children between the ages of 3 and 6 were targeted for the development of normative score data. The large majority of children in the standardization sample fell into this age range, but a handful (41) of subjects rated were either 2 or 7 years of age. These out-of-range subjects were not included in the final age-based PKBS score norms, although their completed ratings were used in the development of the PKBS structure. Tables 3.3 and 3.4 provide details of the number of subjects in each age group.

In order to assess the influence of subject age on PKBS scores, a number of procedures were conducted. First, Pearson product-moment correlations were calculated between subjects' ages and PKBS scores. These correlations, which are presented in Table 3.2, showed a somewhat different pattern than the point-biserial correlations that were obtained between gender and PKBS scores. Correlations between age and PKBS scores were stronger for the Scale A (Social Skills) scores than they were for Scale B (Problem Behavior) scores. An r value of .20 ($p < .001$) was obtained for the correlation between age and the Scale A total score, whereas an r value of only −.06 was obtained for the correlation between age and the Scale B total score. These correlational data appear to indicate that age is correlated modestly but positively with the development of social skills, whereas age has a negative, though much weaker relationship with the exhibition of various behavioral and emotional problems. Thus, social skills and problem behaviors appear to be developmentally related constructs, with children normally developing better social skills and somewhat fewer problem behaviors as they get older.

Because the correlations between age and PKBS scores were higher than the correlations obtained for gender, and because subsequent analysis of variance and t-test procedures (which are discussed in Chapter 4) found that age-related effect sizes for PKBS scores were considerable in some cases, a

Table 3.2

Correlations of PKBS Scores with Gender and Age for the Standardization Sample (N = 2,855)

PKBS score	Gender	Age
A1 Social Cooperation	.11**	.13**
A2 Social Interaction	.14**	.21**
A3 Social Independence	.09**	.18**
AT SOCIAL SKILLS TOTAL	.13**	.20**
B1 Self-Centered/Explosive	-.08**	-.05*
B2 Attention Problems/Overactive	-.19**	-.07**
B3 Antisocial/Aggressive	-.17**	-.00
Externalizing Problems	-.15**	-.05*
B4 Social Withdrawal	-.10**	-.14**
B5 Anxiety/Somatic Problems	-.02	-.03
Internalizing Problems	-.07**	-.09**
BT PROBLEM BEHAVIOR TOTAL	-.14**	-.06*

*$p < .01$; **$p < .001$.

decision was made to develop separate norm tables based on the age of subjects. After experimenting with several different age divisions for norm tables, a final decision was reached to have separate norm tables for two different age divisions: children ages 3 and 4, and children ages 5 and 6. The PKBS score conversion tables in Chapter 2 are thus based on this age division. Using age as a unit of normative analysis in assessing children's social skills and problem behavior is a matter that was not only statistically indicated, but makes practical comm ... sense. As children progress from one developmental level (e.g., age, year) to another, the expectations that their behaviors are compared against tend to change, with family and preschool rules normally following these expectations. In most cases, the behavioral expectations for 6-year-old children are considerably higher than for 3-year-old children.

Although the PKBS score conversion procedure described in Chapter 2 employs separate norms for two different age groups (3-4 and 5-6), the gender-by-age descriptive statistics for the PKBS that are found in Table 3.3 and Table 3.4 make possible the comparison of a particular child's score with a specific age/year group. Similar to the case of gender-specific comparisons, it is recommended that these supplementary gender and age tables be used for occasional comparison purposes rather than as usual procedure, because the number of subjects within each gender by age cell is much smaller, and thus less reliable than the combined group normative data.

Race and Ethnicity

During the PKBS standardization process, raters were asked to identify the racial or ethnic group of the children who were rated. Table 3.5 presents the racial distribution of the 2,855 children whose

Table 3.3

PKBS Raw Score Means and Standard Deviations for Males, by Age Level

PKBS score	Age level							
	3 (n = 110)		4 (n = 258)		5 (n = 508)		6 (n = 578)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
A1	25.61	6.15	28.33	5.86	28.48	6.31	29.26	5.12
A2	21.67	7.31	23.77	6.21	24.28	6.46	26.48	5.26
A3	25.54	4.96	26.83	5.17	27.38	4.97	28.70	3.96
AT	72.82	16.05	78.93	14.78	80.14	15.66	84.44	12.33
B1	14.33	7.52	10.76	7.85	11.38	7.99	11.30	6.92
B2	11.83	5.22	9.43	5.80	9.90	6.15	9.54	5.18
B3	7.22	5.11	6.05	5.07	6.70	5.35	6.75	5.04
EXT	33.38	16.14	26.24	17.47	27.98	18.26	27.60	15.81
B4	6.60	4.07	5.91	4.10	5.60	4.23	4.83	3.75
B5	7.91	4.78	6.24	4.57	6.69	4.51	6.82	4.05
INT	14.40	8.10	12.15	8.02	12.29	7.87	11.64	7.04
BT	47.88	21.97	38.40	23.80	40.27	24.24	39.24	20.86

ratings comprised the PKBS normative group. Based on a comparison with U.S. Bureau of the Census data (1990), the PKBS normative sample is slightly overrepresentative of Whites (85.2%, as compared to 80.3% nationally), is exactly representative of the nationwide percentage of Blacks (12.1%), and is modestly underrepresentative of other U.S. racial groups (see Table 3.5). In terms of the Hispanic population, the Bureau of the Census does not list Hispanic as a separate racial group, and normally includes it under the White racial category as a specific ethnic group. In the PKBS standardization sample, 5.2% of the subjects were identified as Hispanic. This figure is somewhat less than the nationwide figure of 9%, based on additional analyses supplied by the Census Bureau's Racial Statistics division.

To evaluate the effect of subject race on PKBS ratings, the racial data from the normative sample were recoded into two general categories (White and all non-White minority groups combined), and Pearson product-moment correlations were calculated between race and PKBS scores. The resulting correlation coefficients were quite small, ranging from .01 to −.10 for the PKBS subscale and area scores, and with the r values for the total scale scores being −.09 for Scale A (Social Skills) and .03 for Scale B, respectively. These results are consistent with previous findings (e.g., Achenbach & Edelbrock, 1981; Merrell, 1993b), which indicate that in most cases, race or ethnicity tends to not be a critical factor in influencing the specific direction of behavioral ratings, particularly when the effects of socioeconomic status are taken into account.

Table 3.4

PKBS Raw Score Means and Standard Deviations for Females, by Age Level

PKBS score	Age level							
	3 (n = 89)		4 (n = 219)		5 (n = 507)		6 (n = 493)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
A1	26.40	7.23	29.56	5.81	30.28	5.29	30.31	4.62
A2	22.71	6.71	25.32	5.55	26.68	5.45	27.81	4.80
A3	25.63	4.67	27.91	4.18	28.52	4.29	29.14	3.68
AT	74.75	15.85	82.79	13.46	85.49	13.19	87.26	11.24
B1	12.92	8.26	10.05	7.11	9.38	6.94	10.44	6.75
B2	9.89	5.71	7.68	5.03	7.34	5.18	7.46	4.89
B3	6.05	5.38	4.60	4.25	4.55	4.31	4.95	4.18
EXT	28.85	17.91	22.34	15.35	21.27	15.39	22.85	14.63
B4	6.46	4.00	5.27	3.73	4.45	3.92	4.11	3.59
B5	7.57	4.02	6.16	3.94	6.50	4.29	6.76	4.24
INT	14.03	7.17	11.43	6.99	10.95	7.49	10.87	7.09
BT	42.89	22.80	33.77	20.79	32.22	21.25	33.72	20.06

Disability Condition

Of the 2,855 subjects in the PKBS normative group, 2,542 (89%) were not identified as having a disability, 239 (8.4%) were identified as developmentally delayed, and 73 (2.6%) were in the process of being assessed for a developmental disability or delay at the time the rating was completed. Significant differences in PKBS scores were found between the identified/referred children and the non-disability group, and these differences are discussed in more detail in Chapter 4. The overall percentage of children in the normative group who were identified as having a disability is very similar to the national rate of preschool- and kindergarten-aged children who have been identified has having disabilities (U.S. Department of Education, 1990).

Socioeconomic Status

One of the demographic items requested on the PKBS research forms was occupation of subjects' parents, which was included as an indicator of socioeconomic status (SES). Parent occupation was utilized as the primary SES indicator because (a) it was more readily available in completing the research forms than other SES indicators such as yearly income or educational level, and (b) a significant amount of research has suggested that occupational category is strongly linked to social class and socialization (Kohn & Schooler, 1973, 1978, 1983). Usable parent occupational information was available for 82% of the 2,855 subjects in the PKBS standardization sample. The parent occupation information for these

Table 3.5

Racial Distribution of the PKBS Standardization Sample with Comparative Information from the General U.S. Population

Racial group	Percent in PKBS standardization sample	Percent in U.S. population*
White	85.2%	80.3%
Black	12.1%	12.1%
American Indian, Eskimo, or Aleut	.01%	.8%
Asian or Pacific Islander	1.5%	2.9%
Other Race	1.2%	3.9%

*SOURCE: *Decennial Census Summary, 1990 Population Profile for the United States*. Washington, DC: U.S. Bureau of the Census.

Note. The U.S Bureau of the Census does not list Hispanic as a racial group; it is considered as an ethnic group under the White category. Of the PKBS normative population, 5.2% were identified as Hispanic, as compared to 9% nationally, based on the Bureau's breakdown of the 1990 census data which was conducted by the Racial Statistics division.

subjects was then coded according to the seven occupational categories used by the U.S. Bureau of the Census, and percentages of this population in each of the seven categories were calculated. Table 3.6 shows the percentages of subject parents from the PKBS standardization sample in each of the seven occupational categories, with comparative information from the general U.S. population. As the data in Table 3.6 indicate, the percentages of subjects' parents from the PKBS standardization sample in each occupational category are very similar to the percentages for the general U.S. population. The only occupational category where the PKBS population appears to be substantially different from the general U.S. population is category VI (operators, fabricators, and laborers), which accounted for 24% of the PKBS standardization data, as compared with 11.5% of the general U.S. population.

To evaluate the effects of parent socioeconomic status on children's PKBS ratings, the occupational information for the PKBS standardization sample was recoded into a four-factor division of socioeconomic status ranging from high (category I) to middle (categories II, IV, and V) to lower (categories III and VI), and lowest (category VII). Pearson product-moment correlations were then computed between occupational-socioeconomic class and PKBS scores. The resulting coefficients were quite low, ranging from .05 (subscale B1) to .11 (Scale A total score). These results indicate that the effects of socioeconomic status are quite minimal in influencing children's behavioral ratings.

Raters and Settings

The target population for raters during the PKBS norming and standardization process was preschool or kindergarten teachers and parents. Table 3.7 presents the distribution of raters who completed PKBS forms. As these data indicate, the PKBS normative data reflect ratings provided by a broad array of informants in both the preschool/kindergarten and home settings. Because it was expected, based on previous research (Achenbach, McConoughy, & Howell, 1987; Wright & Piersel, 1992), that there might

Table 3.6

Occupational Categories of Parents of PKBS Standardization Sample Subjects: Percentages of Sample in Seven Occupational Categories, with Comparative Information from the General U.S. Population.

Occupational category	Percent in PKBS sample	Percent in U.S. population*
I. Managerial and professional workers	30%	25.1%
II. Technical, sales, and administrative support workers	20%	26.2%
III. Service workers	10%	11.6%
IV. Farming, forestry, and fishing workers	1%	2.1%
V. Precision production workers, craftsmen, and repairmen	9%	10.9%
VI. Operators, fabricators, and laborers	24%	11.5%
VII. Not currently in labor force, others	6%	12.5%

*SOURCE: *Decennial Census Summary, 1990 Population Profile for the United States.* Washington, DC: U.S. Bureau of the Census.

be modest differences between school-based and home-based behavior ratings of children, this area was further investigated. First, correlations were obtained between rating setting and PKBS scores by collapsing all rater categories into the two simple categories of school-based and home-based, and then calculating Pearson product-moment correlations between the setting and PKBS scores. The obtained correlation coefficients were relatively weak, ranging from .02 to .19. The correlation between rating setting and PKBS total scores was .18 for Scale A, and .13 for Scale B. To determine whether the differences between raters and settings were substantial enough to justify separate sets of normative data, a series of *t*-tests was conducted on the PKBS scores from a subset of 102 preschool-aged children in the normative sample who were rated by both their parents and teachers. Although the mean PKBS scores of these children were significantly different on 4 of the 10 PKBS subscales, these differences were inconsistent and relatively small. When significant differences were found, they tended to indicate that parents rated the children as having greater levels of both social competence and problem behavior than did teachers. However, there were no significant differences between parent and teacher ratings on 6 of the 10 PKBS subscales, and no clear pattern of greater scores was found between parent and teacher ratings in these cases. Thus, a decision was reached to not divide the normative data by rater or setting. Therefore, although behavior rating scales tend to produce both source and setting variance, the evidence regarding PKBS ratings is that these sources of variance do not lead to any predictable directionality or extreme differences, but tend to reflect a mixed and unpredictable variety of differences in perceptions, behavioral expectations, and actual child behavior across different settings.

Summary of Standardization Procedures

The information presented in this chapter indicates that the PKBS was developed and standardized using ratings of a very large and diverse group of children throughout the United States. Although

Table 3.7

Distribution of Informants Who Provided Child Ratings as Part of the PKBS Norming and Standardization Process

Rating source	Frequency	Percentage
Unspecified	29	1.0
Teacher	1178	41.3
Teacher aide	93	3.3
Parent, unspecified	169	5.9
Mother	1119	39.2
Father	104	3.6
Completed by both parents, jointly	162	5.7
TOTAL	**2,855**	**100**

the composition of the standardization sample is not identical to the general U.S. population on the various demographic variables of interest, the differences between the PKBS standardization sample and the general makeup of the U.S. population do not appear to be substantial enough to cast any serious doubts about the generalizability or utility of the PKBS normative data. During the development of the PKBS, a division of the test normative data into two age groups (3-4 and 5-6) with no further divisions based on gender or rating source was deemed to be the most warranted and practical approach. Included in this chapter are gender-by-age descriptive statistics for the PKBS scores from the normative sample, so that individual cases of PKBS ratings may be interpreted with the addition of more specific comparative data when occasional circumstances seem to dictate the desirability of such a practice.

4

TECHNICAL PROPERTIES

During the development of the PKBS a number of studies and procedures designed to assess the psychometric properties and technical adequacy of the scales were conducted. This chapter of the manual includes the results of these diverse procedures, as well as discussions of the practical implications of these findings for using the PKBS. Specifically, a number of reliability and validity studies are presented and discussed, including basic reliability estimates, evidence for the validity of the PKBS, and information pertaining to the development of the PKBS subscale and area score structure.

Reliability

Reliability is a term used to refer to the consistency or stability of a measure or, more specifically, how consistently the test scores generalize across three possible domains: different item samples, different times, and different scorers (Salvia & Ysseldyke, 1991). These three types of reliability are more commonly referred to as internal consistency, test-retest, and interrat reliability. An addition to these three forms of reliability, the Standard Error of Measurement (SE*m*) of a test is directly related to the consistency or stability of a measure, in that it provides information regarding how much random error is likely to be associated with obtained test scores. Each of these three types of reliability procedures, as well as the SE*m*, was investigated during the development of the PKBS and is discussed within this section of the chapter. Alternate forms reliability, a method of constructing two equivalent forms of a test and then correlating scores between the forms, is another widely used specific reliability method, but was not appropriate for the PKBS, as only one form of the scales exists.

Internal Consistency Reliability

The two most common methods of determining internal consistency reliability, namely Cronbach's (1951) coefficient alpha and the Spearman-Brown split-half reliability formula, were calculated using data from the entire PKBS standardization sample. Each of these two procedures provides measures of internal consistency, but they differ in that coefficient alpha is based on intercorrelations of all comparable parts of the same test, whereas split-half reliability divides the scale into two equivalent halves, and estimates the consistency between forms.

As the data in Table 4.1 indicate, both methods produced uniformly high coefficients of internal consistency. The coefficients for the two methods ranged from .81 to .97 for the PKBS subscale and area scores, whereas the coefficients for the two total scores ranged from .94 to .97. The slightly lower coefficients of internal consistency for the subscale and area scores as opposed to the total scores reflect the fact that internal consistency reliability is positively related to the number of items in a test (Salvia

& Ysseldyke, 1991). The differences in internal consistency between Scale A (Social Skills) and Scale B (Problem Behavior) scores are not significant. The general level and pattern of the obtained reliability coefficients suggests that the PKBS has strong internal consistency.

Standard Error of Measurement

The Standard Error of Measurement (SE*m*) is closely related to internal stability, in that the internal consistency coefficient is used in calculating the SE*m* (by multiplying the standard deviation by the square root of 1 minus the internal consistency coefficient), and because as the coefficient of internal consistency decreases, the SE*m* increases (Sattler, 1988). Unlike the coefficient of internal consistency, which is not helpful in directly interpreting individual scores, the SE*m* is useful in determining the limits of a "true score." The SE*m* is based on the notion that any *obtained score* consists of both a *true score* and an *error score* (the amount of random and systematic error or unreliability that is present in the test). The SE*m* can be used to provide a band of error that can be placed around an obtained test score to create a range within which the examinees' true score is likely to fall within some given level of probability (e.g., 85%, 90%, 95%).

Using the alpha coefficients from Table 4.1, SE*m*'s were calculated for the subscale, area, and total scores of the entire PKBS standardization sample. The resulting SE*m*'s also are presented in Table 4.1. The SE*m*'s for the PKBS scores are relatively small and suggest that the range of error that surrounds obtained PKBS scores is quite modest. As would be expected, the PKBS score areas with fewer items and lower internal coefficients produced SE*m*'s that were a larger percentage of the standard deviation for that score (e.g., 38% and 40% of the standard deviations for the internalizing problems subscales

Table 4.1

*Internal Consistency and Split Half Reliability of the PKBS, with Standard Error of Measurement (SE*m*) Estimates*

PKBS scores	**Coefficient alpha**	**Split-half**	**SE*m***
A1 Social Cooperation	.94	.92	1.93
A2 Social Interaction	.92	.90	1.61
A3 Social Independence	.88	.86	1.56
AT SOCIAL SKILLS TOTAL	.96	.94	2.82
B1 Self-Centered/Explosive	.94	.91	1.80
B2 Attention Problems/Overactive	.92	.92	1.56
B3 Antisocial/Aggressive	.91	.91	1.46
Externalizing Problems	.97	.95	2.86
B4 Social Withdrawal	.85	.85	1.54
B5 Anxiety/Somatic Problems	.84	.81	1.72
Internalizing Problems	.90	.89	2.37
BT PROBLEM BEHAVIOR TOTAL	.97	.96	3.84

B4 and B5) than did the PKBS scores with more items and higher internal consistency coefficients (e.g., 20% of the standard deviation for the Scale A total score, and 17% of the standard deviation for the Scale B total score).

Test-Retest Reliability

During the development of the PKBS, two related procedures were conducted to determine the test-retest reliability of the instrument, an index of temporal stability. Teachers of 82 children ages 3–5 who enrolled in private preschools rated these children at three different time intervals: a baseline rating point, a retest 3 weeks after baseline, and a retest 3 months after the baseline measure. Pearson product-moment correlations were calculated between the baseline PKBS scores and the PKBS scores obtained at the two retest periods. The resulting coefficients of stability, which are presented in Table 4.2, were all in the moderate to high range, and were significant at the $p < .001$ level. It is interesting to note that the test-retest reliabilities for the Scale B (Problem Behavior) scores were somewhat higher than for the Scale A (Social Skills) scores, indicating that in the early childhood/preschool populations, behaviors relating to social competence may be somewhat less stable over time than problem behaviors. Another interesting finding from the test-retest studies was that the stability coefficient for the Anxiety/Somatic Problems subscale on Scale B decreased sharply between the 3-week and the 3-month interval, declining from .81 to .36. This decline suggests that physical or somatic complaints that young

Table 4.2

Test-Retest Reliability of the PKBS at Three-Week and Three-Month Intervals, Based on Preschool Teachers Ratings of 82 Children

PKBS scores	Three week	Three month
A1 Social Cooperation	.65	.70
A2 Social Interaction	.62	.70
A3 Social Independence	.66	.66
AT SOCIAL SKILLS TOTAL	.58	.69
B1 Self-Centered/Explosive	.87	.75
B2 Attention Problems/Overactive	.86	.74
B3 Antisocial/Aggressive	.81	.75
Externalizing Problems	.87	.78
B4 Social Withdrawal	.68	.63
B5 Anxiety/Somatic Problems	.81	.36
Internalizing Problems	.80	.70
BT PROBLEM BEHAVIOR TOTAL	.86	.78

Note. All correlations are significant at $p < .001$.

Table 4.3

Interrater Reliability of the PKBS Using Scores of 82 Preschool Students Rated by Preschool Teachers and Teacher Aides

PKBS scores	**Reliability coefficients**
A1 Social Cooperation	.61**
A2 Social Interaction	.47**
A3 Social Independence	.36**
AT SOCIAL SKILLS TOTAL	.48**
B1 Self-Centered/Explosive	.62**
B2 Attention Problems/Overactive	.51**
B3 Antisocial/Aggressive	.62**
Externalizing Problems	.63**
B4 Social Withdrawal	.46**
B5 Anxiety/Somatic Problems	.42**
Internalizing Problems	.46**
BT PROBLEM BEHAVIOR TOTAL	.59**

*$p < .01$; **$p < .001$.

children experience (e.g., stomachaces, dizziness, aches and pains) may be relatively time-limited, and in many cases may be due to actual short-term illnesses rather than ongoing emotional turmoil. For the PKBS total scores, the stability coefficients were .58 and .69 for Scale A (Social Skills) at the two intervals, and .86 and .78 for Scale B (Problem Behavior), respectively. The fact that the coefficients of stability are somewhat lower than the coefficients of internal consistency is a typical finding in psychological and educational measurement, as test-retest reliability is subject to error through repeated measurement and behavioral change over time.

Interrater Reliability

To assess the interrater reliability of the PKBS, studies were conducted wherein score comparisons were made with two different pairs of raters. The first study included ratings of 82 children ages 3–5 from private preschools who were rated during a common 1-week time period by both teachers and classroom aides. Pearson product-moment correlations were calculated between the pairs of scores from the two raters, and the resulting coefficients are presented in Table 4.3. The resulting coefficients were all in the moderate range, with a fluctuation between .36 and .61 for the Scale A (Social Skills) scores, and .42 to .63 for Scale B (Problem Behavior) scores. These modest correlations are indicative of a fair

Table 4.4

Interrater Reliability of the PKBS Using Scores of 102 Preschool Students Who Were Rated by Both Preschool Teachers and Parents

PKBS scores	Reliability coefficients
A1 Social Cooperation	.57**
A2 Social Interaction	.27*
A3 Social Independence	.20
AT SOCIAL SKILLS TOTAL	.38**
B1 Self-Centered/Explosive	.28*
B2 Attention Problems/Overactive	.48**
B3 Antisocial/Aggressive	.47**
Externalizing Problems	.42**
B4 Social Withdrawal	.21
B5 Anxiety/Somatic Problems	.17
Internalizing Problems	.13
BT PROBLEM BEHAVIOR TOTAL	.16

*$p < .01$; **$p < .001$.

amount of cross-rater variance, which tends to indicate rater biases, different experiences with the subjects, and possible differences in following the rating instructions. As in the case of test-retest reliability, interrater reliability is typically substantially lower than internal consistency in behavioral assessment due to these sources of error.

The second study included PKBS scores of 102 children ages 3–5 who were rated during a common 1-week time frame by both their preschool teachers and one of their parents. The resulting Pearson product-moment correlation coefficients (shown in Table 4.4) were in the weak to moderate range, and were in some cases noticeably lower than the correlations obtained between teachers and classroom aides. The range of coefficients for Scale A (Social Skills) scores was .20 to .57, whereas the range of correlations for Scale B (Problem Behavior) scores was .13 to .48. Several of the interrater reliability coefficients in this study were statistically significant, but several were also extremely low. It is interesting to note that the lowest correlations were on the Scale B internalizing problems scales, suggesting that internalizing behavioral and emotional problems are more likely to fluctuate across settings. The lower coefficients obtained in these comparisons are likely due to the effects of source variance coupled with the effects of setting variance, and seem to suggest that preschool-age children may behave substantially differently across the home and preschool settings.

Table 4.5

Factor Loadings, Item-Subscale Correlations, and Item-Total Score Correlations for Items in Scale A, Social Skills

A1 Social Cooperation				A2 Social Interaction				A3 Social Independence			
Item	Load	Item-Sub.	Item-Total	Item	Load	Item-Sub.	Item-Total	Item	Load	Item-Sub.	Item-Total
2	.75	.77	.63	5	.73	.73	.65	1	.45	.48	.50
7	.80	.79	.66	14	.49	.63	.62	3	.61	.63	.58
10	.74	.75	.62	15	.54	.62	.61	4	.67	.68	.63
12	.60	.66	.61	17	.66	.69	.61	6	.47	.64	.68
16	.60	.61	.55	19	.51	.65	.69	8	.42	.50	.51
22	.70	.68	.56	20	.81	.79	.70	9	.62	.76	.71
23	.82	.80	.63	21	.68	.73	.69	11	.51	.67	.69
25	.61	.69	.68	24	.56	.44	.39	13	.52	.33	.31
28	.64	.69	.68	27	.62	.68	.73	18	.58	.63	.61
29	.77	.77	.62	33	.77	.70	.60	26	.49	.49	.48
30	.71	.75	.67	34	.73	.72	.68	31	.69	.69	.63
32	.67	.74	.69								

Validity

Test validity is usually defined along the lines of the usefulness of a test or, in other words, how well the test measures what it is purported to measure (Salvia & Ysseldyke, 1991). In order to be valid, a test must also have adequate reliability, but reliability in and of itself does not guarantee validity. Making judgments about the validity of a test is a process of evaluating the test against a variety of standards that must be met in order to make corresponding inferences about test results. The validity of a test generally cannot be ascertained through a single research procedure, but must be established through a variety of procedures conducted over time that provide specific types of validity evidence. During the development of the PKBS, several methods of test validation were researched, and the results of these procedures are reported in this section. These procedures include the following three major forms of test validity: (a) content validity, (b) construct validity, and (c) criterion-related validity.

Content Validity

In order to determine the existence of content validity, a judgmental examination of a test and how its contents are relevant to the construct purported to be measured must occur (Cronbach, 1990). As

is described in detail elsewhere in this manual, several methods of content validation were conducted in detail during the development of the PKBS. Items were developed and selected for possible inclusion in the PKBS due to their heuristic relationship with the overall dimensions of social skills and problem behavior as well as the underlying areas within each overall dimension. This initial item development stage was conducted in an exact and methodical manner, through investigating many published studies and other sources regarding social-behavioral development in the early childhood/preschool population. Following the development of a large pool of initial items, each potential item was analyzed and evaluated in detail by members of a 16-person panel of expert judges, resulting in a narrowing and reworking of the item pool. Following the factor analyses that were conducted on the PKBS scores of the standardization sample, a few more items were dropped from inclusion in the test because they did not appear to fit adequately within the theoretical parameters of the PKBS. The final pool of items for the PKBS following an extensive content validation process included 34 items in Scale A (Social Skills) and 42 items in Scale B (Problem Behavior). Again, this information on the development of the PKBS is presented in detail in Chapter 3.

An additional method of determining content validity of a test is to evaluate statistically how well individual items fit within the domains in which they are placed. Salvia and Ysseldyke (1991) have suggested that individual test items that do not correlate at least .25 to .30 or more with the total score of a test (and presumably with their respective subscale or area score as well) probably do not belong within the content domain that is being assessed by the test. To evaluate the content validity of the PKBS along these lines, Pearson product-moment correlations between individual PKBS items and their related subscale, area, and scale totals were calculated. The resulting correlation coefficients are found in Tables 4.5 through 4.8. As an inspection of these coefficients indicates, the relationships between individual PKBS items and the related subscale, area, and total scores meet, and in most cases greatly exceed, the minimum criteria suggested by Salvia and Ysseldyke. None of the items in Scale A (Social Skills) correlated with the total scale score at less than .31, and with their respective subscale at less than .33, and most coefficients were well in excess of these minimum numbers. On Scale B (Problem Behavior) none of the test items correlated at less than .38 with the scale total score, .35 with the related area score, or .48 with the related subscale score. As was true in the case of the Scale A items, the item-domain correlations for the vast majority of Scale B items were well in excess of the minimum recommended levels. These results suggest that the items grouped together within each PKBS scale and the related area and subscale are relatively homogeneous and fit within the appropriate domain of content.

The various procedures presented within this section provide substantial evidence for the content validity of the PKBS. The procedures utilized in item and scale development appear to be technically sound and the resulting evidence suggests that the PKBS contents sample theoretically and statistically cohesive traits and constructs.

Construct Validity

Construct validity refers to a test's ability to measure a theoretical trait or construct. A psychological construct cannot be observed or measured directly, but must be inferred by observing its effects. Likewise, construct validity is inferred through a systematic accumulation of evidence which shows that the test actually measures the underlying construct(s) that are purported to be measured. Compared to content validity and criterion-related validity, there are many more possible methods of establishing construct validity. Each method must be analyzed on its own merit, and then the accumulated evidence obtained through several methods is examined to make an overall judgment of construct validity. During the development of the PKBS, a variety of research procedures were conducted for either the primary or secondary purpose of establishing the construct validity of the test. These procedures and findings will be presented individually within this section, and then analyzed in a summary fashion at the end of the section.

Table 4.6

Broad-Band Factor Loadings, Item and Broad-Band Scale Correlations, and Item-Total Score Correlations for Items in Scale B, Problem Behavior

Externalizing Problems			
Item	Load	Item-Scale	Item-Total
1	.71	.74	.70
3	.66	.60	.56
6	.69	.69	.65
7	.72	.76	.75
8	.72	.75	.73
10	.66	.73	.73
11	.78	.79	.74
13	.67	.73	.71
14	.73	.77	.73
15	.52	.60	.62
16	.76	.81	.77
19	.70	.77	.76
20	.73	.76	.72
21	.64	.69	.67
22	.73	.79	.77
25	.66	.69	.67
26	.62	.64	.60
29	.70	.69	.63
31	.64	.74	.75
32	.59	.67	.68
34	.62	.70	.69
35	.55	.65	.71
37	.60	.64	67
39	.75	.77	73
40	.59	.69	.63
41	.69	.76	.74
42	.80	.80	.75

Internalizing Problems			
Item	Load	Item-Scale	Item-Total
2	.61	.55	.70
4	.48	.49	.47
5	.58	.53	.43
9	.64	.65	.56
12	.66	.61	.49
17	.67	.65	.58
18	.75	.66	.44
23	.49	.51	.49
24	.57	.51	.55
27	.70	.61	.61
28	.70	.62	.48
30	.71	.67	.55
33	.51	.35	.61
36	.56	.60	.62
38	.60	.51	.38

Internal consistency and interrelationships among PKBS scales. One way of helping to establish construct validity is by showing that individual test items correlate highly with the total test score. Demonstrating such a pattern of correlations provides evidence that the test measures a single broad construct. In the content validity discussion in this section of the chapter, the relationships between individual test items and total test scores have previously been discussed. As was noted in that section, all PKBS items on both scales were correlated with their respective total scores above the minimum levels suggested by Salvia and Ysseldyke (1991) for this purpose. These item-total and item-scale correlations (which are presented in Tables 4.5 through 4.8) thus provide evidence for the construct validity of the PKBS as well as its content validity.

A related method of demonstrating construct validity of a test is to assess the interrelationships among scales within a test. The purpose of this validation method is to determine whether or not the relationships among subscales are strong enough, and at the same time independent enough, to show that the underlying constructs measured by subscales are being drawn upon adequately. To provide evidence of construct validity in this manner, Pearson product-moment correlations were computed on the entire PKBS standardization sample within each of the two SSBS Scales to develop matrices of scale intercorrelations. The scale intercorrelation matrix for Scale A (Social Skills) is presented in Table 4.9, and these data for Scale B (Problem Behavior) are presented in Table 4.10. The range of correlations between the Social Skills subscales is between .58 and .76, whereas the subscale to total score

Table 4.7

Factor Loadings and Item-Subscale Correlations for Items in Externalizing Problem Subscales (B1, B2, B3) of Scale B, Problem Behavior

B1 Self-Centered/ Explosive			**B2 Attention Problems/ Overactive**			**B3 Antisocial/Aggressive**		
Item	Factor loading	Item-subscale	Item	Factor loading	Item-subscale	Item	Factor loading	Item-subscale
7	.77	.80	1	.69	.76	3	.72	.66
8	.60	.74	6	.62	.69	11	.48	.76
10	.62	.72	14	.52	.69	21	.61	.71
13	.70	.75	15	.75	.68	26	.73	.70
19	.70	.79	16	.65	.79	29	.73	.76
22	.64	.77	20	.78	.82	34	.58	.69
31	.50	.71	25	.76	.76	40	.61	.64
32	.53	.67	39	.64	.77	42	.58	.77
35	.72	.71						
37	.66	.68						
41	.69	.77						

Table 4.8

Factor Loadings and Item-Subscale Correlations for Items in Internalizing Problem Subscales (B4, B5) of Scale B, Problem Behavior

B4 Social Withdrawal			B5 Anxiety/ Somatic Problems		
Item	Factor loading	Item-subscale	Item	Factor loading	Item-subscale
4	.58	.52	2	.71	.60
12	.82	.65	5	.66	.57
17	.81	.72	9	.58	.62
27	.57	.57	18	.60	.66
28	.76	.65	23	.77	.57
30	.60	.62	24	.57	.51
33	.47	.52	36	.55	.55
			38	.43	.48

correlations range from .84 to .89. For the Problem Behavior subscale scores, the correlations range from .46 to .80. The correlations between subscale scores and their respective area scores for Scale B are in the .90 to .95 range, whereas the subscale scores correlate with the total Problem Behavior score from .74 to .94. These patterns of correlations provide additional evidence for the construct validity of the PKBS. The correlations between subscales in each scale are in the moderate to moderately high range, whereas the correlations between subscale and total scale (including area) scores are substantially higher. This pattern suggests that the subscales are all strongly related to the general constructs being measured (social skills or problem behavior), and are at the same time somewhat independent from each other, measuring specific facets of social skills or problem behavior.

As a final measure of this type of construct validity, the correlation between the total scores of the two PKBS scales (Social Skills and Problem Behavior) was computed. The resulting coefficient of -.56 is both negative and moderate, suggesting that children with strong social skills are likely to have lower levels of problem behavior than are children with poor social skills. However, the moderate strength of this correlation suggests that an evaluator cannot always predict a child's functioning in one behavioral domain simply from making an inference from his or her functioning in the other domain. There are likely to be many exceptions to the rule, such as gregarious children who are skilled at interacting with their peers but intimidate and bully them at the same time, and socially unskilled withdrawn children who exhibit few overt problem behaviors. Given that social skills and problem behavior are two theoretically separate, but related domains, the correlation of −.56 between the two scales provides additional evidence for the construct validity of the PKBS.

Factorial validity. If a test is found to have an empirically sound factor structure then it is said to have "factorial validity," which is another form of construct validity. During the development of the

Table 4.9

Intercorrelations Among PKBS Scale A (Social Skills) Scores

Social Skills Scores	A1	A2	A3	AT
A1 Social Cooperation	1.00	.59	.60	.84
A2 Social Interaction		1.00	.76	.90
A3 Social Independence			1.00	.88
AT SOCIAL SKILLS TOTAL				1.00

Note. All correlations are significant at $p < .001$.

PKBS, the factor structure of the scales was investigated in detail using a variety of factor analytic methods to find the most suitable final factor solution. During the initial factorial investigation the PKBS item level data were analyzed using the entire PKBS standardization sample as well as subsamples based on age levels, gender, and disability status breakdowns. Both orthogonal (varimax) and oblique (direct oblimin) rotation methods were utilized during this exploratory phase to help determine a final factor solution that was both clinically useful and statistically sound. With only a few exceptions between methods and samples, the same PKBS items tended to cluster together into factors. The factor solution for the PKBS that was deemed the most appropriate was an analysis of data from the entire standardization sample, utilizing a principle components analysis (Harmon, 1976) with the Kaiser normalization method (Kaiser, 1958).

On Scale A (Social Skills) the varimax rotation converged in 6 iterations, producing a final factor solution with 3 factors. The first factor, which consisted of 12 items, accounted for 40% of the variance (Eigenvalue = 14.01) and was labeled *Social Cooperation* (A1). The second factor, which consisted of 11 items, accounted for 9.8% of the variance (Eigenvalue = 3.44), and was labeled *Social Interaction* (A2). The final Social Skills factor consisted of 11 items and accounted for 4.3% of the variance (Eigenvalue = 1.50); this factor was labeled *Social Independence* (A3). The total amount of variance explained in this three-factor solution for Scale A was 54.1%. Each of the items in Scale A had a factor loading into its respective factor of .42 or higher. Of the final 34 items used in Scale A, 7 had cross loadings into more than one factor at .40 or higher. In the end, these 7 items were assigned to only 1 factor each, with the final assignment determined by the highest factor loading. The factor loadings, item-subscale correlations, and item-total correlations for each of the 34 Scale A items are presented in Table 4.5.

The factor structure for Scale B (Problem Behavior) proved to be considerably more complex than the final solution that was obtained for Scale A, and yielded both broad-band (area) and narrow-band (subscale) factors. After allowing the computer analysis program to determine automatically how many factors to select (which provided uninterpretable results), several factor extraction methods were utilized with predetermined numbers of factors specified. The factor solution that proved to be the most statistically cohesive and clinically useful was a principal components analysis with two factors prespecified. Using this analysis method the varimax rotation produced 2 factors in 3 iterations. The first factor accounted for 38.6% of the variance (Eigenvalue = 17.00) and included 27 items. This factor was labeled *Externalizing Problems*. The second factor accounted for 6.9% of the variance (Eigenvalue = 3.05) and included 15 items. This factor was labeled *Internalizing Problems*. The total variance explained through this factor analysis was 45.5%. Three of the initial 45 items developed for the Problem Behavior scale were dropped following the factor analysis, as they lacked specificity and had very weak loadings on

Table 4.10

Intercorrelations Among PKBS Scale B (Problem Behavior) Scores

Problem Behavior Scores	B1	B2	B3	EXT	B4	B5	INT	BT
B1 Self-Centered/Explosive	1.00	.79	.80	.95	.61	.62	.68	.94
B2 Attention Problems/Overactive		1.00	.78	.92	.55	.50	.58	.88
B3 Antisocial/Aggressive			1.00	.91	.54	.46	.55	.87
Externalizing Problems				1.00	.62	.59	.66	.97
B4 Social Withdrawal					1.00	.64	.89	.76
B5 Anxiety/Somatic Problems						1.00	.91	.74
Internalizing Problems							1.00	.83
BT PROBLEM BEHAVIOR TOTAL								1.00

both factors. All of the remaining 42 items loaded into their respective factors at .48 or higher, and only one of the remaining items (B36, "Is overly sensitive to criticism or scolding") cross-loaded into both factors at .40 or higher. After analyzing the content and factor loadings, this item was assigned to the Internalizing Problems factor, due to a substantially higher factor loading there. The factor loadings for individual Problem Behavior items into this broad-band two-factor solution are presented in Table 4.6, along with item-scale and item-total correlations.

Although the initial two-factor solution was very strong from a psychometric standpoint, and verified the internalizing-externalizing problem behavior dichotomy upon which the PKBS was conceptualized, the large diversity and number of items in the two broad-band factors (particularly in the Externalizing Problems factor) presented some potential difficulties from a practical and clinical standpoint. Of particular concern was the potential clinical usefulness of the two factors in assessing a variety of child behavior disorders. Thus, a decision was made to conduct separate factor analyses of the items in each of the two broad-band factors in order to determine whether "narrow-band" factors existed. This broad-band/narrow-band approach to factor analyses of child behavior problems is consistent with the methodology used to develop the Child Behavior Checklist (Achenbach, 1991a) and the Teacher's Report Form (Achenbach, 1991b), two of the most sophisticated and respected child behavior assessment instruments currently available.

Through a principal components analysis of the 27 Externalizing Problems items, a varimax rotation converged in 8 iterations, producing three narrow-band factors. The first factor accounted for 53.1% of the variance in the analysis (Eigenvalue = 14.33) and contained 11 items. This factor was labeled *Self-Centered/Explosive* (subscale B1) The second factor accounted for 5.7% of the variance (Eigenvalue = 1.53), included 8 items, and was labeled *Attention Problems/Overactive* (subscale B2). The third factor accounted for 4.8% of the variance (Eigenvalue = 1.29), also included 8 items, and was labeled *Antisocial/Aggressive* (B3). This particular factor analysis accounted for 63.6% of the variance in the included items. A varimax rotation of the 15 Internalizing Problems items converged in 3 iterations and produced 2 narrow-band factors. The first factor (B4, *Social Withdrawal*) accounted for 42.4% of the variance (Eigenvalue = 6.36) and included 7 items. The second factor (B5, *Anxiety/Somatic Problems*) accounted for 8.9% of the explained variance (Eigenvalue = 1.34) and included 8 items. The total explained variance ac-

counted for by the 15 Internalizing Problems items was 7.69%. The narrow-band factor loadings for the Problem Behavior subscales are presented in Tables 4.7 and 4.8, along with the item-subscale and item-total correlations. All items in Scale B loaded into their respective narrow-band factors at .43 higher. A few of the items loaded onto more than one subscale at .40 or higher, and as in the case of the Scale A items, these items were assigned to the subscale for which they had the highest loading.

To conduct a preliminary investigation of the stability of the PKBS factor structure, cases from the entire standardization sample (N = 2,855) were randomly split into two equal halves, and separate factor analyses were conducted with each of these samples based on the final factor solutions that were obtained for the two PKBS scales (principal components analysis, varimax rotation, with three factors specified for Scale A and two broad-band factors specified for Scale B). There was a high degree of congruence between the factor solutions obtained with each sample. For Scale A, 91% of the items clustered together in similar groups across both samples, whereas there was 100% similarity between the factor placement of items on Scale B. These findings provide some initial support for the stability of the PKBS factor structure across different samples of subjects. These findings should be considered preliminary, as future research on the PKBS factor structure using structural equation modeling and invariance analyses will provide more definitive data on the stability of the structure across samples.

In sum, the factor analytic data from the PKBS provide additional evidence of the construct validity of the instrument. The final factor solutions that were utilized in developing the PKBS subscales appear to be psychometrically strong, clinically useful, and sufficiently stable across samples. The Scale A (Social Skills) factors, although not built around the peer-related and adult-related forms of social adjustment that were reflected in the development of items, provide evidence for three separate forms of social skills development (cooperation, interaction, independence) that fit within the broad construct of social skills. The items within Scale B (Problem Behavior) easily fit within the internalizing-externalizing dichotomy upon which the scale was designed, and through additional analysis, also fit within separate narrow-band domains of problem behavior that may be very useful from a clinical standpoint.

Assessment of group differences. Another measure of a test's construct validity is the degree of difference in magnitude between scores obtained from groups of persons who are known to differ from each other in a way that meaningfully relates to the underlying psychological constructs measured by the test. Given that the PKBS was developed to measure social skills development and both externalizing and internalizing problem behaviors in children, three group membership variables were identified for further analysis due to previous evidence suggesting that PKBS scores would be influenced by group membership. The three variables that were identified for further investigation of group membership include age, gender, and special education or disability status. Given that social skills and problem behaviors are considered to be developmental constructs, it would be expected that significant PKBS score differences would be found between younger and older children. In terms of gender differences in social skills and problem behavior in the early childhood/preschool population, there is a substantial body of research suggesting that as a group, girls are rated as having better social skills and exhibiting fewer problem behaviors than boys (e.g., Gresham, Elliott, & Black, 1987; Merrell, 1993a; Merrell, Merz, Johnson, & Ring, 1992). Thus, in order to demonstrate construct validity, the PKBS would be expected to show significant gender differences. Additionally, there is considerable evidence that disability or special education status is an important variable in social skills and problem behavior development. Most studies investigating this topic have shown that as a group, students with developmental and learning disabilities have poorer social skills and exhibit more problem behavior than do their peers without disabilities (e.g., Bryan, Pearl, Donahue, Bryan, & Pflaum, 1983; Merrell, Sanders, & Popinga, 1993; Swanson & Malone, 1992). Therefore, it would be expected that significant differences in PKBS ratings would be found when scores of developmentally delayed and nondisabled children are compared.

To investigate the effects of gender and age on PKBS ratings, a two-way multivariate analysis of variance (MANOVA) was conducted on data from the entire standardization sample, using gender and age as independent variables and the combined PKBS subscales for both scales as dependent measures. For age, 3- and 4-year-old subjects were combined into one group and 5- and 6-year-old subjects were combined into a second group. An overall significant effect was found for gender, $F(8,2772) = 19.17$, $p < .0001$. Follow-up univariate ANOVAs showed significant gender effects for each of the PKBS subscales except for B5 (Anxiety/Somatic Problems). Female subjects' PKBS subscale scores showed significantly greater Social Skills ratings and significantly lower Problem Behavior ratings than those of male subjects on each of the PKBS subscales except B5 ($p < .0001$). As in the case of gender, an overall significant MANOVA effect was found for age, $F(8, 2772) = 18.41$, $p < .0001$. Follow-up univariate ANOVAs indicated significant differences between the age 3-4 and age 5-6 groups on most of the PKBS subscales, with the older group being rated as having better social skills development on all Social Skills subscales ($p < .0001$), fewer problem behaviors on Problem Behavior subscales B1 ($p = .006$), B2 ($p = .001$), and B4 ($p < .0001$). The differences between the two age groups on subscales B3 (Antisocial/Aggressive) and B5 (Anxiety/Somatic Problems) were not significant. The overall interaction between gender and age was not significant in this analysis, $F(8,2772) = .80$, $p = .60$.

To investigate the effects of disability or special education status on PKBS ratings, mean PKBS subscale scores of all subjects in the standardization sample who were identified as having a developmental disability ($n = 313$) were compared with ratings of the subjects who were not identified as having a disability ($n = 2542$). This comparison was accomplished by using a one-way MANOVA with disability status as the independent variable and the combined PKBS subscales as dependent measures. A significant overall effect was found for group membership, $F(8,2846) = 56.33$, and follow-up ANOVAs conducted separately with each PKBS subscale showed large, statistically significant ($p < .0001$) differences between the two groups on all PKBS subscale scores, with the nondisability group rated as having significantly better social skills and significantly fewer problem behaviors than the disability group.

In sum, significant group differences in PKBS scores were found based on the grouping variables of gender, age level, and disability status. These findings are consistent with previous research related to social skills and problem behavior development in children, and provide additional strong support for the construct validity of the PKBS.

Convergent and divergent construct validity. A final form of construct validity that was investigated during the development of the PKBS is an assessment of the convergent and divergent properties of the test. Convergent construct validity is ascertained by finding strong correlations with similar types of measures (i.e., measures that "converge" with the construct under investigation), whereas divergent validity is ascertained by showing evidence of weak relationships between different measures that purport to measure dissimilar constructs. Given that the PKBS is purported to measure both social skills and externalizing/internalizing problem behaviors, measures were selected for comparison that have been demonstrated to measure one or both of these constructs. Comparative investigations of the PKBS with four different behavior rating scales were ultimately conducted.

In the first convergent/divergent validity study, correlations between the PKBS and the preschool form of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990) were obtained using ratings of 86 preschool-age children (ages 3–5) who had been referred for special education child-find screenings in a large urban public school district. These children were rated on both measures by one of their parents (in most cases their mother). The parent form of the preschool-level SSRS includes a 39-item social skills scale with four subscales (Cooperation, Assertion, Responsibility, Self-Control) and a total score, as well as a brief 10-item problem behavior rating, which includes externalizing and internalizing items and a total score. Relationships between the raw scores of the PKBS and SSRS for this sample were obtained by computing Pearson product-moment correlation coefficients. The resulting coefficients are shown in Table 4.11. The social skills scores from the two measures were found to correlate at a moderate to strong level, with the coefficients ranging from .32 to .76. The strongest relationship was

found to exist between the total social skills scores of the two measures (.76). The problem behavior scores of the two measures were found to correlate to widely varying degrees, with a range of coefficients of .25 to .83. As in the case of the social skills scores, the strongest correlation for problem behavior scores (.83) was found for the two total scores. Not surprisingly, stronger correlations were found between internalizing to internalizing and externalizing to externalizing problem scores of the two measures. Negative correlations were obtained between the social skills and problem behavior scores of the two measures, and these correlations were weak to moderately strong, with the coefficients ranging from −.10 to −.66.

Table 4.11

Correlations Between PKBS Scores and Scores on the Social Skills Rating System for 86 Preschool-Aged Children Who Were Rated with Both Instruments by Parents or Guardians as Part of a Child-Find Screening Process

PKBS scores	Social Skills Rating System scores							
	C	A	R	SC	SST	E	I	PBT
A1 Social Cooperation	.60**	52**	32*	.73**	.59**	- 66**	-.36**	- 61**
A2 Social Interaction	.58**	.65**	.64**	.45**	.68**	- 32*	- 20	.31*
A3 Social Independence	.61**	.75**	.58**	.51**	.67**	-.42**	-.38**	.46**
AT SOCIAL SKILLS TOTAL	.71**	.74**	.60**	.69**	.76**	-.57**	-.37**	-.56**
B1 Self-Centered/Explosive	-.39**	-.33*	- 23	-.62**	-.41**	.80**	48**	77**
B2 Attention Problems/Overactive	- . .**	34**	-.33*	.63**	- 48**	73**	.40**	68**
B3 Antisocial/Aggressive	.33*	.23	- 13	.48**	- 32*	.69**	.44**	68**
Externalizing Problems	- 42**	- 32*	24	61**	42*	82**	46**	77**
B4 Social Withdrawal	-.33	- 34**	17	- 29*	.31*	.33*	41**	42**
B5 Anxiety/Somatic Problems	-.26*	.21	- 10	- 21	- 20	25*	38**	.35**
Internalizing Problems	-.48**	-.45**	- 29*	45**	- 45*	62*	53**	67**
BT PROBLEM BEHAVIOR TOTAL	-.48**	.41**	- 27*	-.62**	- 47*	82**	57**	83**

*$p < .01$; **$p < .001$.

Table Key for SSRS Scores:
C = Cooperation
A = Assertion
R = Responsibility
SC = Self-Control
SST = Social Skills Total
E = Externalizing Problems
I = Internalizing Problems
PBT = Problem Behavior Total

The second convergent/divergent validity study included teacher ratings of 116 developmentally delayed children (ages 3–6) served in special education preschool settings in a large metropolitan area school district. These children were rated by their teachers using both the PKBS and the Matson Evaluation of Social Skills with Youngsters (MESSY; Matson, Esvelt-Dawson, & Kazdin, 1983; Matson, Rotari, & Helsel, 1983, 1985). The MESSY teacher form is a 65-item scale that includes two factor scores, Inappropriate Assertiveness/Impulsivity and Appropriate Social Skills, that appear to sample both major domains of behavior measured by the PKBS. Pearson product-moment correlations obtained between the scores of the two measures (see Table 4.12) were moderate to very strong for the social skills scores (.62 to .85), and relatively weak to quite strong for the problem scores (.22 to .72). The correlations between the PKBS total scores and the related MESSY factors were substantial (.84 between PKBS Scale A total score and MESSY Factor II; .64 between PKBS Scale B total score and MESSY Factor I). Correlations between opposing construct scores of both measures were for the most part negative, and were very weak to moderate in strength.

The third convergent/divergent validity study included a comparison of the PKBS with the 39-item version of the Conners Teacher Rating Scales (CTRS-39; Conners, 1990). The CTRS-39 is a well-researched and widely used problem behavior rating scale that has been in use for approximately two decades.

Table 4.12

Correlations Between PKBS Scores and Scores on the Matson Evaluation of Social Skills with Youngsters (MESSY) Teacher Report for 116 Developmentally Delayed Preschool-Aged Children Who Were Rated with Both Instruments by Their Special Education Preschool Teachers

PKBS scores	MESSY acher report scores	
	I. Inappropriate Assertiveness/Impulsivity	II. Appropriate Social Skills
A1 Social Cooperation	-.23*	.62**
A2 Social Interaction	.30**	.85**
A3 Social Independence	.10	.78**
AT SOCIAL SKILLS TOTAL	.08	.84**
B1 Self-Centered/Explosive	.68**	-.17
B2 Attention Problems/Overactive	.51**	-.16
B3 Antisocial/Aggressive	72**	- 03
Externalizing Problems	70**	-.15
B4 Social Withdrawal	.22*	- 56**
B5 Anxiety/Somatic Problems	40**	- 14
Internalizing Problems	35**	- 38**
BT PROBLEM BEHAVIOR TOTAL	64**	- 25*

*$p < .01$; **$p < .001$.

The 39 items are divided into 6 subscale or factor scores (Hyperactivity, Conduct Problems, Emotional Overindulgent, Anxious-Passive, Asocial, Daydream-Attention Problem), as well as a 10-item "Hyperactivity Index" which is composed of various items from each subscale that have been found to be most sensitive to pharmacological effects in treatment studies with hyperactive children. The sample for this study consisted of 46 students (ages 5 and 6) in regular Kindergarten classes in a medium-sized suburban school district who were rated by their teachers with both the PKBS and CTRS-39. The resulting Pearson product-moment correlation coefficients, which are displayed in Table 4.13, are indicative of a pattern of weak to extremely strong relationships between the two measures, depending on the par-

Table 4.13

Correlations Between PKBS Scores and Scores on the Conners Teacher Rating Scale-39 for 46 Regular Kindergarten Students Who Were Rated with Both Instruments by Their Teachers

PKBS scores	CTRS-39 scores						
	A	B	C	D	E	F	I
A1 Social Cooperation	-.83**	-.70**	-.46**	-.08	-.64**	-.54**	-.77**
A2 Social Interaction	-.59**	-.36*	-.45**	-.43*	-.63**	-.71**	-.52**
A3 Social Independence	-.43*	-.20	-.49**	-.54**	-.68**	-.59**	-.42*
AT SOCIAL SKILLS TOTAL	-.74**	-.50**	-.55**	-.41*	-.77**	-.73**	-.68**
B1 Self-Centered/Explosive	.82**	.90**	.76**	.11	.54**	.45**	.85**
B2 Attention Problems/Overactive	.83**	.78**	.56**	.10	.47**	.35*	.84**
B3 Antisocial/Aggressive	.43*	.34*	.35*	.18	.60**	.60**	.41*
Externalizing Problems	.85**	.87**	.67**	.12	.57**	.44*	.85**
B4 Social Withdrawal	.59**	.53**	.67**	.50**	.60**	.53**	.52**
B5 Anxiety/Somatic Problems	.61**	.47**	.74**	.59**	.41*	.50**	.56**
Internalizing Problems	.66**	.55**	.78**	.61**	.56**	.57**	.59**
BT PROBLEM BEHAVIOR TOTAL	.85**	.83**	.76**	.29	.62**	.52**	.84**

*$p < .01$; **$p < .001$.

Table Key for CTRS-39 Scores:

A = Hyperactivity
B = Conduct Problems
C = Emotional Overindulgent
D = Anxious-Passive
E = Asocial
F = Daydream-Attention Problem
I = Hyperactivity Index

ticular scale comparisons that are made. In general, very strong relationships were found between scores that appear to measure similar constructs. For example, the Externalizing Problems score from the PKBS correlated highly with the CTRS-39 scale scores that appeared to be more "externalizing" in their focus: .85 with Hyperactivity, .87 with Conduct Problem, and .85 with the Hyperactivity Index. On the other hand, the Internalizing Problems score from the PKBS correlated at .78 with the CTRS-39 Emotional/Indulgent scale, and .61 with the Anxious/Passive scale, which both appear to be more "internalizing" in their focus. Although the primary focus of this investigation was the PKBS Problem Behavior scale (Scale B), the negative correlations between the CTRS-39 and PKBS Scale A (Social Skills) scores, many of which are quite strong in magnitude, are also of interest.

The final convergent/divergent validity study conducted during the development of the PKBS was a comparison of the PKBS with the School Social Behavior Scales (SSBS; Merrell, 1993b). The SSBS is a school-based social behavior rating instrument for K-12 students that is conceptually very similar to the PKBS. It includes two separate scales: a 32-item social competence scale (Scale A) and a 33-item antisocial behavior scale (Scale B). Both of these scales include three empirically derived subscales. In addition to the differences in targeted age ranges between the two instruments, the major conceptual difference involves the problem behavior scales. Scale B of the PKBS was designed to measure both the externalizing and internalizing domains of problem behavior, whereas Scale B of the SSBS was designed specifically to measure problem behaviors that are *antisocial* in nature, and thus more closely related to the externalizing domain. The sample for this study included ratings of 47 regular Kindergarten students (ages 5 and 6) from a large metropolitan school district who were rated with both instruments by their Kindergarten teacher. Pearson product-moment coefficients obtained between the two measures showed evidence of moderate to very strong consistency between the social skills scores of the PKBS and the social competence scores of the SSBS; the median correlation was .68, whereas the correlation between the two total scores was .86. Given the differences between the constructs being measured, the correlations between the PKBS Problem Behavior scores and the SSBS antisocial behavior scores were understandably more variable in nature. The externalizing problems score of the PKBS was substantially correlated with the SSBS antisocial behavior scores (.75 to .83), whereas the correlations between the SSBS antisocial behavior scores and PKBS internalizing problems scores evidenced a weaker relationship (.36 to .45). The total scores of the two scales were correlated at a relatively strong .77. As in the case of the previous three convergent validity studies, the relationships between social skills and problem behavior scores of different instruments were negative, and quite variable in strength, ranging from very weak to very strong.

In sum, the four studies reviewed in this section provide strong additional evidence of the construct validity of the PKBS. In general, moderately strong to very strong relationships were found to exist between PKBS scores and scores from established rating scales measuring similar constructs (convergent validity). The weaker relationships that were found to exist between PKBS scores and scores from other measures tapping different constructs (e.g., the correlation of .12 between the PKBS externalizing problems score and the Anxious/Passive scale of the CTRS-39) provide evidence of divergent validity, which is another way of showing construct validity.

Criterion-Related Validity

A test is considered to have criterion-related validity when scores from that test have been found to correlate highly with scores or decisions from some external criterion. When discussing the criterion-related validity of a test, the test is often referred to as the "predictor." Criterion-related validity may be either *concurrent* or *predictive* in nature. Concurrent criterion-related validity data are collected when the criterion and predictor data are collected at a similar point in time. Predictive criterion-related validity data are collected when the purpose of the validation study is to determine if the test can predict future performance on the criterion. Thus, predictive validity studies necessitate a sufficient length of time between the obtaining of predictor and criterion scores.

Table 4.14

Correlations Between PKBS Scores and Scores on the School Social Behavior Scales for 47 Regular Kindergarten Students Who Were Rated with Both Instruments by Their Teachers

PKBS scores	SSBS scores							
	A1	A2	A3	AT	B1	B2	B3	BT
A1 Social Cooperation	.47**	.73**	.73**	.68**	-.48*	- 61**	-.69**	-.63**
A2 Social Interaction	.85**	.58**	.53**	.79**	-.20	-.32	-.38*	-.31
A3 Social Independence	.81**	.46**	.54**	.74**	-.16	-.25	-.34*	-.26
AT SOCIAL SKILLS TOTAL	.83**	.68**	.68**	.86**	-.31	-.44**	-.53**	-.45
B1 Self-Centered/Explosive	-.28	-.74**	-.56**	-.54**	.76**	68**	.69**	.78**
B2 Attention Problems/Overactive	-.29	-.62**	-.55**	-.53**	.54**	.74**	.75**	70**
B3 Antisocial/Aggressive	-.19	-.65**	-.42*	-.43*	.74**	.83**	.61**	.77**
Externalizing Problems	- 30	- 76**	-.58**	-.57**	.75**	.82**	.76**	.83**
B4 Social Withdrawal	-.72**	-.62**	-.50**	-.73**	.40*	.42*	.50**	48**
B5 Anxiety/Somatic Problems	-.49**	-.43*	-.40*	-.53**	.34*	.24	.36**	35*
Internalizing Problems	-.64**	-.55**	-.47**	-.67**	.39*	.36	.45**	.44*
BT PROBLEM BEHAVIOR TOTAL	-.48**	-.77**	-.61**	-.68**	.70**	.73**	.73**	.77**

$^{*}p < .01$; $^{**}p < .001$.

Table Key for SSBS Scores:
A1 = Interpersonal Skills
A2 = Self-Management Skills
A3 = Academic Skills
AT = Social Competence Total
B1 = Hostile-Irritable
B2 = Antisocial-Aggressive
B3 = Disruptive-Demanding
BT = Antisocial Behavior Total

During the development of the PKBS, a concurrent criterion-related validity study was conducted wherein the ability of the PKBS to predict the special education status of rated children was analyzed. The sample for this study included 1,771 children who attended either a preschool or kindergarten class and who were rated on the PKBS by their teacher. Of this particular population sample, 192 received special education services through their having been identified as developmentally delayed, whereas 1,579 had not been identified for special education purposes. Using the PKBS ratings from this sample, the ability of the PKBS to predict special education classification correctly was analyzed through a two-group linear discriminant function analysis using the direct entry method. Discriminant function analysis is a procedure wherein the accuracy of classification rates is determined by analyzing linear combinations of the predictor variables that may be used to classify subjects into groups. Group membership

in this analysis was determined by the presence or absence of special education classification, whereas the variables in the analysis were the combined PKBS subscales for both Scale A and Scale B (the area and total scores were excluded from the analysis to eliminate collinearity problems).

The discriminant procedure produced an overall significant classification effect, Wilks' Lambda = .84, $\chi^2(8) = 297.87$, $p < .0001$. The classification results table produced through the discriminant procedures indicated that 90.18% of the grouped cases had been correctly classified in special education status groups based solely on the critical statistical properties of the PKBS scores. It is interesting to note that the PKBS subscale having the largest correlation with the discriminant function was Scale A2 (Social Interaction), with a pooled within-groups correlation of .82 on the structure matrix. The PKBS subscale that had the smallest correlation with the discriminant function was Scale B3 (Antisocial/Aggressive), which yielded a pooled within-groups correlation of −.11. These results provide some tentative support for the criterion-related validity of the PKBS as an instrument for use in early childhood special education classification. However, additional studies using expectancy tables with positive "hit rates" would be useful in validating the PKBS specifically for the purpose of special education classification, and test users should keep in mind that no single procedure is to be used to determine special education eligibility.

Reliability and Validity Summary

Although the PKBS is a new instrument and the evidence reviewed in this chapter should be viewed as initial or prelimi ary, these data provide substantial evidence for the psychometric and technical properties of the instrument. Several forms of test reliability for the PKBS were reported, including internal consistency, test-retest, and interrater reliability. The internal stability of the PKBS is very strong, with subsequently low Standard Errors of Measurement (SE*m*). The test-retest reliability of the PKBS should be considered moderate to relatively strong at this point, providing evidence that behavioral ratings fluctuate somewhat over time, but remain relatively constant within raters. At the present time, the interrater reliability of the PKBS should be considered to be weak to moderate, indicating that both source and setting variance influence behavior ratings to a substantial amount, which is consistent with other research on rating scale technology. Very strong evidence supporting the content and construct validity of the PKBS has been presented, and along these lines, the PKBS appears to have a clinically useful and statistically sound factor structure. Preliminary evidence for the criterion-related validity of the PKBS has also been presented, suggesting that there may be a strong relationship between PKBS ratings and the external criterion of special education classification status, and that the instrument may be useful as one part of a complete assessment battery for classification and eligibility purposes.

Request for Research Data

Research with the PKBS is ongoing and future revisions of the PKBS manual are expected as additional evidence accrues. The author is greatly interested in the results of new studies using the PKBS, and would appreciate receiving copies of research reports (i.e., journal articles, conference presentations, thesis and dissertation summaries, and unpublished research reports). Please send such information to: Kenneth W. Merrell, Ph.D., Department of Psychology, Utah State University, Logan, UT 84322-2810. The author is also willing to provide consultation and technical advice to investigators considering using the PKBS in their own research.

REFERENCES

Achenbach, T. M. (1985). *Assessment and taxonomy of child and adolescent psychopathology*. Beverly Hills, CA: Sage.

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Edelbrock, C. S. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monographs for the Society for Research in Child Development, 46* (1, serial no. 188).

Achenbach, T. M., McConoughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

AERA, APA, & NCME. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Barnett, D. W., & Casey, K. T. (1992). *Designing interventions for preschool learning and behavior problems*. San Francisco: Jossey-Bass.

Bryan, T., Pearl, R., Donahue, M., Bryan, J., & Pflaum, S. (1983). The Chicago Institute for the Study of Learning Disabilities. *Exceptional Education Quarterly, 41*, 1-23.

Campbell, S. B. (1991). Active and aggressive preschoolers. In D. Cicchetti & S. L. Toth (Eds.), *Internalizing and externalizing expressions of dysfunction* (pp. 57-89). Hillsdale, NJ: Lawrence Erlbaum Associates.

Campbell, S. B., & Werry, J. S. (1986). Attention deficit disorder (hyperactivity). In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (3rd ed., pp. 111-155). New York: Wiley.

Children's Defense Fund. (1992). *The state of America's children, 1992*. Washington, DC: Author.

Cicchetti, D., & Toth, S. L. (1991). A developmental perspective on internalizing and externalizing disorders. In D. Cicchetti & S. L. Toth (Eds.), *Internalizing and externalizing expressions of dysfunction* (pp. 1-19). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cone, J. E., & Hawkins, R. P. (1977). *Behavioral assessment: New directions in clinical psychology*. New York: Brunner/Mazel.

Conners, C. K. (1990). *Conners Rating Scales Manual*. North Tonawanda, NY: Multi-Health Systems.

Conners, C. K., & Werry, J. S. (1979). Pharmacotherapy. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (2nd ed.). New York: Wiley.

Cronbach, L. (1951). Coefficient Alpha and the internal structure of tests. *Psychometrika, 16,* 297-334.

Cronbach, L. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper and Row.

Cullinan, D., Epstein, M. H., & Kauffman, J. M. (1984). Teacher's ratings of students' behaviors: What constitutes behavior disorder in schools? *Behavioral Disorders, 10,* 9-19.

Gresham, F. M. (1986). Conceptual issues in the assessment of social competence in children. In P. Strain, M. Guralnick, & H. Walker (Eds.), *Children's social behavior: Development, assessment, and modification* (pp. 143-179). New York: Academic Press.

Gresham, F. M., & Elliott, S. N. (1990). *The Social Skills Rating System.* Circle Pines, MN: American Guidance.

Gresham, F. M., Elliott, S. N., & Black, F. L. (1987). Teacher-rated social skills of mainstreamed mildly handicapped and nonhandicapped children. *School Psychology Review, 16,* 78-88.

Gresham, F. M., & Reschly, D. J. (1987). Dimensions of social competence: Method factors in the assessment of adaptive behavior, social skills, and peer acceptance. *Journal of School Psychology, 25,* 367-381.

Harmon, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.

Kaiser, H. F (1958). The varimax criteria for analytic rotation in factor analysis. *Psychometrika, 23,* 187-200.

Kauffman, J. M. (1989). *Characteristics of behavior disorders of children and youth* (4th ed.). Columbus, OH: Merrill.

Kauffman, J. M. (1992). *Characteristics of behavior disorders of children and youth* (5th ed.). Columbus, OH: Merrill.

Kazdin, A. E. (1979). Situational specificity: The two-edged sword of behavioral assessment. *Behavioral Assessment, 1,* 57-75.

Kerr, M. M., & Nelson, C. M. (1989). *Strategies for managing behavior problems in the classroom* (2nd ed.). Columbus, OH: Merrill.

Kohn, M. L., & Schooler, C. (1973). Occupational experience and psychological functioning: An assessment of reciprocal effects. *American Sociological Review, 38,* 97-118.

Kohn, M. L., & Schooler, C. (1978). The reciprocal effects of substantive complexity of work and intellectual flexibility: A longitudinal assessment. *American Journal of Sociology, 84,* 24-52.

Kohn, M. L., & Schooler, C. (1983). *Work and personality: An inquiry into the impact of social stratification.* Norwood, NJ: Ablex.

Landau, S., & Milich, R. (1990). Assessment of children's social status and peer relations. In A. M. LaGreca (Ed.), *Through the eyes of the child* (pp. 259-291). Boston: Allyn & Bacon.

Lanyon, R. I., & Goodstein, L. D. (1982). *Personality assessment* (2nd ed.). New York: Wiley.

Martin, R. P. (1988). *Assessment of personality and behavior problems.* New York: Guilford Press.

Martin, R. P., Hooper, S., & Snow, J. (1986). Behavior rating scale approaches to personality assessment in children and adolescents. In H. Knoff (Ed.), *The assessment of child and adolescent personality* (pp. 309-351). New York: Guilford Press.

Masser, J. D., & Cloninger, C. R. (1990). *Comorbidity of mood and anxiety disorders.* Washington, DC: American Psychiatric Press.

Matson, J. L., Esvelt-Dawson, K., & Kazdin, A. E. (1983). Validation methods for assessing social skills in children. *Journal of Clinical Child Psychology, 12,* 174-180.

Matson, J. L., Rotari, A. F., & Helsel, W. J. (1983). Development of a rating scale to measure social skills in children: The Matson Evaluation of Social Skills with Youngsters (MESSY). *Behaviour Research and Therapy, 21,* 335-340.

Matson, J. L., Rotari, A. F., & Helsel, W. J. (1985). Development of a rating scale to measure social skills in children: The Matson Evaluation of Social Skills with Youngsters (MESSY). *Psychopharmacology Bulletin, 21,* 855-867.

McMahon, R. J. (1984). Behavioral checklists and rating scales. In T. H. Ollendick & M. Herson (Eds.), *Child behavioral assessment: Principles and practices* (pp. 80-105). New York: Pergamon.

Merrell, K. W. (1993a). Using behavior rating scales to assess social skills and antisocial behavior in school settings: Development of the School Social Behavior scales. *School Psychology Review, 22,* 115-133.

Merrell, K. W. (1993b). *School Social Behavior Scales.* Brandon, VT: Clinical Psychology Publishing Company.

Merrell, K. W. (1994). *Assessment of behavioral, social, and emotional problems.* White Plains, NY: Longman.

Merrell, K. W., Merz, J. M., Johnson, E. R., & Ring, E. N. (1992). Social competence of students with mild handicaps and low achievement: A comparative study. *School Psychology Review, 21,* 125-137.

Merrell, K. W., Sanders, D. E., & Popinga, M. R. (1993). Teacher ratings of student social behavior as a predictor of special education status: Discriminant validity of the School Social Behavior Scales. *Journal of Psychoeducational Assessment, 11,* 220-231.

Oden, S. L., & Asher, S. R. (1977). Coaching children in social skills for friendship making. *Child Development, 48,* 496-506.

Odom, S. L., McConnell, S. R., & McEvoy, M. A. (1992). *Social competence of young children with disabilities.* Baltimore: Paul H. Brookes.

Peacock Hill Working Group. (1991). Problems and promises in special education and related services for children and youth with emotional or behavioral disorders. *Behavioral Disorders, 16,* 299-313.

Quay, H. C. (1986). Classification. In H. C. Quay & J. R. Werry (Eds.), *Psychopathological disorders of childhood* (3rd ed., pp. 1-34). New York: Wiley.

Salvia, J., & Ysseldyke, J. E. (1991). *Assessment* (5th ed.). Boston: Houghton-Mifflin.

Sattler, J. M. (1988). *Assessment of children* (3rd ed.). San Diego, CA: Jerome M. Sattler.

Swanson, H. L., & Malone, S. (1992). Social skills and learning disabilities: A meta-analysis of the literature. *School Psychology Review, 21,* 427-442.

Tuma, J. M. (1989). Mental health services for children: The state of the art. *American Psychologist, 44,* 188-199.

U.S. Bureau of the Census. (1990). *Decennial Census Summary, 1990 Population Profile for the United States.* Washington, DC: U.S. Bureau of the Census.

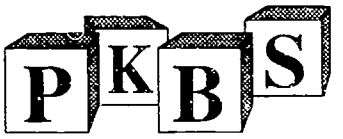
U.S. Department of Education. (1990). *Annual report to Congress on the implementation of Public Law 94-142.* Washington, DC: Federal Government Printing Office.

Walker, H. M., & McConnell, S. R. (1988). *The Walker-McConnell Scale of Social Competence and School Adjustment.* Eugene, OR: University of Oregon Center on Human Development.

Walker, H. M., McConnell, S. R., & Clarke, J. Y. (1985). Social skills training in school settings: A model for social integration of handicapped children into less restrictive settings. In R. McMahon & R. D. Peters (Eds.), *Childhood disorders: Behavioral-developmental approaches* (pp. 140-168). New York: Brunner-Mazel.

Wright, D., & Piersel, W. C. (1992). Components of variance in behavior ratings from parents and teachers. *Journal of Psychoeducational Assessment, 10,* 310-318.

Preschool and Kindergarten Behavior Scales



Kenneth W. Merrell, Ph.D.

Child Information

Child Name _____
Last First

Age: Years Months Sex: M F

Is this child receiving services in a school or in a school-related program (e.g., Preschool, Head Start, etc.)? ____Yes ____No

If Yes, what is the name of the school and the program?

If this child has a disability, please list the special education service category or classification:

Rater Information

Rated By _____

Relationship To Child

Date Completed

List the setting(s) in which you observe or interact with the child:

Instructions

Please rate the child on each of the items on pages 2 and 3 of this rating form. Ratings should be based on your observations of this child's behavior **during the past three months**. The rating points after each item appear in the following format:

Never	Rarely	Sometimes	Often
0	1	2	3

Never If the child does not exhibit a specified behavior, or if you have not had an opportunity to observe it, circle 0, which indicates *Never*.

Rarely If the child exhibits a specified behavior or characteristic, but only very infrequently, circle 1, which indicates *Rarely*.

Sometimes If the child occasionally exhibits a specified behavior or characteristic, circle 2, which indicates *Sometimes*.

Often If the child frequently exhibits a specified behavior or characteristic, circle 3, which indicates *Often*.

Please complete all items, and do not circle between numbers.

CPPC

PKBS

Scale A
Social Skills

	Never	Rarely	Sometimes	Often	Scoring Key		
1. Works or plays independently	0	1	2	3			
2. Is cooperative	0	1	2	3			
3. Smiles and laughs with other children	0	1	2	3			
4. Plays with several different children	0	1	2	3			
5. Tries to understand another child's behavior ("Why are you crying?")	0	1	2	3			
6. Is accepted and liked by other children	0	1	2	3			
7. Follows instructions from adults	0	1	2	3			
8. Attempts new tasks before asking for help	0	1	2	3			
9. Makes friends easily	0	1	2	3			
10. Shows self-control	0	1	2	3			
11. Is invited by other children to play	0	1	2	3			
12. Uses free time in an acceptable way	0	1	2	3			
13. Is able to separate from parent without extreme distress	0	1	2	3			
14. Participates in family or classroom discussions	0	1	2	3			
15. Asks for help from adults when needed	0	1	2	3			
16. Sits and listens when stories are being read	0	1	2	3			
17. Stands up for other children's rights ("That's his!")	0	1	2	3			
18. Adapts well to different environments	0	1	2	3			
19. Has skills or abilities that are admired by peers	0	1	2	3			
20. Comforts other children who are upset	0	1	2	3			
21. Invites other children to play	0	1	2	3			
22. Cleans up his/her messes when asked	0	1	2	3			
23. Follows rules	0	1	2	3			
24. Seeks comfort from an adult when hurt	0	1	2	3			
25. Shares toys and other belongings	0	1	2	3			
26. Stands up for his/her rights	0	1	2	3			
27. Apologizes for accidental behavior that may upset others	0	1	2	3			
28. Gives in or compromises with peers when appropriate	0	1	2	3			
29. Accepts decisions made by adults	0	1	2	3			
30. Takes turns with toys and other objects	0	1	2	3			
31. Is confident in social situations	0	1	2	3			
32. Responds appropriately when corrected	0	1	2	3			
33. Is sensitive to adult problems ("Are you sad?")	0	1	2	3			
34. Shows affection for other children	0	1	2	3			
				Totals			
					A1	A2	A3

PKBS

Scale B
Problem Behavior

	Never	Rarely	Sometimes	Often
1. Acts impulsively without thinking	0	1	2	3
2. Becomes sick when upset or afraid	0	1	2	3
3. Teases or makes fun of other children	0	1	2	3
4. Does not respond to affection from others	0	1	2	3
5. Clings to parent or caregiver	0	1	2	3
6. Makes noises that annoy others	0	1	2	3
7. Has temper outbursts or tantrums	0	1	2	3
8. Wants all the attention	0	1	2	3
9. Is anxious or tense	0	1	2	3
10. Will not share	0	1	2	3
11. Is physically aggressive (hits, kicks, pushes)	0	1	2	3
12. Avoids playing with other children	0	1	2	3
13. Yells or screams when angry	0	1	2	3
14. Takes things away from other children	0	1	2	3
15. Has difficulty concentrating or staying on task	0	1	2	3
16. Disobeys rules	0	1	2	3
17. Has problems making friends	0	1	2	3
18. Is afraid or fearful	0	1	2	3
19. Must have his/her own way	0	1	2	3
20. Is overly active; unable to sit still	0	1	2	3
21. Seeks revenge against others	0	1	2	3
22. Defies parent, teacher, or caregiver	0	1	2	3
23. Complains of aches, pain, or sickness	0	1	2	3
24. Resists going to preschool or day care	0	1	2	3
25. Is restless and "fidgety"	0	1	2	3
26. Calls people names	0	1	2	3
27. Is difficult to comfort when upset	0	1	2	3
28. Withdraws from the company of others	0	1	2	3
29. Bullies or intimidates other children	0	1	2	3
30. Seems unhappy or depressed	0	1	2	3
31. Has unpredictable behavior	0	1	2	3
32. Is jealous of other children	0	1	2	3
33. Acts younger than his/her age	0	1	2	3
34. Destroys things that belong to others	0	1	2	3
35. Is moody or temperamental	0	1	2	3
36. Is overly sensitive to criticism or scolding	0	1	2	3
37. Whines or complains	0	1	2	3
38. Gets taken advantage of by other children	0	1	2	3
39. Disrupts ongoing activities	0	1	2	3
40. Tells lies	0	1	2	3
41. Is easily provoked; has a "short fuse"	0	1	2	3
42. Bothers and annoys other children	0	1	2	3

Scoring Key

Totals	B1	B2	B3	B4	B5

Additional Information

Please use the following lines to provide any additional information about this child that you believe would be useful.

PKBS Score Grid

PKBS Score	Raw Score	Standard Score	Percentile Rank	Functional Level
SOCIAL SKILLS				
A1 Social Cooperation				
A2 Social Interaction				
A3 Social Independence				
AT SOCIAL SKILLS TOTAL				
PROBLEM BEHAVIOR				
B1 Self-Centered/Explosive				
B2 Attention Problems/Overactive				
B3 Antisocial/Aggressive				
Externalizing Problems				
B4 Social Withdrawal				
B5 Anxiety/Somatic Problems				
Internalizing Problems				
BT PROBLEM BEHAVIOR TOTAL				

Note: For Scale A, higher scores indicate greater levels of social skills.
For Scale B, higher scores indicate greater levels of problem behaviors.

ISBN 0-88422-152-0