MEASURES OF INDIVIDUAL FUNCTIONING: STRESS AND SELF-ESTEEM

Project Intern: Katie Kensinger
Project Director: Shelley M. MacDermid

Project EASE:
Evaluation Assistance Services

A partnership with the United Way of Greater Lafayette
1998

Strengthening services to families with research, education, and outreach
Table of Contents

Project EASe II - Summer 1998
Katie Kensinger, Intern

I. Table of Contents

II. Final Report
   Summary of Project EASe II
   Summary of Intern’s Activities
   Recommended Instruments with Accompanying Information

III. Appendix A
   Additional Resources
Project EASe II - Summer 1998

Summary of Project EASe II

The purpose of Project EASe II was to assist United Way funded agencies with different aspects of outcome-based evaluation. The goals for this phase of Project EASe were three-fold: 1) to identify and obtain measurement tools that would be useful to the majority of United Way agencies, 2) to review research about two different intervention types to determine reasonable outcome expectations, and 3) to provide one-on-one consultation with selected United Way agencies.

Summary of Intern’s Activities

One task addressed by this intern was to identify and obtain reliable, valid, affordable, and useful measurement tools for two of the four constructs chosen by Project EASe II staff from the list of constructs nominated by United Way agencies (stress and self-esteem).

Before beginning the search, Project EASe II staff established a set of criteria for choosing instruments that measure stress, self-esteem, parenting skills, and family functioning. It was determined that measures needed to be inexpensive enough to be affordable for United Way agencies to purchase. It was also felt that it was important that instruments be easy to administer and score in terms of time and level of understandability. In addition to these practical considerations, it was determined that the instruments had to have proof of reliability and validity. Reliability refers to the consistency of the measure. Some things that can threaten the reliability of an instrument are questions that are not understandable, questions that are ambiguous, questions that do not take into account the cultural context of the participant, and/or too many questions on the instrument. Validity refers to the trustworthiness of the instrument in terms of getting accurate answers from the participant. Some things that can threaten validity are participants’ concerns about the
confidentiality of their answers, and social desirability of answers (the participant answers in the way he/she feels the question should be answered rather than truthfully).

After an extensive library search, the following instruments were chosen.
Stress

The Perceived Stress Scale

This instrument is designed to be used as "an outcome measure of experienced levels of stress" (Cohen, Kamarck, & Mermelstein, 1983, pg. 385). This, more generally stated, means that it evaluates the "degree to which situations in one's life are appraised as stressful" (Cohen, Kamarck, & Mermelstein, 1983, pg. 385). This instrument could be useful to practitioners in determining how "stressed out" clients perceive themselves to be, as well as help to predict how clients will perceive other stressful situations in the future.

This instrument was chosen based on many factors. It is a short instrument that is easy to understand and score. The items and scoring criteria are provided in the authors' article which makes the cost of this instrument minimal. The general nature of the questions on the instrument reduce cultural biases, which can be a threat to validity. This means that it can be used with many different cultural backgrounds. These general questions also reduce biases of gender or age.

In summary, this instrument can be used with many populations for minimal cost. It is easy to administer and score and provides useful information for agencies interested in the stress level of their clients.


This is a paper which presents reliability and validity information on the Perceived Stress Scale. It provides information on the correlations it has with other measurement types, as well as providing some ideas for its predictive value.
A Global Measure of Perceived Stress

SHELDON COHEN
Carnegie-Mellon University

WILLIAM KAMARCK
University of Oregon

ROBIN MERMELSTEIN
University of Oregon


This paper presents evidence from three samples, two of college students and one of participants in a community smoking-cessation program, for the reliability and validity of a 14-item instrument, the Perceived Stress Scale (PSS), designed to measure the degree to which situations in one's life are appraised as stressful. The PSS showed adequate reliability and, as predicted, was correlated with life-event scores, depressive and physical symptomatology, utilization of health services, social anxiety, and smoking-reduction maintenance. In all comparisons, the PSS was a better predictor of the outcome in question than were life-event scores. When compared to a depressive symptomatology scale, the PSS was found to measure a different and independently predictive construct. Additional data indicate adequate reliability and validity of a four-item version of the PSS for telephone interviews. The PSS is suggested for examining the role of nonspecific appraisal of stress in the etiology of disease and behavioral disorders and as an outcome measure of experienced levels of stress.

It is a common assumption among health researchers that the impact of "objectively stressful" events is, to some degree, determined by one's perceptions of their stressfulness; e.g., see Lazarus (1966, 1977). Surprisingly, this theoretical perspective has not been accompanied by development of psychometrically valid measures of perceived stress. This article discusses the limitations of objective and subjective measures of stress used in the assessment of both global and event-specific stress levels. It argues that a psychometrically sound global measure of perceived stress could provide valuable additional information about the relationship between stress and pathology. Data are presented on the psychometric properties of the Perceived Stress Scale (PSS), an instrument developed in response to these issues. The PSS measures the degree to which situations in one's life are appraised as stressful.

Research reported in this paper was supported by grants from the National Science Foundation (BNS 7923453) and the Heart, Lung and Blood Institute (HL 29547). The authors are indebted to Edward Lichtenstein and Karen McIntyre for their collaboration on the smoking cessation project; the staff of University of Oregon dormitory housing for their cooperation in securing Student Sample II; Pam Birell for her cooperation in obtaining class participants for Sample III; and to Susan Fiske, Lew Goldberg, Edward Lichtenstein, Karen Matthews, Michael Scheier, and Tom Wills for comments on earlier drafts of this manuscript.

Address communications to: Sheldon Cohen, Department of Psychology, Carnegie-Mellon University, Pittsburgh, PA 15206.
Research on the role of psychosocial and environmental stressors as risk factors in both physical and psychological illness has typically employed relatively objective stressor measures. This work includes studies of the effects of specific stressful events, such as unemployment (Cobb and Kasl, 1977; Dooley and Catalano, 1980; Gore, 1978), bereavement (Stroebe et al., 1982), and exposure to intense levels of noise (Cohen and Weinstein, 1981) and high levels of population density (Sundstrom, 1978). In addition, there is an abundant literature on the cumulative effect of objective stressful life events (Dohrenwend and Dohrenwend, 1974, 1981). In these studies, various versions of life-event scales (original scale developed by Holmes and Rahe, 1967) are used to produce a cumulative stress score. These scores are usually based on either the number of events that have occurred within the specified temporal framework (in most cases, six to 12 months) or on a sum of event weights that are based on judges’ ratings of the difficulty of adjusting to these events.

There are some clear advantages to objective measures of stressful events. First, such measures permit an estimate of the increased risk for disease associated with the occurrence of easily identifiable events. Second, the measurement procedure is often simple, e.g., did this event occur during the last six months? and in many cases, persons experiencing a particular event can be identified without the necessity of asking them about the occurrence of the event, e.g., persons living in noise-impacted communities. Third, these measurement techniques minimize the chance of various subjective biases in the perceptions and reporting of events.

On the other hand, the use of objective measures of stress implies that events are, in and of themselves, the precipitating cause of pathology and illness behavior. This implication is counter to the view that persons actively interact with their environments, appraising potentially threatening or challenging events in the light of available coping resources (Lazarus, 1966, 1977). From this latter perspective, stressor effects are assumed to occur only when both (a) the situation is appraised as threatening or otherwise demanding and (b) insufficient resources are available to cope with the situation. The argument is that the causal “event” is the cognitively mediated emotional response to the objective event, not the objective event itself (Lazarus, 1977; Mason, 1971). An important part of this view is that this response is not based solely on the intensity or any other inherent quality of the event, but rather is dependent on personal and contextual factors as well.

The assumed centrality of the cognitive appraisal process suggests the desirability of measuring perceived stress as opposed or in addition to objective stress. For example, comparison of the predictive validities of objective and subjective measures could help to clarify the role of the appraisal process in the relationship between objective stressors and illness. Perceived stress scales could also be used in conjunction with objective scales in an effort to determine whether factors such as social support (Pearlin et al., 1981), hardiness (Kobasa, 1979), and locus of control (Johnson and Sarason, 1979) protect people from the pathogenic effects of stressful events by altering stressor appraisal or by altering the process or processes by which appraised stress results in physiological or behavioral disorders (Gore, 1981). Finally, perceived stress can be viewed as an outcome variable—measuring the experienced level of stress as a function of objective stressful events, coping processes, personality factors, etc.

Previous work has employed a number of approaches to assess both global and event-specific levels of perceived stress. For example, several investigators have modified event scales in an attempt to measure global perceived stress. The modification involved asking respondents to rate the stressfulness or impact of each experienced event. In general, life-stress scores based on self-ratings of event stressfulness are better predictors of health-related outcomes than are scores derived from either a simple counting of events (unit-weighting) or normative adjustment ratings (Sarason et al., 1978; Vinokur and Selzer, 1975). However, the increases in predictability provided by these ratings are small. It is noteworthy that any increase in predictability of a weighted event score over a simple count of events is likely to be small since alternative weighting schemes yield composite scores that are substantially correlated with the event count (Lei and Skinner, 1980). In short, cal-
A Global Measure of Perceived Stress

Calculating global perceived stress levels on the basis of reactions to individual events assumes that perceived stress levels are very highly correlated with the number of reported events. Other weaknesses of global perceived stress scales that are based on a specific list of events include an insensitivity to chronic stress from ongoing life circumstances, to stress from events occurring in the lives of close friends and family, from expectations concerning future events, and from events not listed on the scale.

Subjective measures of response to specific stressors have also been widely used, e.g., measures of perceived occupational stress (Kahn et al., 1964). There are, however, some practical and theoretical limitations of measures of specific stressors. Practically, it is difficult and time-consuming to adequately develop and psychometrically validate an individual measure every time a new stressor is studied. Theoretically, there are an issue of whether measures of perceived response to a specific stressor really assess a person's evaluations of that stressor. There is, in fact, evidence that people often misattribute their feelings of stress to a particular source when that stress is actually due to another source (Gochman, 1979; Keating, 1979; Worchel, 1978; Worchel and Teddie, 1976). Another problem with measures of response to specific stressors is that such measures imply the independence of that event in the precipitation of disease. However, it is likely that the illness process is affected by a person's global stress level, not by his/her response to a particular event.

The above discussion indicates the desirability of developing an instrument to measure a global level of perceived stress. This article presents data on the Perceived Stress Scale, a 14-item measure of the degree to which situations in one's life are appraised as stressful. PSS items were designed to tap the degree to which respondents found their lives unpredictable, uncontrolable, and overloading. These three issues have been repeatedly found to be central components of the experience of stress (Averill, 1973; Cohen, 1978; Glass and Singer, 1972; Lazarus, 1966, 1977; Seligman, 1975). The scale also includes a number of direct queries about current levels of experienced stress. The PSS is an economical scale that can be administered in only a few minutes and is easy to score. Because levels of appraised stress should be influenced by daily hassles, major events, and changes in coping resources, the predictive validity of the PSS is expected to fall off rapidly after four to eight weeks.

Evidence is presented from three samples—two of college students and one of a more heterogeneous community group—for the concurrent and predictive validities and the internal and test-retest reliabilities of the new scale. The paper also examines the relative predictive validity of the PSS and two life-event instruments. Our premise is that the PSS should provide a better predictor of health outcomes than does a global measure of objective stressors, such as life-event scales. This should occur because a perceived stress instrument provides a more direct measure of the level of stress experienced by the respondent. Presumably, it is this level of appraised stress, not the objective occurrence of the events, that determines one's response to a stressor(s) (Lazarus, 1966, 1977). Also, the new measure is more global than life-event scales. That is, it is sensitive to chronic stress deriving from ongoing life circumstances, to stress from expectations concerning future events, to stress from events not listed on a particular life-events scale, and to reactions to the specific events included on any scale.

METHODS

Validation data were collected in three samples—two consisting of college students and one consisting of a more heterogeneous group enrolled in a smoking-cessation program. The samples and assessment procedures are described below.

Perceived Stress Scale (PSS)

The 14 items of the PSS are presented in...
education. The items are easy to understand and the response alternatives are simple to grasp. Moreover, as noted above, the questions are quite general in nature and hence relatively free of content specific to any subpopulation group.

The data reported in this article are from somewhat restricted samples, in that they are younger, more educated, and contain fewer minority members than the general population. In light of the generality of scale content and simplicity of language and response alternatives, we feel that data from representative samples of the general population would not differ significantly from those reported below.

**College Student Sample I**

The respondents were 332 (121 male, 209 female, two with sex not specified) freshman college students living in dormitories at the University of Oregon. The mean age of the sample was 19.01 with a standard deviation of 2.75. All respondents gave written consent allowing access to their student health center records.

**Measures.** Respondents completed five scales: one measured life events; another social anxiety; a third depressive symptomatology, fourth, physical symptomatology; and finally perceived stress (the PSS). All instruments were completed during a one and one-half hour session.

A modified version of the College Student Life-Event Scale (CSLES) was used as a measure of stressful life events; the original scale was developed by Levine and Perkins (1980). This scale is composed of 99 items that represent events that fall into 14 different categories characterizing the adjustment demands of college students, e.g., academic affairs, male-female relationships, and family matters. Nine items dealing with health-related issues were not used in calculating life-stress scores because of the possibility that these items were measuring the same thing as items in the symptom checklists. Analyses including the unused items indicated that their exclusion did not affect the results reported below.

Respondents were asked to indicate whether each event had occurred during the last year. They were asked to rate the impact of events that had occurred on a scale ranging from -3 (extremely negative) to +3 (extremely positive), the format used by Sarason et al. (1978). Separate scores were generated based self-rated impact and unweighted events. The unweighted score was the total number of life events checked. The score based on impact ratings was the summed impact of checked events. The impact score is not a pure measure of the occurrence of objective events, but rather takes into account the respondents' perceptions of the events.

The Center for Epidemiologic Studies Depression Scale (CES-D) was employed as a measure of current level of depressive symptomatology (Radloff, 1977). Twenty items, each representing a state characteristic or not characteristic of a depressed person, are rated on a four-point scale to indicate the frequency of their occurrence during the last week. Response options range from "rarely or none of the time" to "most or all of the time."

Physical symptomatology was measured by the Cohen-Hoberman Inventory of Physical Symptoms (CHIPS). The CHIPS is a list of 39 common physical symptoms (Cohen and Hoberman, 1983). Items were carefully selected to exclude symptoms of an obviously psychological nature, e.g., felt nervous or depressed. The scale, however, is primarily made up of symptoms, such as headache, back ache, acid stomach, that have been traditionally viewed as psychosomatic. Each item is rated for the degree to which that problem bothered or distressed the individual during the past two weeks. Items are rated on five-point scale from "not at all" to "extremely." The CHIPS has been found to have adequate reliability and to predict use of student health services in the seven-week period following completion of the scale (r = .29 and .22 in independent samples).

The Social Avoidance and Distress Scale (SADS) was used to measure social anxiety (Watson and Friend, 1969). This 28-item true-false scale taps both the desire to avoid others (social avoidance) and the experience of distress in social interactions (social distress).

Utilization of the student health center was also monitored. All university students are required to pay a fee that provides for outpatient medical care. The physicians at the student health center routinely fill out a standardized form describing the problem(s) for which the student was treated at each based on the International Diseases, 9th Revision (Commission and Hospital Activity) each visit to be classified (codes 001-779), injury- and (codes 800-999), or "other receive prophylactic vaccine"

For each student, the mean of each of the three categories to three symptoms or problems were all in the same category for the problems were in different categories. When total visit categories were calculated, counted as one, irrespective of whether a single problem or a problem in different categories. Bothness visits and the number analyzed in this report.

The number of visits was of two independent time periods preceding the testing session, following the testing session period was used as an indication of visits.

**College Student Sample II**

Respondents in the second members of a class in intro- psychology (53 females, 60 sex not specified) who received five items to participating in the study. A sample was 20.75 with a standard deviation of 4.41. These students completed questionnaires as those in a during a one and one-half the second week of the S this sample, the data on questionnaires in the testing session divided into the testing session and the testing session. The 90th as an indicator of the base

**Smoking-Cessation Sample**

**Subjects.** Subjects were males participating in a program run by the University
A Global Measure of Perceived Stress

student was treated at each visit. The form is based on the International Classification of Diseases, 9th Revision (Commission on Professional and Hospital Activities, 1980) and allows each visit to be classified as illness-related (codes 001–779), injury- and poisoning-related (codes 800–999), or “other” (V codes), e.g., receive prophylactic vaccination.

For each student, the number of visits in each of the three categories was recorded. Up to three symptoms or problems could be checked for any individual visit. If the problems were all in the same category, only one visit was recorded for that category. If the problems were in different categories, a separate visit was recorded for each of the pertinent categories. When total visits (collapsing over categories) were calculated, each visit was counted as one, irrespective of whether there was a single problem or a number of problems in different categories. Both the number of illness visits and the number of total visits are analyzed in this report.

The number of visits was calculated for each of two independent time blocks: the 44 days preceding the testing session and the 46 days following the testing session. The initial 44-day period was used as an indicator of the base-rate of visits.

College Student Sample II

Respondents in the second sample were 114 members of a class in introductory personality psychology (53 females, 60 males, and one with sex not specified) who received class credit for participating in the study. The mean age of the sample was 20.75 with a standard deviation of 4.41. These students completed the same five questionnaires as those in the previous study during a one and one-half hour session during the second week of the Spring Quarter. For this sample, the data on health services utilization were divided into the 90 days preceding the testing session and the 46 days following the testing session. The 90-day period was used as an indicator of the base-rate of visits.

Smoking-Cessation Sample

Subjects. Subjects were 27 males and 37 females participating in a smoking-cessation program run by the University of Oregon Smok-

ing-Control Program. Participants were solicited through newspaper, television, and radio advertisements. To qualify for participation in the program, subjects either had to be married or living with a partner. The mean age of the subjects was 38.4 years (s = 11.57). Thirty-seven percent of the sample made over $25,000 per year, and 74 percent had formal education beyond high school. They had been smoking for an average of 20.9 years (s = 11.82). Only three of the participants were students; all of these were graduate students, 25 to 31 years old. The mean self-monitored baseline smoking rate was 25.6 cigarettes a day.

Treatment. Treatment groups met for six consecutive weekly sessions lasting approximately two hours each. The target quitting date occurred on the fourth session. Intervention strategies included behavioral self-management techniques, nicotine-fading, and a cognitive-behavioral relapse prevention program. Overall, 64 percent (N = 41) of the subjects were abstinent at the end of treatment.

Measures. During a pretreatment testing session, subjects completed a life-event scale, a physical-symptom checklist (CHIPS), and the PSS. The life-event scale consisted of 71 normatively negative events chosen from the Unpleasant Events Schedule (Lewinsohn and Talkington, 1979). This scale replaced the one used with student samples because it provided an item pool appropriate to the community population. Subjects identify the events that have happened to them in the last six months and rate the impact of each event on a seven-point scale, ranging from extremely negative to extremely positive. As in the college student samples, both the number of events and the event-impact scores were analyzed.

Subjects completed the PSS and the CHIPS prior to the first treatment session and at the end of the six-week treatment. Both tests were administered in the same manner as with the college student samples, except that with the group of smokers, a one-week time frame was used for the CHIPS.

RESULTS

Means, Variance and Reliability Estimates

Mean scores on the PSS for the complete samples (males and females combined) were
23.18 and 23.67 in the student samples and 25.0 in the smoking-cessation sample. Standard deviations were 7.31, 7.79, and 8.00, and ranges were 6 to 50, 5 to 44, and 7 to 47. Mean PSS scores for females were 23.57 and 25.71 in the student samples and 25.6 in the community sample. Standard deviations were 7.55, 6.20, and 8.24. Mean PSS scores for males were 22.38 and 21.73 in the student samples and 24.0 in the community sample. Standard deviations were 6.79, 8.42, and 7.80, respectively. Although the mean PSS score for females was slightly higher than the mean score for males in all three samples, this difference did not approach statistical significance in any sample.

Age was unrelated to PSS in all three samples. Since the age distribution in the college student samples was severely skewed, a correlation between the PSS and age was unlikely. The correlations between age and PSS were .04 and .08 in the college samples and .02 in the smoking-cessation sample.

Coefficient alpha reliability for the PSS was .84, .85, and .86 in each of the three samples. For a state measure, test-retest correlations should be much higher for short retest intervals than for longer ones. For the PSS, two intervals are available, two days and six weeks. The PSS was administered, on two occasions separated by two days, to 82 college students enrolled in courses at the University of Oregon. When responding to the retest, the subjects were asked to strive for accuracy rather than for consistency across time. The test-retest correlation in this sample was .85, whereas the correlation was only .55 for the 64 subjects in the smoking study who were retested after six weeks.

Evidence for Concurrent and Predictive Validity

Separate correlations between the PSS and validity criteria were calculated for males and females in each sample. The relative magnitudes of the correlations for males and females with each of the criteria were compared by transforming the correlations to z scores (Fisher’s transformation) and dividing the difference between the z scores by the standard error of the difference between the z coefficients (cf. Guilford, 1965). None of the resulting z scores were significantly different from 0 at the p < .05 level. Since there were no differences between males and females, only data for the entire sample are reported below.

Separate correlations between the PSS and validity criteria were also calculated for those below and above median age in the smoking cessation sample. These analyses seemed unnecessary in the student samples, since 98 percent of the students were between 16 and 25 years old. The 31 persons in the “young” community group ranged from 22 to 35. The 33 persons in the “old” group ranged from 36 to 70. Correlations for young and old were compared with the same procedure described in the foregoing. Only one of these comparisons indicated a statistically significant difference at the p < .05 — actually p < .01 — level. Data on this comparison are presented in the appropriate section in the following and addressed in the discussion section.

Correlations between PSS and Life-Event Scores. Since perceived stress should generally increase with increases in objective cumulative stress levels, the PSS should be related to the number of life events. Moreover, these correlations should be higher when the life-event scores are based on the self-rated impact of the events, since impact scores reflect some of the same stressor appraisal measured by the PSS. As apparent from Table 1, there is a small to moderate correlation between number of life events and the PSS in all three samples. Moreover, in all but one case, that correlation increases when the scale score takes into account the respondent’s perception of the events. Hotelling’s statistical significance of correlations, indicate the significant (p < .05) for the smoking-cessation sample of treatment.

There was a difference between PSS and number of life events for young and old participants. For the sample was .65 (p < .05); in the smoking-cessation sample. As expected, the PSS woul...
A Global Measure of Perceived Stress

TABLE 2. Correlations of Stress Measures with Depressive Symptomatology

<table>
<thead>
<tr>
<th></th>
<th>College Student Sample I</th>
<th>College Student Sample II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Events</td>
<td>.18*</td>
<td>.14</td>
</tr>
<tr>
<td>Impact of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Events</td>
<td>.29*</td>
<td>.33*</td>
</tr>
<tr>
<td>Perceived</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress Scale</td>
<td>.76*</td>
<td>.65*</td>
</tr>
</tbody>
</table>

* p < .001.

of the events. Hotelling t-tests, testing the statistical significance of differences between correlations, indicate that this increase is significant (p < .05) for Student Sample I and for the smoking-cessation sample at the beginning of treatment.

There was a difference in the correlation between PSS and number of negative events for young and old participants in the smoking cessation sample. For the young, the correlation was .65 (p < .05); for the old, it was .19.

**PSS Versus Life Events as a Predictor of Symptomatology.** As noted earlier, we expected that the PSS would be a better predictor of the various health outcomes than would stressful life-event scores. The data presented in Tables 2 and 3 support this prediction in the case of both depressive and physical symptomatology. Hotelling t-test (p < .05) provide statistical support for these differences in all cases. Since these are cross-sectional correlations, no causal inferences are implied. For example, it is possible that increased symptomatology caused increased stress, rather than that the stress caused the symptomatology.

In regard to establishing the validity of the PSS, it is important to note the substantial correlations between the scale and both symptomatology measures. There is probably some overlap between what is measured by the depressive symptomatology scale and what is measured by the PSS, since the perception of stress may be a symptom of depression. This may, to some degree, account for the magnitude of that correlation. In light of the very high correlation between the PSS and the CES-D (depressive symptom scale), it is desirable to demonstrate that these scales are not measuring the same thing. Hence, partial correlations were calculated: in these, depressive symptomatology was partialled out of the correlations between the PSS and physical symptomatology, and the PSS was partialled out of the correlation between depressive symptomatology and physical symptomatology. In the case of PSS and physical symptomatology, the correlation was .16, p < .01, for sample I and .17, p < .07, for sample II. In the case of the CES-D and physical symptomatology, the correlation was .31, p < .01, in sample I and .38, p < .01, in sample II. Hence, even with the very high correlation between the PSS and CES-D, both scales still independently predicted physical symptomatology.

**PSS Versus Life Events as a Predictor of Utilization of Health Services.** Table 4 presents the correlations between the PSS and utilization of health services, both before and after administration of the scale. In Sample I, the PSS significantly predicted utilization during the five-week period after completing the

<table>
<thead>
<tr>
<th></th>
<th>College Student Sample I</th>
<th>College Student Sample II</th>
<th>Smoking-Cessation Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Events</td>
<td>.31</td>
<td>.36</td>
<td>.40</td>
</tr>
<tr>
<td>Impact of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Events</td>
<td>.23</td>
<td>.32</td>
<td>.51</td>
</tr>
<tr>
<td>Perceived</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress Scale</td>
<td>.52</td>
<td>.65</td>
<td>.70</td>
</tr>
</tbody>
</table>

* p < .001 for all correlations.

TABLE 3. Correlations of Stress Measures with Physical Symptomatology

TABLE 4. Correlations of PSS and Health Center Utilization Before and After Completing Scale

<table>
<thead>
<tr>
<th></th>
<th>Student Sample I</th>
<th>Student Sample II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before Scale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Illness Visits</td>
<td>.08</td>
<td>-.06</td>
</tr>
<tr>
<td>All Visits</td>
<td>.11*</td>
<td>-.05</td>
</tr>
<tr>
<td><strong>After Scale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Illness Visits</td>
<td>.17**</td>
<td>.04</td>
</tr>
<tr>
<td>All Visits</td>
<td>.20***</td>
<td>.12</td>
</tr>
</tbody>
</table>

* p < .05.
** p < .01.
*** p < .001.
scale. In Sample II, there was a nonsignificant correlation between the PSS and all visits after administration of the scale. Correlation between the PSS and physical illness visits after administration of the scale, with before-administration visits partialled out, was .15, p < .007, for Sample I and -.02 for Sample II. Similar partial correlation for the PSS and all visits after administration, with all preadministration visits partialled out, was .16, p < .01, for Sample I and .13 for Sample II. These correlations suggest that the PSS is predictive of changes in health center utilization. Correlations of life-event scores with utilization were not significant in both samples for both physical illness and all visits: these correlations ranged from -.04 to +.03.

**PSS Versus Life Events as a Predictor of Social Anxiety.** Levels of perceived stress in college students, especially freshmen, should be related to their ability to become integrated into the university community. We would expect that the poorer the integration, the greater the perceived stress. The social anxiety scale provides a trait measure that presumably taps difficulty in making friends and social contacts, i.e., the ability to integrate into the community.

In both student samples, increases in social anxiety were associated with increases in perceived stress (.37 and .48, p < .001 for both). Although number of life events was unrelated to social anxiety in both samples, there were small correlations between event impact and social anxiety (.13, p < .02, -.26, p < .01). Again, these are cross-sectional data and no causal inferences are implied.

**PSS and Smoking-Reduction Maintenance:** The Four-Item Scale. To examine the role of the PSS as a predictor of maintenance of smoking-rate reduction, perceived stress level was also assessed one and three months following treatment. Since posttreatment data were collected by telephone interview, a short version of the scale, consisting of the four items (numbers 2, 6, 7, and 14) that were correlated most highly with the 14-item scale, was employed. The mean score on the four-item scale was 5.6 at one month and 5.9 at three months. Standard deviations were 3.6 and 4.0, and the scores ranged from 0 to 15 and 0 to 14. Mean PSS scores for males were 4.8 at one month and 5.9 at three, while mean scores for females were 6.2 and 5.9, respectively. The coefficient alpha reliability estimate for the four-item PSS was .72. The test-retest reliability of the four-item scale over a two-month interval was .55.

The four-item scale at one month was correlated (.31, p < .01) with the average number of cigarettes smoked per day at one month; the latter information was from a self-report confirmed with a carbon monoxide measure. Similarly, the four-item scale at three months was correlated (.37, p < .001) with smoking rate at three months. Finally, the short PSS scale at one month after treatment predicted smoking rate three months after treatment (.39, p < .001). In all cases, the greater the PSS score, the more cigarettes smoked. A number of partial correlations were calculated to clarify the nature of the relationship between the abridged version of the PSS and smoking rate. In the first partial correlation, the end-of-treatment PSS score and smoking rate were partialled out of the correlation between the four-item PSS and smoking rate at one month following treatment. The partial correlation was .29, p < .01. In the second correlation, the four-item PSS and smoking rate at one month following treatment were partialled out of the correlation between the four-item scale and smoking rate at three months. The partial correlation was .34, p < .01. These analyses indicate that changes in perceived stress as measured by the PSS are predictive of changes in smoking rate. Another partial correlation indicated that the abridged PSS predicted changes in smoking rate over a two-month period. Specifically, a correlation of .26 (p < .05) was found when smoking rate at the one-month follow-up was partialled out of the correlation between the four-item PSS at one month and smoking rate at three months. These data suggest that the four-item scale provides a useful measure of perceived stress for use in telephone interviews and other situations where a very short scale is required.

**DISCUSSION**

The PSS has adequate internal and test-retest reliability and is correlated in the expected manner with a range of self-report and behavioral criteria. Moreover, the PSS is more closely related to a life-event impact score which is to some degree base student's appraisal of the event, objective measure of the num curring within a particular tin also proved to be a better predictor and health-related outcomes: two life-event scales. Final though highly correlated symptomatic, was found different and independently present.

It is noteworthy that the relationship between the life-event scale and symptomatology outcomes (.18 equivalent to, if not better than, notions reported in the literature and Stroten, 1976; Tausig, superior predictability of the usable to psychometric weak event scales that were ecritatic aspects of the sample. Moreover, from an absol PSS correlations with sympto scores are quite high (.52 to the case of depressive symptoms correlation may be somewhat overlap in the operational ceived stress and of symptomatology (cf. Dohrenwend 1981).

In general, the relations and the validity criteria were or age. The one exception relationship between the PSS events for the young and relationship for the old. The a difference in the role of mining stress levels for the That is, other chronic stress etc. may be more importan respondents.

The PSS differs from li numbers of ways. First, the shorter period, one month usual six to 12 months cov event scales. It is worth subjective scale, the short sufficient since perceived a month should reflect any care are still affecting respondents.

A second difference bet life-event scale is the per mi nistration that the scale of health-related outcome
A Global Measure of Perceived Stress

which is to some degree based on the respondent's appraisal of the event, rather than on the more objective measure of the number of events occurring within a particular timeframe. The PSS also proved to be a better predictor of health and health-related outcomes than either of the two life-event scales. Finally, the PSS, although highly correlated with depressive symptomatology, was found to measure a different and independently predictive construct.

It is noteworthy that the level of correlation between the life-event scales and the symptomatological outcomes (.18 to .36 range) is equivalent to, if not better than, similar correlations reported in the literature (cf. Rabkin and Struening, 1976; Tausig, 1982). Hence, the superior predictability of the PSS is not attributable to psychometric weaknesses in the life-event scales that were employed or to idiosyncratic aspects of the samples under study. Moreover, from an absolute perspective, the PSS correlations with symptomatological measures are quite high (.52 to .76). However, in the case of depressive symptomatology, the correlation may be somewhat inflated by the overlap in the operational definitions of perceived stress and of depressive symptomatology (cf. Dohrenwend et al., 1978; Gore, 1981).

In general, the relationships between PSS and the validity criteria were unaffected by sex or age. The one exception was the strong relationship between the PSS and number of life events for the young and the lack of such a relationship for the old. These data may reflect a difference in the role of life events in determining stress levels for these two age groups. That is, other chronic stressors, expectations, etc. may be more important for the older respondents.

The PSS differs from life-event scales in a number of ways. First, the PSS asks about a shorter period, one month as opposed to the usual six to 12 months covered by typical life-event scales. It is worth noting that with a subjective scale, the shorter period should be sufficient since perceived stress during the last month should reflect any objective events that are still affecting respondents' stress levels.

A second difference between the PSS and a life-event scale is the period of time after administration that the scale provides predictions of health-related outcomes. Presumably, life-event scales will be predictive over fairly long periods, such as several months to several years. We have examined the predictive ability of the PSS over four to 12 week periods after administration. These data suggest that the best predictions occur within a one- or two-month period. It is our feeling that as this period is lengthened, the predictive validity of the scale will fall. After all, perceived levels of stress should be influenced by daily hassles, major events, and changes in the availability of coping resources, all of which are quite variable over a short period. In fact, test-retest reliability analyses indicate that test-retests involving a very short time (two days) result in fairly substantial correlations, whereas administrations six weeks later produce more moderate test-retest correlations.

As mentioned earlier, the PSS can be used to determine whether "appraised" stress is an etiological (or risk) factor in behavioral disorders or disease. It can also be used to look more closely at the process by which various moderators of the objective stressor/pathology relationship operate. For example, we could determine whether social support protects one from the pathogenic effects of stressful events by altering the appraisal of those events or by altering the process by which appraised stress causes an illness outcome. This second kind of analysis, however, is limited to the degree that the PSS reflects responses to events outside of those measured by the objective event instrument. That is, it is limited to the degree that it is more global than the objective stressor measure. Finally, the PSS can be used as an outcome variable, measuring people's experienced levels of stress as a function of objective stressful events, coping resources, personality factors, etc.

The four-item version of the PSS provides a useful tool when data must be collected over the phone. This scale makes repeated measures of perceived stress in large samples feasible. It should be noted, however, that because of the limited number of items, the abridged scale suffers in internal reliability and thus provides a less adequate approximation of perceived stress levels than the entire scale.

Although not tested in this study, the PSS may also provide an economical tool for assessing chronic stress level. Either the abridged version of the PSS or the complete
14-item scale could be used. Monthly administrations of the scale could be summed or averaged, providing a reliable, i.e., based on more samples, measure of chronic stress, as well as a predictor that represents a longer term than the one-month period covered by the scale.

Two commonly used measures of nonspecific psychological distress, the nine-scale, 54-item PERI Demoralization Measure (Dohrenwend et al., 1980) and the 28-item General Health Questionnaire (Goldberg, 1972), include a number of items that are similar to those in the PSS. These scales, however, are much broader in scope. They are designed as epidemiological measures of symptomatology, and both include a broad range of items tapping common psychiatric symptoms, such as hostility, diminished self-esteem, depression, and anxiety, as well as psychosomatic complaints. Although appraised stress may be symptomatic of psychological disorder when viewed in combination with elevated scores on other psychiatric symptoms, it is our contention that the perception of stress itself, as assessed by the PSS, is not a measure of psychological symptomatology. This contention is, in fact, supported by the data indicating independent predictive validities of the PSS and a depressive symptomatology scale. Hence, the PSS can be viewed as assessing a state that places people at risk of, i.e., is antecedent to, clinical psychiatric disorder even though that state is also part of a diverse set of feelings and states that are characteristic of disorder.

In sum, the PSS is a brief and easy-to-administer measure of the degree to which situations in one's life are appraised as stressful. It has been proven to possess substantial reliability and validity; thus, it provides a potential tool for examining issues about the role of appraised stress levels in the etiology of disease and behavioral disorders.

NOTES

1. There are two recently developed life-event scales that assess events over a one-month period, the Hassle Scale (Kanner et al., 1981) and the Unpleasant Events Scale (Lewinsohn and Tallington, 1979).
2. Pearlman et al. (1981) have pointed out that there is no single time frame that is optimal for observing the effects of diverse life events. However, existing works, including the Pearlman et al. study, examine rather long periods.

APPENDIX A:

Items and Instructions for Perceived Stress Scale

The questions in this scale ask about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between them and you should treat each one as a separate question. The best approach is to answer each question fairly quickly. That is, don't try to count up the number of times you felt a particular way, but rather indicate the alternative that seems like a reasonable estimate.

For each question choose from the following alternatives:

0. never
1. almost never
2. sometimes
3. fairly often
4. very often

1. In the last month, how often have you been upset because of something that happened unexpectedly?
2. In the last month, how often have you felt that you were unable to control the important things in your life?
3. In the last month, how often have you felt nervous and "stressed"?
4. In the last month, how often have you dealt successfully with irritating life hassles?
5. In the last month, how often have you felt that you were effectively coping with important changes that were occurring in your life?
6. In the last month, how often have you felt confident about your ability to handle your personal problems?
7. In the last month, how often have you felt that things were going your way?
8. In the last month, how often have you found that you could not cope with all the things that you had to do?
9. In the last month, how often have you been able to control irritations in your life?
10. In the last month, how often have you felt that you were on top of things?

REFERENCES


Commission on Professional 1980 International Classification of Clinical Psychoi Ann Arbor, Michigan and Hospital


Dooley, D., and R. Catalone 1980 "Economic change in..."
A GLOBAL MEASURE OF PERCEIVED STRESS

APPENDIX A (Continued)

11. In the last month, how often have you been angered because of things that happened that were outside of your control?
12. In the last month, how often have you found yourself thinking about things that you have to accomplish?
13. In the last month, how often have you been able to control the way you spend your time?
14. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

* Scored in the reverse direction.

REFERENCES

Averill, J. R.

Cobb, Sidney, and Stanley V. Kasl

Cohen, S.

Cohen, S., and H. Hoberman

Cohen, S., and N. Weinstein

Commission on Professional and Hospital Activities

Dohrenwend, Barbara S., and Bruce P. Dohrenwend (eds.)

Dohrenwend, B. S., L. Krassof, A. R. Askenasy, and B. P. Dohrenwend

Dohrenwend, B. P., P. E. Shrout, G. Egri, and F. S. Mendelsohn

Dooley, D., and R. Catalano

Glass, David C., and Jerome E. Singer

Gochman, I.

Goldberg, David P.

Gore, S.


Guilford, Joy Paul

Holmes, T. H., and R. H. Rahe

Johnson, J. H., and I. G. Sarason

Kahn, Robert L., Donald M. Wolfe, Robert P. Quinn, J. Diedrick Sneek, and Robert A. Rosenthal

Kanner, A. D., J. C. Coyne, C. Schaefer, and R. S. Lazarus

Keating, J.
Kobasa, S. C.  

Lazarus, Richard S.  


Lei, H., and H. A. Skinner  

Levine, M., and D. V. Perkins  
1980 "Tailor making life events scale." Presented at the meeting of the American Psychological Association, Montreal.

Levinsohn, P. M., and J. Talkington  

Mason, J. W.  

Pearlin, L. I., M. A. Lieberman, E. G. Menaghan, and J. T. Mullan  

Rabkin, J. G., and E. L. Struening  

Radloff, L.  

Sarason, I. G., J. H. Johnson, and J. M. Siegel  

Seligman, Martin E. P.  

Stroebe, W., M. S. Stroebe, K. J. Gergen, and M. Gergen  

Sundstrom, E.  

Tausig, M.  

Vinokur, A., and M. L. Seltzer  

Watson, D., and F. Friend  

Worchel, S.  

Worchel, S., and C. Teddie  

Below are the names:  
1983. Without the able called upon with burde tasks of the journal. I:

Nancy Adler  
Marilyn L. Bowman  
B. Bradford Brown  
John G. Bruhn  
Robert H. Coombs  
Sandra K. Danziger  
Maradee A. Davis  
Alfred Dean  
David Dooley  
John W. Finney, Jr.  
Joanne C. Gersten  
Charles W. Given  
Mary E. W. Goss  
Martin D. Hanlon  
Joseph M. Holtzman  
Howard B. Kaplan
The State-Trait Anxiety Inventory

This is an extremely well-known and widely used instrument designed to measure both the level of anxiety at the time of testing, as well as the general proneness to anxiety that the respondent possesses in terms of his/her personality and/or general make-up. A limitation to this instrument is that it is easy to "fake" or give false answers. Agencies must be aware of this, and do their best to determine which of their clients would be most likely to attempt this, and in which direction.

This particular instrument is available in both English and Spanish versions, and is relatively low in cost. It takes approximately 10-20 minutes to complete the test. It can be used with populations in the 9th grade or above.

It was chosen based on its extensive reputation and popularity in the scholarly literature. One other characteristic that helped this instrument make it to the list of finalists is it's availability in a second language.

This instrument could provide useful information to agencies on their clients level of stress as both an emotional state and a personality trait, providing that the administrators are aware of the limitations of the instrument.

For more information, see attached: see Bucky, S. F., Speilberger, C.D., & Bale, R.M. (1972).


This paper presents a study conducted on how instruction affects flight students' levels of both state and trait anxiety. It provides an example of how the State-Trait Anxiety Inventory can be used to differentiate between state and trait anxiety.

Murphy, L.L., Conoley, J.C., & Impara, J.C. (Eds.). (1994). Tests in print IV: An index to tests, test reviews, and the literature on specific tests. Lincoln, Nebraska: The University of Nebraska Press.
This presents a brief description of the test, as well as provides additional reference information for anyone interested in learning more about this specific instrument.


Provides a brief description of the test, references for additional information, and a write up evaluating the strengths and weaknesses of this measure.
**SELF-EVALUATION QUESTIONNAIRE**

Developed by Charles D. Spielberger  
in collaboration with  
R. L. Gorsuch, R. Lushene, P. R. Vagg, and G. A. Jacobs  
STAI Form Y-1

| Name ___________________________________________ | Date __________ S ___ |
| Age __________ Sex: M ____ F _____ | T ___ |

DIRECTIONS: A number of statements which people have used to describe themselves are given below. Read each statement and then blacken in the appropriate circle to the right of the statement to indicate how you feel right now, that is, at this moment. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I feel calm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I feel secure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. I am tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. I feel strained</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. I feel at ease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I feel upset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. I am presently worrying over possible misfortunes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. I feel satisfied</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. I feel frightened</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. I feel comfortable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. I feel self-confident</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. I feel nervous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. I am jittery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. I feel indecisive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. I am relaxed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. I feel content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. I am worried</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. I feel confused</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. I feel steady</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. I feel pleasant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Short Notes

Interestingly, the knowledge of the layoff was not detrimental to the companies. There appears to be no reason why management cannot be more liberal in granting of advance notice of layoff.

The results bear out this unorthodox recommendation. Managements are not unaware that a short notice of layoff does hurt employee feelings and morale. Yet, they have traditionally felt that such a hurt was less important than the expected increased costs due to production drop and the increased lateness and absenteeism.

Having already described what happens when workers hear a layoff, an important question deserves further consideration. Are there any limitations in accepting the results of the study?

The temptation to generalize from the results and then become predictive must be tempered by the realization that there may be a factor common to these four companies that may make them unlike any other four companies. This immediately implies a type of management that is sympathetic to research and perhaps somewhat different from other managements. If they are different, then is not the working climate different? And if so, might not employee behavior as a result be different? There is no guarantee that the results obtained in any four companies would be similar to the results obtained from any other four companies.

Further, the layoffs of these four companies occurred during a period when jobs were not too difficult to get. The question must then be raised as to whether or not the results would be the same in a tight labor market.

The results of this study raise questions concerning management's traditional practice of announcing layoffs as close as possible to the last day to be worked. In cases where organizations know who must be laid off over a long period of time, it would appear that it is mutually advantageous to give employees as much notice as possible.

(Received April 1, 1971)

Journal of Applied Psychology
1971, Vol. 56, No. 3, 175-176

EFFECTS OF INSTRUCTIONS ON MEASURES OF STATE AND TRAIT ANXIETY IN FLIGHT STUDENTS 1

STEVEN F. BUCKY 2

Nasal Aerospace Medical Institute, Pensacola, Florida

CHARLES D. SPIELBERGER

Florida State University

RONALD M. BALE

Nasal Aerospace Medical Research Laboratory, Pensacola, Florida

The State-Trait Anxiety Inventory was administered twice to 134 flight students during their routine entrance aviation physical examination, once with standard instructions and then under an artificially induced anxiety situation in which each student was instructed to respond "as if you had just made your first landing on an aircraft carrier." As expected, flight students were lower in trait anxiety and higher in state anxiety than a group of 253 previously tested male college students. Unexpectedly, however, with the induced anxiety act, both state and trait anxiety scores were significantly lower than the scores obtained during the initial administration of the test.

Recent research (Bucky, 1971) has suggested that flight students approach psychological testing with considerable apprehension and anxiety. Since these men are strongly motivated to become aviators, it is reasonable to expect that they may be somewhat defensive during the administration of such tests and hesitant to admit to symptomatology that may jeopardize their careers.

1 The opinions expressed herein are those of the author and not necessarily reflect the position of the United States Navy.

2 Requests for reprints should be sent to Steven F. Bucky, Aerospace Medical Institute, Naval Aerospace Medical Institute, Pensacola, Florida 32512.

Anxiety also may be a critical factor in determining the way in which the flight student will function as a student and later when he completes training. To evaluate the manifestation of anxiety, two different, though related, constructs are necessary—trait and state anxiety. According to Spielberger, Gorsuch, and Lushene (1970), trait anxiety (A-Trait) refers to relatively stable individual differences in anxiety proneness, that is, a disposition to perceive a wide range of circumstances as personally threatening. State anxiety (A-State) refers to a transitory emotional condition that is characterized by subjective feelings of tension.
and apprehension and heightened autonomic nervous system activity.

Based on Spielberger's theoretical discussion of state and trait anxiety, it was expected that the flight student tested early in his training would be lower in A-Trait but higher in A-State than a group of male college students. The A-Trait scores of the flight students were expected to be relatively low because flight students are a highly select group of individuals who have volunteered for a demanding aviation training program. Furthermore, they are also likely to have test-taking attitudes that lead them to consistently aim to look good. The prediction of relatively high A-State scores is based on the fact that the flight student is in a stressful situation, one in which a constant adjustment to a strenuous way of life must be made.

Spielberger et al. (1970) report that when giving the State-Trait Anxiety Inventory (STAI) with an induced anxiety set, A-State significantly increases. The authors initially administered the test to college students following the standard instructions and then were asked to respond according to how they believe they would feel "just prior to the final examination in an important course." If the flight student given an anxiety test with instructions to answer "as if you had just made your first landing on the aircraft carrier," it is expected that his A-State score will be markedly elevated since aircraft carrier landings are generally regarded as perhaps the most stressful part of the training program. However, it is hypothesized that A-Trait scores will remain the same because a personality trait should not change as a function of the circumstance under which the measure is obtained.

In this time "as if you had just made your first landing on an aircraft carrier."

Results and Discussion. The results in Table 1 indicate that when the STAI was administered with instructions, flight students had lower A-Trait scores (t = 3.233, p < .01) but higher A-State scores (t = 9.433, p < .01) than the group of college students. During the second administration of the test, the in which Ss had volunteered just landed on an aircraft carrier, A-Trait scores (t = 6.287, p < .01) and A-State scores (t = 2.579, p < .01) declined from those obtained with the first administration of the test.

The results provide several interesting implications for understanding what it is like to be a flight student. First, S's appear to be generally less anxious than college students, but they admit to higher levels of A-State in their present situation. When reported feelings associated with landing on an aircraft carrier, however, which bears directly on their future performance and careers, A-Trait scores declined (no significant change was predicted). What the flight student was being is, "While I am a relative unknown person, I feel quite anxious in my present stressful situation (being a flight student, adjusting to a way of life); but when I fly, particularly when I'm on an aircraft carrier, not only will I be a less anxious individual in general, but I will feel much less anxious than I am presently experiencing."

In conclusion, the results suggest that being a student is anxiety arousing for men who generally less anxious than the average college student. We confronted with a situation in which they feel particularly vulnerable and where their entire future may be at stake, the defensive tendency to "look good" increases and anxiety is denied. The probable presence of test-taking attitudes suggests that the relatively low A-State scores obtained during the initial administration of the STAI may actually underestimate the level of A-State experienced by flight students. Recent is presently being conducted in order to determine whether the STAI will be helpful in identifying the student who voluntarily drops out of the flight program.

REFERENCES


Values of the Subjective Prob

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>joint average</td>
<td>2.15</td>
<td>2.38</td>
</tr>
<tr>
<td>relative probability of success, per grade point average</td>
<td>.50</td>
<td>.60</td>
</tr>
</tbody>
</table>
State-Trait Anxiety Inventory

Purpose: Designed to assess anxiety as an emotional state (S-Anxiety) and individual differences in anxiety proneness as a personality trait (T-Anxiety).

Population: Grades 9-16 and adults.

Publication Dates: 1968-84.

Acronym: STAI.

Scoring: 2: State Anxiety, Trait Anxiety.

Administration: Group.

Forms: 2: X, Y.

Parts: 2: parts for each form labeled Form 1 (State), Form 2 (Trait).

Price Data: $5.50 per 25 expendable test booklets; $2 per scoring key; $9 per Spanish scoring key; $7 per 25 Spanish test booklets.


Time: (10-20) minutes.

Comments: Title on test is Self-Evaluation Questionnaire.

Authors: Charles D. Spielberger; Form Y and manual prepared in collaboration with R. L. Gorsuch, R. Lushene, P. R. Vagg, and G. A. Jacobs.

Publisher: Consulting Psychologists Press, Inc.

Cross References: See 9:1186 (158 references) and T2:2300 (277 references); for reviews by Ralph Mason Drager and Edward S. Katkin, see 8:683 (268 references); see also T2:1391 (45 references) and 7:141 (20 references).

TEST REFERENCES


THE EIGHTH MENTAL MEASUREMENTS YEARBOOK

REFERENCES

1-17. See 6:178.
18-35. See 5:245.
36-61. See 3:259; also includes a cumulative name index to the first 66 references and 2 reviews for this test.
62-70. See 7:159.
71-80. See 8:172.
81-84. See 9:183.
85-90. See 10:194.
91-94. See 11:205.
111-112. See 16:760.
113-114. See 17:881.
115-117. See 18:992.
118-119. See 19:1103.
120-121. See 20:1214.
122-123. See 21:1325.
126-127. See 23:1547.
128-129. See 24:1658.
132-133. See 26:1880.
136-137. See 28:2102.
138-139. See 29:2213.
140-141. See 30:2324.
142-143. See 31:2435.
144-145. See 32:2546.
146-147. See 33:2657.
148-149. See 34:2768.
150-151. See 35:2879.
152-153. See 36:2990.
154-155. See 37:3101.
156-157. See 38:3212.
158-159. See 39:3323.
160-161. See 40:3434.
162-163. See 41:3545.
164-165. See 42:3656.
166-167. See 43:3767.
168-169. See 44:3878.
170-171. See 45:3989.
172-173. See 46:4090.
174-175. See 47:4101.
176-177. See 48:4212.
180-181. See 50:4434.
182-183. See 51:4545.
184-185. See 52:4656.
188-189. See 54:4878.
190-191. See 55:4989.
194-195. See 57:5191.
196-197. See 58:5292.
198-199. See 59:5393.
200-201. See 60:5494.
202-203. See 61:5595.
204-205. See 62:5696.
206-207. See 63:5797.
208-209. See 64:5898.
210-211. See 65:5999.
216-217. See 68:6293.
220-221. See 70:6495.
222-223. See 71:6596.
224-225. See 72:6697.
226-227. See 73:6798.
228-229. See 74:6899.
230-231. See 75:6990.
232-233. See 76:7091.
234-235. See 77:7192.
236-237. See 78:7293.
238-239. See 79:7394.
240-241. See 80:7495.
244-245. See 82:7697.
246-247. See 83:7798.
248-249. See 84:7899.
250-251. See 85:7990.
252-253. See 86:8091.
254-255. See 87:8192.
256-257. See 88:8293.
258-259. See 89:8394.
260-261. See 90:8495.
262-263. See 91:8596.
264-265. See 92:8697.
266-267. See 93:8798.
268-269. See 94:8899.
270-271. See 95:8990.
high score suggests high state or trait anxiety. The manual presents a short discussion of the nature of state and trait anxiety, fairly clear instructions for administration and scoring, and norms for screened high school juniors, college freshmen, and college students enrolled in introductory psychology courses, neuropsychiatric patients, general medical and surgical patients, and prisoners. Norms are presented separately for the male and female students.

Test-retest reliabilities are reported for state (Form X-1) and trait (Form X-2) scores, separately by males and females, as follows—one-hour interval: .33 (males) and .16 (females) for state, .84 and .76 for trait; 20 days: .54 and .27 for state, .86 and .76 for trait; 104 days: .33 and .31 for state, .73 and .77 for trait. The high reliabilities (.84 and .76) for state scores—following one hour of “relaxation training; a difficult IQ test; and a film that depicted accidents resulting in serious injury or death”—suggest that Form X-2 (trait) is measuring considerable state anxiety, even allowing for theoretically predicted fluctuations in trait anxiety. Alpha reliability coefficients for the normative samples (high school juniors, college freshmen, introductory psychology students) range from .83 to .92 for state scores and .86 to .92 for trait scores; alpha coefficients are more suitable reliability indicators for X-1 than test-retest coefficients.

Validities for trait scores were estimated by correlating the scores with the IPAT Anxiety Scale, Manifest Anxiety Scale, and Affect Check List. For 126 college women, coefficients were .75, .80, and .52, respectively.

An important question, whether X-2 scores are related to everyday reality, and not just to other purported measures of reality, is confronted only tangentially in the manual. Though the high volume of research both prior and subsequent to the present STAI forms might suggest many assessments of validity, they cannot really be considered basic validity studies. (What would we think of a carpenter who says, “My tape measure indicates that this board, which has independently been proven to be five feet long, is five feet long; therefore, my tape measure is accurate.”)

The hand scoring keys are a bit cumbersome. The reviewer uses punched templates which fit over the response matrices with the appropriate forward and reverse scoring digits above the holes. Student scorers make fewer mistakes with these templates than with the keys provided with the manual.
measure is a valid measure of length?"

Instead, however, the authors do provide data from their standardization samples that show (by the reviewer's calculations) that the trait scores are related to real-life criteria. Normative means and standard deviations are given for specifically diagnosed neuropsychiatric patients, general hospital patients, and prisoners. The patients without psychiatric complications (at least, without secondary psychiatric diagnoses) can be compared with all other subgroups. As one might expect, in all the neuropsychiatric classifications but one (character disorders) the trait means for the neuropsychiatric patients are significantly above the means for the nonpsychiatric patients. Thus, on independently constituted normative samples, trait measures differ in the direction one would expect. (State comparisons cannot be made for these samples, for their specific present circumstances are not reported.)

Tempering these validity statements is the recognition that, as instruments that appear to be so patently what they purport to be, STAI is open to faking to a greater extent than more subtle pencil-and-paper tests. In fact, Smith (204) found that both STAI forms were markedly altered toward "faking bad" by simulated stress instructions. The presumption is that "faking good" would be just as easy. Therefore, the user needs to ascertain independently to what degree examinees would be likely to bias their responses and in which direction.

For the state measure, both item and total score comparisons between presumably stressful states are given, with mean scores reflecting changes in the appropriate directions. Alpha coefficients reported for stressful and nonstressful conditions remain substantially the same, again indicating internal consistency. Miscellaneous correlations of both forms of the STAI with several other (composite) temperament and self-assessment tests are not very informative, for the constructs in these tests are so conglomerate that high or low correlations can be interpreted to mean almost anything.

The huge number of studies inspired by or done under the senior author's direction, many of which this reviewer has examined, tend to indicate that the trait-state distinction is a valid one. A-trait seems to be a unitary dimension, though it may not encompass all there is of trait anxiety. A-state, however, is at least duplex (228, 297). This reviewer, in two studies not yet published, had 42 persons take the state form for 60 days, once a day at approximately the same time each day, and factored each subject's results by P-technique. Two to eight meaningful (but idiographic) factors were derived from each record. A comparison with factors from six randomly generated score matrices demonstrated the comparative "reality" of the real subjects' A-state factors. In no case was there a single factor that could be designated as the "A-state factor."

The revised STAI is one of the best standardized of anxiety measures, if not the best. Its faults are suggested in the previous paragraphs.

From the many references in the literature to the STAI it appears that it is a popular test. For instruments of its type it appears to be deservedly popular, in that the reliabilities are nearly as high as one would expect for intelligence scales; it demonstrates expected differences among groups of persons; and its form generates nonrandom factor structures when used over time. The only major reservation this reviewer has to recommending the STAI for both research and applied uses is its openness to faking. Further, it is recommended only if appropriate cautions are observed in administration and particularly interpretation. Users need not be psychometricians but should have some psychometric training, for without such training the very simplicity of the test will invite untrained users to attribute greater validity and reliability to the test than it has.

Edward S. Katkin, Professor of Psychology, State University of New York at Buffalo, Buffalo, New York.

The STAI was developed to provide reliable quantitative measurement of two distinct components of the anxiety concept: trait anxiety and state anxiety. The concept of trait anxiety, as defined by the authors, is that of a relatively enduring personality characteristic—anxiety proneness. A person who is defined as being high in trait anxiety is not a person who necessarily shows evidence of high anxiety at any given point in time, but is rather a person who is prone to respond to certain specified conditions with anxiety responses. State anxiety refers to anxiety which a person experiences in response to certain specific conditions. The anxiety state

State-Trait Anxiety Inventory
consists of phenomenological as well physiological responses to the environment and may not be correlated with the trait of anxiety. That is, both high and low trait-anxious individuals may from time to time experience the state of anxiety. It would be predicted theoretically that high trait-anxious individuals will experience state anxiety in response to a wider array of stimulus conditions than subjects who are low in trait anxiety.

The test manual is a model of what test manuals should be. It outlines concisely and intelligently the theoretical background and definition of the concepts to be measured and then provides explicit instructions on techniques for administration and scoring. In addition, norms for both the state and trait scales are provided for college freshmen, university undergraduates, high school students, neuropsychiatric patients, general medical and surgical patients, and inmates in a state prison. The normative samples for each of these categories are reasonably large and the data are presented in such a way that the test user can compare the scores of his population against normalized and percentile ranked scores for relative norm groups.

Test-retest reliabilities for the A-trait scale for male and female college undergraduates over a six-month period are .73 and .77, respectively, indicating that the trait measure is quite stable. Test-retest reliabilities for the A-state measure are low, as might be expected since the state measure conceptually does not measure a persistent characteristic of the individual. However, internal consistency of the A-state scale, as measured by K-R 20, ranges from .83 to .92. The A-trait scale correlates very highly with the Taylor and IPAT anxiety scales, indicating that the A-trait scale measures essentially the same concept and may be interpreted in the same context. The validity of the A-state scale has been demonstrated in a wide variety of studies, many of which are summarized in the test manual. In general, the technique is to give the A-state scale to subjects under two different instructional sets, one being a normal set and the other being a set to imagine oneself in a stressful situation. Scores on the A-state scale increase dramatically under the stress-imagina- tion condition as compared to the normal condition, indicating that the A-state scale measures changes in subjects' phenomenological experiences of anxiety.

A series of experiments has also been done on subjects in different states of mental stress and, in general, the A-state scale has turned out to be a reliable measure of increases in the state of anxiety resulting from the experimental manipulations.

Research with the State-Trait Anxiety Inventory has been proliferating to the point where there is probably more published research on the STAI, and more ongoing research now on the STAI, than on any other commercially available anxiety inventory. The bulk of the research tends to indicate that the distinction between state and trait anxiety is a useful one for researchers as well as clinicians, and that the STAI scale represents a relatively efficient, reliable, and valid way to assess individual differences in both anxiety-proneness and phenomenological experience of anxiety in normal as well as in patient populations.

In summary, it appears that the STAI is an excellent choice for the clinical psychologist or personality researcher looking for an easy-to-administer, easy-to-score, reliable, and valid index of either individual differences in proneness to anxiety or individual differences in transitory experience of anxiety. The test is carefully described in the manual, there is a voluminous research literature attesting to its reliability and validity in a variety of contexts, and the test is grounded well in psychological theory. For those researchers or clinicians who are interested only in the measurement of trait anxiety, it should be noted that well-known inventories such as the Taylor Manifest Anxiety Scale or the Welsh Anxiety Scale probably offer measurement of the same construct as is measured by the trait scale. The virtue of the STAI lies in the convenience with which one can measure both trait and state concepts of anxiety.

State-Trait Anxiety Inventory for Children. Grades 4-6; 1970-73; STAIC; downward extension of State-Trait Anxiety Inventory; title on test is How-I-Feel Questionnaire; 2 scores: state anxiety, trait anxiety; 1 form (70, 2 pages); preliminary manual (73, 11 pages); $2.50 per 25 tests; 50¢ per key; $3.50 per manual; $4 per specimen set; postage extra; (20) minutes; Charles D. Spielberger in collaboration with C. Drew Edwards, Robert E. Lushene, Joseph Montouri, and Denna Plateck; Consulting Psychologists Press, Inc.*

See T2:1392 (2 references).
MMPI, MMPI-2, & MMPI-A in Court: A Practical Guide for Expert Witnesses and Attorneys

Kenneth S. Popa, Ph.D., James N. Buscher, Ph.D., & Joyce Selahn, R.A., J.D.

This book should be read by any professional who deals with the MMPI-2/A in court. It represents the authoritative forensic source of information regarding research, references, and data on the MMPI-2/A.

This is a practical guide that can be used by both psychologists and lawyers. It pulls together psychometric, legal, scientific, and professional information into one source that can be used by those who testify using psychological assessment as well as those who challenge inadequate assessments and misleading testimony.

It includes numerous appendices containing sample agreements between expert witnesses and lawyers and presents cases involving the MMPI.

K-MMILR MMPI-2 in Court (hardcover, 1993, 450 pages) .................. $68.95

MMPI-2: Assessing Personality and Psychopathology
Second Edition
John R. Graham, Ph.D.

This volume incorporates what has been learned about the original and revised MMPI-2 in the four years since the introduction of the MMPI-2. In this version you are presented with an interpretive strategy, case examples, and pointers on how to present MMPI-2 results to clients. Information is given about both new alcohol abuse scales and a marital distress scale. There is a section on using the MMPI-2 with special populations such as older adults, ethnic minorities, medical patients, correctional groups, and individuals in non-clinical settings.

Computerized use of the MMPI-2 is considered, and a computerized interpretation is compared with a clinician-generated interpretation. The MMPI-A is discussed, along with suggestions for interpretation.

K-M22R MMPI-2 Assessing Personality (hardcover, 1993, 394 pages) .................. $46.95

Manual for Using the MMPI-2 as a Therapeutic Intervention
Stephen E. Pines, Ph.D.

This manual provides a step-by-step procedure, tested in controlled research and illustrated with case examples, for using the MMPI-2 as a therapeutic intervention. All stages of the MMPI-2 assessment are covered, from the initial interview, to the scoring and interpretation of the test and into the feedback sessions.


Bender Report 4.0
Giles Rainwater, Ph.D.
CLINICIAN COMPLETED

The BENDER REPORT facilitates scoring and provides a detailed interpretation of the Bender-Gestalt test. It can be used for both child and adult protocols and includes relevant normative comparisons, associated behavioral/personality characteristics, level of severity, diagnostic indications, and treatment recommendations. With the adult system, based on the work of Max Hutt, you can produce a report that includes a narrative interpretation, and diagnostic screening, and clinical findings sections. The child section interprets are based on the work of Elizabeth Koppitzt and produces a narrative report that covers developmental level, time to complete the test, neurological indicators, emotional indicators, and behavioral indicators.

K-BR1-38 Bender Report MS DOS 36" ................................................... $200.00

State-Trait Anxiety Inventory (STAI)
Charles D. Spielberger
The STAI measures anxiety in individuals aged 13 and older

The STAI differentiates between the temporary condition of "State Anxiety" and the more general and long-lasting quality of "Trait Anxiety." The STAI helps professionals distinguish between a client's feelings of anxiety and depression. The STAI has 40 questions with a range of 4 possible responses to each. It provides norms for clinical patients, high school and college students, and working adults. One can complete the inventory in 10 minutes. A child version for children younger than 13 is available (see page 9).

K-ST1P State-Trait KH (includes Manual, 25 Question Booklets, and Scoring Key) .................................................. $70.00
K-ST2P State-Trait Question Booklets (25/pk) .................. $35.00
K-ST3P State-Trait Scoring Key .................................................. $21.00
K-ST4P STAI Bibliography .................................................. $45.00

STAI Computer Program
The STATE-TRAIT ANXIETY INVENTORY COMPUTER PROGRAM is an administration and scoring program. STAI's two 20-item self-report scales assess anxiety-proneness (trait) and report the current level of anxiety (state). The computer administers the test and can immediately score and graph the results. Data from the program can be stored for future use.

The STAI Computer Program is sold with 50 administrations.

K-ST1-38 STAI MS DOS 36" (50 uses) .................................................. $100.00

908 Niagara Falls Blvd., North Tonawanda, NY 14120-2060
State-Trait Anxiety Inventory for Children

This instrument is an off-shoot of the State-Trait Anxiety Inventory. It is designed to measure the same constructs as the above mentioned instrument, but it's target population is children in grades four through six. It possesses many of the same characteristics of the adult version in terms of length and time to administer, but some adaptations have been made in response styles to make it easier for children to complete.

This test is newer than the adult version, therefore is not as widely accepted. There are still some limitations in terms of reliability issues, but The Eighth Mental Measurements Yearbook (Kramer & Conoley, 1992) states that, “despite the limitations mentioned above, this scale is probably the best scale available for assessing anxiety in children” (pg. 684).

For more information, see attached: see Kramer, J.J., & Conoley, J.C. (Eds.). (1992). The eleventh mental measurements yearbook, Lincoln, Nebraska: The University of Nebraska Press.

Provides a brief description of the test, references for additional information, and a write up evaluating the strengths and weaknesses of this measure.


This presents a study conducted to provide validity information on two different measures of stress and anxiety in children, one of which is the State-Trait Anxiety Inventory for Children. It provides validity information for both, and concludes that both instruments are potentially useful for screening.


This paper presents additional support for the use of the State-Trait Anxiety Inventory for Children. The authors conducted a study to measure the effects of stress on trait and state anxiety in children and found the hypothesized results to be confirmed.
consists of phenomenological as well physiological responses to the environment and may not be correlated with the trait of anxiety. That is, both high and low trait-anxious individuals may from time to time experience the state of anxiety. It would be predicted theoretically that high trait-anxious individuals will experience state anxiety in response to a wider array of stimulus conditions than subjects who are low in trait anxiety.

The test manual is a model of what test manuals should be. It outlines concisely and intelligently the theoretical background and definition of the concepts to be measured and then provides explicit instructions on techniques for administration and scoring. In addition, norms for both the state and trait scales are provided for college freshmen, university undergraduates, high school students, neuropsychiatric patients, general medical and surgical patients, and inmates in a state prison. The normative samples for each of these categories are reasonably large and the data are presented in such a way that the test user can compare the scores of his population against normalized and percentile ranked scores for relative norm groups.

Test-retest reliabilities for the A-trait scale for males and female college undergraduates over a six-month period are .73 and .77, respectively, indicating that the trait measure is quite stable. Test-retest reliabilities for the A-state measure are low, as might be expected since the state measure conceptually does not measure a persistent characteristic of the individual. However, internal consistency of the A-state scale, as measured by K-R 20, ranges from .83 to .92. The A-trait scale correlates very highly with the Taylor and IPAT anxiety scales, indicating that the A-trait scale measures essentially the same concept and may be interpreted in the same context. The validity of the A-state scale has been demonstrated in a wide variety of studies, many of which are summarized in the test manual. In general, the technique is to give the A-state scale to subjects under two different instructional sets, one being a normal set and the other being a set to imagine oneself in a stressful situation. Scores on the A-state scale increase dramatically under the stress-imagina-
tion condition as compared to the normal condition, indicating that the A-state scale measures changes in subjects’ phenomenological experiences of anxiety.

A series of experiments has also been done on subjects in different states of mental stress and, in general, the A-state scale has turned out to be a reliable measure of increases in the state of anxiety resulting from the experimental manipulations.

Research with the State-Trait Anxiety Inventory has been proliferating to the point where there is probably more published research on the STAI, and more ongoing research now on the STAI, than on any other commercially available anxiety inventory. The bulk of the research tends to indicate that the distinction between state and trait anxiety is a useful one for researchers as well as clinicians, and that the STAI scale represents a relatively efficient, reliable, and valid way to assess individual differences in both anxiety-proneness and phenomenological experience of anxiety in normal as well as in patient populations.

In summary, it appears that the STAI is an excellent choice for the clinical psychologist or personality researcher looking for an easy-to-administer, easy-to-score, reliable, and valid index of either individual differences in proneness to anxiety or individual differences in transitory experience of anxiety. The test is carefully described in the manual, there is a voluminous research literature attesting to its reliability and validity in a variety of contexts, and the test is grounded well in psychological theory. For those researchers or clinicians who are interested only in the measurement of trait anxiety, it should be noted that well-known inventories such as the Taylor Manifest Anxiety Scale or the Welsh Anxiety Scale probably offer measurement of the same construct as is measured by the trait scale. The virtue of the STAI lies in the convenience with which one can measure both trait and state concepts of anxiety.

State-Trait Anxiety Inventory

State-Trait Anxiety Inventory for Children. Grades 4-6: 1970-73: STAI-C; downward extension of State-Trait Anxiety Inventory; title on test is How-I-Feel Questionnaire; 2 scores: state anxiety, trait anxiety; 1 form (40, 2 pages); preliminary manual (72, 11 pages); $2.50 per test; 50c per key; $3.50 per manual; $4 per specimen set; postage extra; (20) minutes; Charles D. Spielberger, Robert E. Lushene, Joseph Montuori, and Denna Plattek; Consulting Psychologists Press, Inc. * See T2:1392 (2 references).
The STAI is a 40-item self-administered scale and has no time limits. Twenty items are designed to measure A-state (how you feel right now) and the other 20 items to assess A-trait (how you generally feel). Whereas the adult form has a four-point scale for each item, the STAI has a three-point scale, which includes hardly-ever, sometimes, or often. Examples of some of the trait items are “I feel like crying,” “I am sly,” “My hands get sweaty.” Examples of state items include “I feel very nice, nice or not nice,” “I feel very upset, upset, not upset.”

The STAI was initially developed in 1969. Elementary school children (a total of 1554) in grades 4-6 were used as the normative sample. Both black and white children were represented. The selection of the final set of items from a larger pool of items was based on the joint criteria of internal consistency and concurrent validity of each item. The manual provides separate norms (means and standard deviations) for each grade and for both A-state and A-trait. Percentile ranks and normalized standard scores are provided.

The alpha reliability internal consistency coefficients for A-state are .82 for males and .87 for females; the coefficients for A-trait are .78 and .81, respectively. The test-retest reliability coefficients for A-trait are .65 for males and .75 for females; coefficients for A-state are .31 and .47, respectively. The authors try to explain away the low A-state test-retest coefficients as an index of the transitory nature of anxiety states. However, it is our contention that if subjects are tested under two equivalent conditions, e.g., two neutral or nonstressful conditions or two conditions of equivalent stress, then the coefficients should be higher. Otherwise, the

The STAI-C, a downward extension of the State-Trait Anxiety Inventory, was initially designed as a research tool for investigating anxiety in elementary school children. It includes two separate self-report scales for measuring state anxiety (A-state) and trait anxiety (A-trait). The authors claim that the STAI-C A-state scale may be used as an index of drive level (D) or as an index of the actual levels of A-state invoked by stressful experimental procedures. Furthermore, the A-state measures transitory anxiety experienced by children in psychotherapy and counseling situations and can also be used to assess the effectiveness of behavior therapy procedures. The STAI-C A-trait scale was designed to measure children who differ in anxiety proneness (research purpose) or as a screening technique for detecting neurotic behavioral tendencies in elementary school children.
test-retest reliability of the STAIC A-state is questionable. In other words, the internal consistency of the STAIC scales is fairly good, but the test-retest reliability is poor. In either case, both the internal consistency and stability coefficients of STAIC are not as good as the equivalent coefficients for the adult STAI. In general, the adult STAI is further developed than the STAIC.

The research on the validity of the STAIC is not as far advanced as for STAI. However, the adult form is an older scale and, hopefully, Spielberger and his colleagues will, in the future, provide additional evidence for the validity of the STAIC. The adult form is one of the best researched self-report instruments published and one would expect that in the future the STAI will achieve the same standards. Some of the preliminary evidence for the validity of the STAIC indicates that, for each item, the STAIC A-state score is higher under test conditions (feelings before a final examination) than it is under norm (standard instructions) conditions. This provides some evidence for the construct validity of the STAIC A-state. Evidence for the concurrent validity of the STAIC A-trait is based on correlations between the STAIC A-trait scale and both the CMAS and GASC. For a sample of 75 children the STAIC A-trait correlated .75 with the CMAS and .63 with the GASC. Correlations with nonanxiety scales indicate that the STAIC A-trait is not a good predictor of either aptitude or achievement. There are a number of recently completed experimental studies that show more promise.

A number of criticisms raised about STAIC apply equally as well to STAIC. Endler [683 (224)] and Endler and Magnusson [683 (286)] have provided evidence to indicate that the adult scales, both A-state and A-trait, are multidimensional rather than unidimensional. Yet the scales are treated as unidimensional. For example, Endler and Magnusson present evidence that there are at least three response factors for A-trait and at least three for A-state (e.g., affective discontent, cognitive preoccupation, and avoidance apathy). These factors may operate differentially and, if there were separate factor scores, the predictive validity of the STAI scales might be improved. Furthermore, Endler and Magnusson have indicated that although A-state and A-trait may be conceptually distinct, empirically they appear to be related, at least for Canadian and Swedish college students. It would have been useful if the authors had reported correlations between the A-state and A-trait subscales of the STAIC for the normative sample. What is the long term evidence for the stability of A-trait with children? Are there any sex differences in A-state or A-trait as assessed by the STAIC?

Despite the limitations mentioned above, this scale is probably the best scale available for assessing anxiety in children. I would recommend it over the CMAS and GASC, primarily on the basis of the care and precision with which it has been developed. I would, however, recommend it primarily as a research instrument; it is probably premature to use it in applied clinical work. Hopefully, on the basis of research with this instrument, it may be possible at some future date to use it in a clinical setting. However, for experimental research in anxiety and for research in clinical settings, the STAIC is a useful scale. It has a good theoretical basis, adequate norms, adequate reliability, and moderate validity.

Structured and Scaled Interview to Assess Maladjustment. Mental patients; 1974; SSIAM; social adjustment; 11 ratings (5 deviant behavior, friction with others, 3 distress, 2 inferential) in each of 5 areas (work, social-leisure, family, marriage, sex) plus 11 overall ratings; no data on reliability and validity in manual; no norms; individual; 1 form (27 pages plus 5-page training manual); $24 per 10 tests; $7.50 per specimen set of 3 tests; postage extra; (30) minutes; Barry J. Guriland, Neil J. Yorkston, Anthony R. Stone, and Jerome D. Frank; Springer Publishing Co., Inc.

REFERENCES

Validity of Two Measures of Anxiety in Children

L.E. Montgomery
Middle Peninsula—Northern Neck Mental Health and Mental Retardation Services

A. J. Finch, Jr.¹
Virginia Treatment Center for Children

The State-Trait Anxiety Inventory for Children (STAIC) and the Children's Manifest Anxiety Scale (CMAS) were administered to 60 emotionally disturbed children and 60 normal children matched on mental age. Emotionally disturbed children obtained significantly higher scores on the A-State and A-Trait portions of the STAIC and the anxiety portion of the CMAS, while normal children obtained significantly higher scores on the Lie scale of the CMAS. Cutoff scores which maximized the discrimination between normal and emotionally disturbed children were established. Results are interpreted as demonstrating the potential usefulness of these measures for screening.

The State-Trait Anxiety Inventory for Children (STAIC) was recently developed as a research tool for the study of anxiety in children (Spielberger, 1973). The test consists of two independent 20-item self-report scales designed to measure state anxiety (A-State) and trait anxiety (A-Trait). The A-State scale was designed to measure transitory anxiety states that vary over time, while the A-Trait scale was designed to measure relatively stable individual differences in anxiety proneness. Good reliability has been demonstrated with both normal (Spielberger, 1973) and emotionally disturbed children (Finch, Montgomery, & Deardorff, 1974b).

The Children's Manifest Anxiety Scale (CMAS) was developed by Castaneda, McCandless, and Palermo (1956) as a measure of anxiety in children. This scale consists of 42 anxiety items and 11 additional items which constitute a Lie scale. The reliability of the CMAS has been demonstrated also with normal children (Castaneda et al., 1956; Holloway, 1958; Kitano, 1960; Levitt, 1957).

Manuscript received in final form September 30, 1974.
¹Requests for reprints should be sent to Dr. A. J. Finch, Jr., Virginia Treatment Center for Children, P. O. Box 1-1, Richmond, Virginia 23201.

293
and with emotionally disturbed children (Finch, Montgomery, & Deardoff, 1974a).

Validity data have been summarized by Spielberger (1973) for the STAIC. Platzek (1970) demonstrated the concurrent validity of the STAIC with the CMAS and the General Anxiety Scale for Children. In another study, negative correlations with measures of Verbal IQ, aptitude, and achievement and positive correlations with school grades have been reported (Spielberger, 1973). However, in this same study neither A-State nor A-Trait anxiety was correlated with teacher ratings of behavioral manifestations of anxiety.

Since the validity of the A-State scale depends on its ability to reflect momentary changes in anxiety, validity studies have focused on changes in scores following a threat or other anxiety-producing situation. Spielberger (1973) reported a significant increase in STAIC A-State scores when children were asked to respond according to how they would feel immediately before taking a school examination. Likewise, Gaudry and Poole (1972) reported increases in A-State scores following a failure condition and decreases following a success condition.

Aside from its concurrent validity with the STAIC A-Trait, validity data for the CMAS come from a study by Kitano (1960). Behavior-problem children in an adjustment class were matched with normal children on grade level, socioeconomic status, and school district and were administered the CMAS. Results indicated no significant differences on the Lie scale, but significantly higher scores on the Anxiety scale for the behavior-problem children.

The purpose of the present study was to investigate the predictive validity of the STAIC and the CMAS by comparing scores of emotionally disturbed children in residential treatment and normal children. Specifically it was hypothesized that emotionally disturbed children would obtain significantly higher scores on these anxiety measures than would normal children. In addition, the present study was designed to determine the potential usefulness of these measures of anxiety as diagnostic screening instruments in differentiating emotionally disturbed from normal children.

METHOD

Subjects

A total of 60 emotionally disturbed children (49 males, 11 females) admitted for residential treatment to the Virginia Treatment Center for Children, a short-term residential treatment facility, were matched with 60 normal children on sex and mental age (MA) as measured by the Peabody Picture Vocabulary Test, Form A. Of the cases of emotional disturbance, 25% were diagnosed "overanxious reaction of childhood," 17% were "unsocialized aggressive reaction of childhood," 13% were "withdrawing reaction of childhood," 12% were "adjustment reaction of childhood"; the remaining were distributed over 11 different diagnostic categories. The mean chronological age (CA) of the emotionally disturbed group was 11.28 years (SD = 2.10 years), while the mean MA of the normal group was 10.41 years (SD = 2.73 years). The mean mental age (MA) of the emotionally disturbed group was 11.33 years (SD = 3.49), while the mean MA of the normal group was 11.31 years (SD = 3.48).

Procedure

The subjects were pretested on the Peabody Picture Vocabulary Test and 2 weeks later, individually administered the STAIC and then the CMAS according to standard instructions in a single session.

RESULTS

Table 1 presents the means, standard deviations, and $t$ values for emotionally disturbed and normal children on the STAIC A-State and A-Trait and the CMAS. As predicted, emotionally disturbed children obtained significantly higher scores on the A-State and A-Trait portions of the STAIC and the anxiety portion of the CMAS. In addition, the normal group obtained a significantly higher score on the Lie portion of the CMAS.

Point-biserial correlations calculated between diagnosis (emotionally disturbed or normal) and scores on the Anxiety scales were all significant, as can be seen in Table II. These correlations further attest to the ability of these measures to differentiate between the two groups.

In order to determine the ability of scores obtained on these scales to correctly identify individuals, optimal cutoff scores were established for the

| Table 1. Means, Standard Deviations, and $t$ Values for Emotionally Disturbed and Normal Children on the Anxiety Measures |
| --- | --- | --- |
| STAIC | CMAS |
| A-State | A-Trait | Anxiety | Lie |
| Emotionally disturbed $\bar{X}$ | 34.98 (81st percentile) | 40.73 (68th percentile) | 21.65 | 3.02 |
| SD | 8.17 | 7.81 | 10.37 | 2.28 |
| Normal $\bar{X}$ | 29.78 (53rd percentile) | 37.05 (43rd percentile) | 15.78 | 4.25 |
| SD | 4.74 | 6.50 | 7.17 | 2.69 |
| $t$ value | 4.05$^a$ | 2.82$^b$ | 3.67$^a$ | 2.76$^c$ |

$^a$p < .001.

$^b$p < .005.

$^c$p < .01 (two-tailed).
Table II. Point-Biserial Correlations Between Diagnosis (Emotionally Disturbed or Normal) and Scores on Anxiety Scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>r pb with diagnosis (emotionally disturbed–normal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAIC</td>
<td></td>
</tr>
<tr>
<td>A-State</td>
<td>.36 a</td>
</tr>
<tr>
<td>A-Trait</td>
<td>.25 b</td>
</tr>
<tr>
<td>CMAS</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>.31 a</td>
</tr>
<tr>
<td>Lie</td>
<td>.24 b</td>
</tr>
</tbody>
</table>

a p < .001.

b p < .01.

Table III. Optimal Cutoff Scores for the Anxiety Measures and the Percentage Correctly and Incorrectly Classified

<table>
<thead>
<tr>
<th></th>
<th>A-State</th>
<th>A-Trait</th>
<th>STAIC Anxiety</th>
<th>CMAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal cutoff score</td>
<td>31</td>
<td>39</td>
<td>69</td>
<td>18</td>
</tr>
<tr>
<td>Percent emotionally disturbed correctly identified</td>
<td>65</td>
<td>63</td>
<td>65</td>
<td>62</td>
</tr>
<tr>
<td>Percent normals misclassified</td>
<td>29</td>
<td>37</td>
<td>35</td>
<td>33</td>
</tr>
</tbody>
</table>

A-State, A-Trait, and CMAS Anxiety scales. In addition, since the A-State and A-Trait are usually administered together, scores on these two scales were combined to form a STAIC Anxiety score. These optimal cutoff scores for the anxiety measures and the percentage of correctly and incorrectly classified children are presented in Table III. In addition, it should be noted that on the combined STAIC Anxiety score none of the normals obtained a score greater than 87 while 10 of the emotionally disturbed group did.

A-State scores correlated significantly with A-Trait scores (r = .51, p < .001), CMAS Anxiety scores (r = .47, p < .001), and Lie-scale scores (r = .23, p < .02). A-Trait scores correlated significantly with CMAS Anxiety scores (r = .52, p < .001), MA (r = -.21, p < .02), and CA (r = -.24, p < .01). Lie-scale scores correlated significantly with CMAS Anxiety scores (r = -.19, p < .05), MA (r = .47, p < .001), and CA (r = .47, p < .001).

DISCUSSION

The results of the present study lend additional support to the validity of the STAIC and the CMAS. On both the A-State and A-Trait portions of the STAIC and the anxiety portion of the CMAS, emotionally disturbed children obtained significantly higher scores than did normal children. From examination of individual scores, it appears that both tests hold promise as screening instruments for identifying emotionally disturbed children. For each instrument, the optimal cutoff score resulted in a moderate number of false positives but for screening purposes this number was acceptable.

In addition, Quay (1972) reports that anxiety is not a necessary covariant with emotional disturbance. In reviewing factor-analytic research, he notes that anxiety is not generally reported in children having conduct disorders, and since at least 20% of the present sample of emotionally disturbed children would be considered to have conduct disorders the present results are more encouraging. With the increasing pressure for early identification of children with emotional problems in public school, the use of these scales for screening purposes should prove highly beneficial.

One possible shortcoming of the present study is that it involved only hospitalized emotionally disturbed children. Additional data are needed on the utility of the CMAS and STAIC in differentiating levels of anxiety in other groups of emotionally disturbed children and normals. In the Kitano (1960) study, the absolute differences in the CMAS scores of the normal group and the adjustment class group were not great and individual scores were not examined. Unless individuals can be differentiated, these scales would not provide useful information for clinical purposes since the individual scores are of prime importance rather than differences between groups.

The correlation between the A-Trait and the CMAS was somewhat lower than that reported by Spielberger (1973). The reason for this finding is not known but does indicate the need for further research in the area.

REFERENCES


Sensitivity of the Faschingbauer Abbreviated MMPI to Hospitalized Adolescents

Charles S. Newmark and Lee Glenn
University of North Carolina School of Medicine

Comparisons were obtained between Faschingbauer's abbreviated MMPI (FAM) and the standard MMPI for a sample of hospitalized adolescent psychiatric patients. The results showed close statistical correspondence, highly correlated scale group means, and high comparability in identifying valid and invalid profiles. Additionally, the FAM was successful in predicting several gross indices of psychopathology as well as high-point elevations on the standard-form MMPI. A replication supported the stability of these findings.

The MMPI is one of the most widely used clinical instruments designed to provide an objective assessment of the major personality characteristics that affect personal and social adjustment. Its value is greatly enhanced by the large body of actuarial and statistical data that has become available as a result of research and clinical use. However, there are numerous clinical situations where its extreme length prohibits its use. Nowhere is this more evident than with adolescents on a psychiatric inpatient ward. The approximately 2-hour testing session tends to provoke and irritate restless adolescents, who then express their anger and defiance by seizing on the legitimate option of omitting items or failing to complete the test (Dahlstrom, Welsh, & Dahlstrom, 1972). This submarginal compliance with routine procedures tends to vitiate the results. Occasionally, staff insistence will lead to a completed MMPI but only after the adolescent has grudgingly completed a few items at a time over a period of several days. Often, important changes in the patient's mood or behavior are noted to occur during this test-taking period.

While there have been several recent attempts to develop an abbreviated MMPI, mainly Kincannon's (1968) 71-item Mini-Multi, Dean's (1972) 86-item

Manuscript received in final form October 1, 1974.

Requests for reprints should be sent to Dr. Charles S. Newmark, Department of Psychiatry, University of North Carolina Medical School, Chapel Hill, North Carolina 27514.
The Effects of Stress on State and Trait Anxiety in Emotionally Disturbed, Normal, and Delinquent Children

Jeffrey R. Bedell and John Roitsch
Medical University of South Carolina

The effects of stress on trait and state anxiety of emotionally disturbed, normal, and delinquent children were assessed using the State–Trait Anxiety Inventory for Children (STAIC). This scale measures both transitory feelings of apprehension and tension that vary over time as a function of situational stress (A-State) and individual differences in the disposition to experience such anxiety states (A-Trait). Results indicated that A-State increased as a function of stress, whereas A-Trait was relatively stable and impervious to the temporary effects of stress. Use of the STAIC with normal and deviant populations was supported.

Psychological research has provided considerable support for Spielberger's (1966, 1972) State–Trait Anxiety Theory (see Smith & Lay, 1974). In order to facilitate the application of this theory to research with children, the State–Trait Anxiety Scale for Children (STAIC) was constructed. Reliability and validity data for the STAIC support its use with normal children for whom it has been demonstrated that STAIC A-Trait scores are relatively impervious to stress and remain stable over time, and that STAIC A-State increases as a function of stress (Finch, Kendall, & Montgomery, in press; Spielberger, 1973).

The usefulness of the STAIC for assessing state and trait anxiety in emotionally disturbed children has been questioned by the findings in several recent studies. Finch, Montgomery, and Deardorff (1974) reported that the test–retest reliability of the A-Trait scale was low, and Sitarz (1974) and Finch et al. (in press) found that the STAIC A-Trait scale showed significant increases following situational stress for emotionally disturbed children.

Manuscript received in final form February 5, 1976

Requests for reprints should be sent to Jeffrey R. Bedell, Department of Psychology, University of Virginia, Charlottesville, Virginia 22901.
The explanation most frequently suggested for these findings is that emotionally disturbed children experience state and trait anxiety differently from normal children (Finch et al., in press; Finch, et al., 1974; Sitarz, 1974). However, a comparison of studies supporting State—Trait Anxiety Theory and those reporting unstable A-Trait scores reveals important methodological differences with regard to the type of experiences intervening between STAIC administrations. In studies in which emotionally disturbed children were tested, most of the subjects were involved in inpatient psychotherapeutic treatment for periods of 10 days to 3 months between STAIC assessments. Studies with normal children have had shorter periods of time between assessments, with no intervening experiences such as psychotherapy. Thus, the effects of psychotherapy on the disturbed children and not differences between subject populations may be causing the discrepant findings.

The present study tested the hypothesis that state anxiety in hospitalized, emotionally disturbed children will increase as a function of stress, and that trait anxiety will remain stable when testing occurs over a short period of time so that the effects of psychotherapy are eliminated. In order to compare state and trait anxiety in normal and disturbed samples, the present study included an emotionally disturbed group that was expected to be high in anxiety proneness, a group of normal children, and a group of delinquent children who were expected to be low in anxiety proneness.

**METHOD**

**Subjects**

The subjects were 15 emotionally disturbed children, 15 normal schoolchildren, and 15 juvenile delinquents. The emotionally disturbed children had admitting diagnoses of adjustment to childhood or adolescence and were undergoing inpatient psychiatric treatment. The normals were seventh-grade public school children with no current discernable adjustment problems, and the juvenile delinquents were undergoing outpatient group psychotherapy by order of the juvenile court.

The average ages of the emotionally disturbed, normal, and delinquent groups were 14.62 years (range 13–15), 14.13 years (range 13–15), and 15.05 years (range 13–16), respectively.

**Procedure**

After nonthreatening introductory comments in which the children were requested to complete several tests and questionnaires, groups of subjects were administered the STAIC A-State and A-Trait scales, with standard instructions.

The Arithmetic subtest of the Wide Range Achievement Test (WRAT) was then administered with stressful, ego-involving instructions. The subjects were told that the WRAT would measure their ability in arithmetic and mathematics. The emotionally disturbed and normal children were tested in a school situation and then advised that their WRAT performance would be used by their teachers as part of their school evaluation. The delinquent sample was tested during their weekly group therapy meeting and told that their performance on the WRAT would be considered as part of their posttherapy evaluation that would be reported to their judge.

Upon completion of the WRAT, the subjects were told that their tests were being corrected and that they would shortly receive feedback on their arithmetic ability. While the WRAT was being scored, the STAIC was administered again with standard instructions. Subsequent to the experiment, subjects were told the purpose of the study, and that participation in the experiment would not result in negative reports being sent to schools or courts.

**RESULTS**

The mean pre- and poststress A-State scores for the three diagnostic groups are reported in Table 1. These data were evaluated in a 3 X 2 analysis of variance in which diagnostic group was the between-subjects variable and pre-post testing was the within-subjects variable.

The only significant effect in this analysis was a pre-post stress main effect: $F (1, 42) = 6.90, p < .05$. This finding indicated that state anxiety increased as a function of stress and the three diagnostic groups were not significantly different with respect to changes in A-State.

A-Trait scores were evaluated in the same manner as were the A-State scores, but there were no significant $F$ ratios in this analysis. Thus, there were no differences in the A-Trait scores of the emotionally disturbed, normal, and juvenile delinquent groups, and no changes in A-Trait as a function of stress.

Combining the three experimental groups, the correlations between pre- and posttreatment A-State and A-Trait scores were .56 and .94, respectively. These correlations are consistent with State—Trait Anxiety Theory, and indicate

| Table 1. Mean Pre- and Poststress STAIC A-State and A-Trait Scores for Emotionally Disturbed, Normal, and Juvenile Delinquent Children |
|---|---|---|---|---|
| Group | $N$ | STAI A-State | STAI A-Trait |
| | | Pre | Post | Pre | Post |
| Emotionally disturbed | 15 | 32.33 | 36.47 | 38.67 | 38.20 |
| Normal | 15 | 32.00 | 35.80 | 39.27 | 39.00 |
| Juvenile delinquent | 15 | 35.40 | 34.60 | 37.80 | 37.13 |
that changes in state anxiety are associated with stress, whereas trait anxiety is relatively impervious to stress.

**DISCUSSION**

In the present study, ego-threatening stress produced significant increases in children’s state anxiety, but not their trait anxiety. These findings are consistent with the research involving normal children reported by Spielberger (1973), but not with Finch et al. (in press); Finch et al. (1974); and Sitarz (1974), who observed changes in both state and trait anxiety in repeated administrations of the STAIC to emotionally disturbed children.

While Finch and his colleagues interpreted their results as suggesting that Spielberger’s (1966, 1972) State–Trait Anxiety Theory needed to be revised, it may be that methodological differences more parsimoniously explain previously reported inconsistencies in anxiety research on normal and emotionally disturbed children.

Finch and his co-workers have examined hospitalized emotionally disturbed children over relatively long periods of time during which they were undergoing psychotherapeutic treatment. Such circumstances are quite different from the conditions of the present research and the research reported by Spielberger (1973), whose children were tested over short periods of time during which psychotherapy and other extraneous variables had less opportunity to influence the subjects.

Research using the STAIC with adults has shown that changes in A-Trait may occur as a function of intense experiences such as electroconvulsive therapy (Newmark, 1972), living in crowded confinement (Smith & Haythorn, 1972), or experiencing recent changes in A-State (Gorsuch, 1969). Similarly, it would be expected that children exposed to significant changes in environment, such as intensive psychotherapy for long periods of time, would report changes in A-Trait as a function of these experiences.

**REFERENCES**


Feelings, Attitudes, and Behaviors Scale for Children (FAB-C)

Joseph H. Belchman, M.D.

Even children with limited reading abilities can easily complete the items of the FAB-C.

QuickView

Description: Scale for emotional and behavioral assessment of young children
Ages: 6 to 13
Administration: Self-report
Administration Time: 10 minutes
Qualification Level: b (see page 55)

The FAB-C is a self-report scale designed to assess emotional and behavioral problems in children aged 6 to 13 years. The FAB-C comprises 48 yes-no items that can be completed in individual or group administrations in under 15 minutes, even by children who have limited reading abilities. You can score the test in under 10 minutes using the QuickScore™ Form and then transfer the results to the Profile Form, which helps you convert raw scores to T-scores.

The FAB-C produces five factors that assess the following areas:
- Self-Image
- Negative Peer Relations
- Conduct Problems
- Antisocial Behavior
- Worry

The FAB-C also includes an overall Problem Index, which is useful in identifying children who might benefit from more detailed assessment, and a Life Scale, which gauges last-taking attitudes.

The FAB-C was developed using a normative sample of 1,988 children (1,074 males and 914 females) from rural and metropolitan areas. Separate norms are available for boys and girls in two-year intervals. The 89-page FAB-C Manual describes how to administer, score, and interpret the test. It also gives 6 case studies, describes the rational and empirical stages in the development, and presents extensive normative, reliability, and validity data.

The FAB-C was designed for clinical and research purposes. It can be used as a routine screening device in a number of settings, such as schools, outpatient clinics, residential treatment centers, child protective services, and private practice offices. The FAB-C is a useful tool for providing structured and normed information about a number of potential problem areas in children.

K-FAB1P FAB-C Complete Kit (Includes Manual and 25 QuickScore™ Forms) ............................................. 155.00
K-FAB2P FAB-C Manual ............................................. 90.00
K-FAB3P FAB-C QuickScore™ Forms (25/pkg) ........... 27.00
K-FAB1S Specimen Set (Includes Manual, 3 QuickScore™ Forms) ..................................................... 132.05

State-Trait Anxiety Inventory for Children (STAIC)

Charles Spielberger, Ph.D.

COMPLETED BY PATIENT

The "How I Feel" Questionnaire

The STAIC-TRAIT ANXIETY INVENTORY (STAIC) has been designed for elementary school children ages 9 to 12 years. It contains two 20-item scales measuring "trait" and "stat" anxiety. The STAIC Computer Program administers and scores the test in 8 to 20 minutes providing grade appropriate scores. The data can be collected and stored for future use. The program is sold with 50 administrations.

K-CS1P STAIC Kit (Includes Manual, 25 Questionnaires, and Scoring Key) ............................................ 170.00
K-CS2P STAIC Questionnaire (25/pkg) .......................... 35.00
K-CS1 STAIC MS DOS 3.0 (50 uses) ................................ 90.00

The Journal of Attention Disorders, published by Multi-Health Systems Inc. and edited by C. Keith Conners, Ph.D., is now in its second year of publication. This quarterly journal focuses exclusively on basic and applied research and clinical issues related to attention in children, adolescents, and adults. Included are articles on diagnosis, comorbidity, neuropsychology, functioning, psychopharmacology, classroom management strategies, parent training, behavioral assessment, diet, family therapy, and other areas relevant to problems in attention. Submissions are currently being accepted. For submission guidelines or to subscribe, please contact MHS Inc. 1(800) 456-3003.

Individual $45/yr; Institutional $95/yr. (Add $11 outside the U.S.) All subscriptions payable in U.S. funds. Rates are valid until April 1, 1996.
STATE-TRAIT ANXIETY INVENTORY FOR CHILDREN

Date: 05/13/96

Child's Name: JOHN
Sex: Male
Age: 10
Grade: 4

FORM C-1 (STATE) FORM C-2 (TRAIT)
1. 2 11. 2 1. 2 11. 3
2. 3 12. 1 2. 1 12. 2
3. 2 13. 2 3. 2 13. 3
4. 3 14. 1 4. 2 14. 1
5. 2 15. 1 5. 2 15. 3
6. 1 16. 3 6. 1 16. 3
7. 2 17. 2 7. 2 17. 2
8. 1 18. 2 8. 3 18. 2
9. 2 19. 2 9. 3 19. 3
10. 3 20. 2 10. 2 20. 2

RAW SCORE: 35 RAW SCORE: 44
PERC RANK STATE: 82 PERC RANK TRAIT: 88
NORM 'T' STATE: 59 NORM 'T' TRAIT: 62

PERCENTILE: This child's score is at or above the 82nd percentile of STATE anxiety and 88th percentile of TRAIT anxiety for children.

NORM: This child's score falls within the average range of STATE anxiety and above the average range of TRAIT anxiety.

******************************************************************************

State-Trait Anxiety Inventory for Children Graph (T-Scores)

0 10 20 30 40 50 60 70 80 90 100
!----------------------------------------! T-Score

STATE Anxiety 59

TRAIT Anxiety 62

******************************************************************************


Self-esteem

The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children

This instrument is better known as the Harter Scale. It is designed to measure childrens' perceptions of their own self-competence. The target population for this instrument is children in preschool through the second grade. One limitation of this instruments is that it needs to be administered individually and will need the availability of a staff person to conduct this as a one-on-one interview with the child. However, it is easy to score and administer. Another limitation is that the score derived is based on how the child perceives him or herself to be like or unlike the depicted character. It does not address how the child feels about that likeness or unlikeness. This may affect results because whether the child feels like or unlike the depicted character is only important in relation to how desirable or undesirable that trait is to the child. The test assumes which traits are desirable to children. Despite this, it is still widely used among researchers.

This instrument is useful in determining how competent a child feels in four different areas: 1) cognition (or intellect), 2) physical, 3) peer acceptance, and 4) maternal acceptance. It was selected based on its ease of use, low cost, and popularity amongst researchers.


This is the test authors' write up of the description, potential usefulness, validity, and reliability of the Pictorial Scale of Perceived Competence and Social Acceptance.

This is a paper describing a study conducted to measure the relationship between academic readiness and children's perceptions of competence and acceptance. The authors used the Pictorial Scale of Perceived Competence and Social Acceptance to measure this component and a standardized test to measure academic readiness.

(Another version of this test is also available for the school-aged population. It is called The Perceived Self-Competence Scale, and uses words instead of pictures. See the attached article for more information on this instrument:

The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children

Susan Harter and Robin Pike
University of Denver

HARBER, SUSAN, and PIGE, ROBIN. The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children. CHILD DEVELOPMENT, 1984, 55, 1969–1982. A new pictorial scale of perceived competence and social acceptance for young children, a downward extension of the Perceived Competence Scale for Children, is described. There are 2 versions of this instrument, 1 for preschoolers and kindergartners and a second for first and second graders, each tapping 4 domains: cognitive competence, physical competence, peer acceptance, and maternal acceptance. Factor analyses reveal a 2-factor solution. The first factor, general competence, is defined by the cognitive and physical competence subscales. The second factor, social acceptance, comprises the peer and maternal acceptance subscales. The psychometric properties were found to be acceptable. Weak correlations between children's and teachers' judgments are discussed in terms of the young child's tendency to confuse the wish to be competent or accepted with reality. It is urged that this instrument not be viewed as a general self-concept scale but be treated as a measure of 2 separate constructs, perceived competence and social acceptance.

Introduction

Constructs such as self-concept and self-esteem have had a long history within the field of psychology, although in recent years there has been a revival of interest in topics involving the self and self-description (see Harter, 1983; Harter, in press). Despite this interest, relatively little attention has been devoted to the sensitive measurement of such constructs, particularly in the young child (see also Wylie, 1979). The present article describes the construction of a new scale designed to assess perceived competence and social acceptance in young children, ages 4–7.

Our conceptual approach to the assessment of self-judgments has been domain-specific, unlike the frameworks adopted by other test constructors (e.g., Coopersmith, 1967; Fiers & Harris, 1969), who have sought to assess self-concept or self-esteem primarily through the calculation of a single score, summing items across diverse domains. In contrast, we have sought to assess children's self-judgments separately within specific domains in order to provide a profile of self-perceptions across these domains. This approach is based on the assumption that children do not view themselves as equally adequate in all domains; rather, we have assumed that they are capable of making meaningful distinctions between different domains.

Support for this assumption has been obtained in measurement efforts with children and adolescents between the ages of 8 and 18. In the construction of the Self-Perception Profile for Children (Harter, Note 1), we demonstrated that children differentiate the following five domains as revealed through factor-analytic procedures: scholastic competence, athletic competence, peer acceptance, physical appearance, and conduct or behavior. In addition to judgments in these specific domains, children aged 8 and older can also make reliable judgments about their general worth as a person. Thus, the current structure of the Self-Perception Profile for older children contains five separate domain-specific subscales as well as a sixth subscale tapping self-worth.

In more recent efforts to devise a self-
report instrument for young children, we adopted a similar approach in that we sought to identify meaningful domains in the child's life and to construct separate subscales for each. We also opted for utilization of a similar type of question format that (a) provides a greater range of responses for each item (four choices rather than the more typical two-choice true/false format) and (b) reduces children's tendency to give the socially desirable response (see Harter [1982] for a more complete description).

However, in devising a developmentally appropriate downward extension of the scale for 4-7-year-olds, several aspects of our procedure were different from those used with the older children. First, a pictorial format was devised rather than a written questionnaire. Experience shows that, on questionnaires, young children's inability to read as well as to understand the items, coupled with related attentional problems, attenuates both the reliability and validity of such instruments. In contrast, the pictorial format engages the young child's interest, is understandable, sustains the child's attention, and leads to more meaningful responses.

The pictorial format also allows us to depict skills and specific activities concretely. Whereas, at older ages, trait labels and general descriptions of skill or adequacy can be employed—such as terms like smart, popular, athletic, and good-looking—the young child has not yet acquired these forms of self-description (see Harter [1983] for a theoretical discussion of these developmental shifts). Rather, the young child's self-judgments involve the behavioral description of their specific abilities, such as completing puzzles, running fast, and playing with friends. Therefore, the graphic presentation of these actions and activities facilitates the young child's understanding of the task since these forms of self-description are developmentally appropriate.

Another difference involves the number of versions of the scale required. At older ages, one version can be utilized across a wide range of ages. For the younger ages, however, it was necessary to devise one version for preschoolers and kindergartners (4- and 5-year-olds) and a separate version for first and second graders (6- and 7-year-olds). This was necessitated by the fact that the specific skills that define or connote competence and social acceptance change rather dramatically within this 4-year age range. For example, puzzles may be indicative of cognitive competence during the preschool and kindergarten years, but more scholastically oriented skills such as being able to spell, read, or add are better measures of cognitive competence in the first and second grades.

The younger child's instrument also differs from the older version in that it contains no self-worth subscale. Both theory (see Harter, 1983) and empirical findings have led to the conclusion that children are not capable of making judgments about their worth as persons until approximately the age of 8. The very concept of "personness" is not yet firmly established among younger children, nor is the notion that the self, so defined, can be evaluated as a global entity.

There is another developmental contrast that involves the degree to which we can expect children's self-judgments to be accurate. Developmental frameworks such as those offered by Piagetians or the proponents of psychoanalytic theory would alert us that the judgments of the young child may not be realistic. That is, young children confuse the wish to be competent with reality; they blur the distinction between their ideal self-image and the real self (Stipek, 1981). Related findings by Ruble and her colleagues (see Ruble, 1983) indicate that it is not until approximately 9 years of age that children make use of social comparison for the purposes of judging their own competence. Thus, certain cognitive limitations appear to interfere with the young child's ability to make realistic judgments about the self.

Given that young children may not be very accurate judges of their competence or social acceptance, comparisons of their scores with objective indexes should not be examined as an index of the validity of the instrument. This lack of convergence is an interesting finding, in and of itself, one that bears on the self-descriptive capabilities during this developmental period. Other forms of validity—such as discriminant, convergent, and predictive validity—would appear to be more appropriate, as will be demonstrated.

The specific content of the scale to be described involves two general constructs, perceived competence and perceived social acceptance. The measure contains two subscales within each of these domains. Perceived competence is divided into two subscales, cognitive competence and physical competence. Social acceptance is divided into two subscales, peer acceptance and maternal acceptance. While these particular subscales appear to define salient domains in the life of the young child, obviously there are others,
three of which (paternal acceptance, teacher acceptance, and conduct) will appear in subsequent versions of this scale.

Given the structure of this scale, we strongly urge that the scale not be viewed as an index of self-concept or self-esteem per se. That is, certain judgments such as the perception of one's cognitive or physical competence or one's behavior may well reflect judgments about the self's capabilities. However, perceptions concerning the degree to which one has friends or obtains support from parents or teachers do not necessarily imply judgments about the adequacy of the self. For example, one may conclude that something about the self is responsible for one's lack of friends. On the other hand, the cause may reside in certain characteristics of one's peers; that is, they are not nice or are not friendly. Similarly, lack of parental support might be because one perceives the self as unlovable, yet, on the other hand, one may perceive one's parents as unloving. Thus, the very basis on which children make such judgments is an interesting empirical issue itself. However, until further research has clarified this issue, we would do well not to assume that all of these seeming self-judgments are based on characteristics that reside in the self. For this reason we have urged that the scale be treated as a measure of what the title indicates, perceived competence and perceived social acceptance, rather than treating it as a singular measure of "self-concept" or "self-esteem."

Method
This particular instrument has undergone numerous revisions in terms of scale structure, item content, and question format, based on extensive piloting with large numbers of subjects. In this article, we will restrict our description to the final version of the subscale instrument.

Subjects
Subjects were 90 preschoolers (mean age = 4.45), 56 kindergartners (mean age = 5.54), 65 first graders (mean age = 6.32), and 44 second graders (mean age = 7.41), approximately equally divided by gender within each group. These samples provided the primary data for the factor analyses, means, standard deviations, internal consistency reliability data, and subscales intercorrelations reported. For an additional sample of 77 preschoolers, 25 kindergartners, and 38 first and second graders, scores for both self-report and teacher ratings were available. All subjects were drawn from schools in middle-class neighborhoods. In terms of ethnic composition, 96% of the children were white, with the remaining 4% Hispanic, black, and Oriental.

Scale Description
Scale structure.—The scale contains four separate subscales—cognitive competence, physical competence, peer acceptance, and maternal acceptance. Each subscale contains six items. There are two versions of the scale, one for preschool-kindergarten and one for first and second grades. These two versions are not completely unique. Rather, for certain subscales, there are overlapping or common items across the two versions. Table 1 provides a master list of all items for each version, grouped according to subscale. Asterisks next to the item number indicate which items are common to both forms.

As can be seen in Table 1, none of the items that define the cognitive subscales at the two developmental levels overlap. The preschool-kindergarten form contains a number of rudimentary readiness skills (knowing colors, the alphabet, being able to count) in addition to performance on puzzles and obtaining stars on papers. The first-second grade version includes those scholastic skills initially encountered in the early primary grades (reading, writing, and arithmetic).

For the domain of physical skills, four items occur on both versions (swinging, climbing, skipping, and running). Two of the preschool-kindergarten skills (tying shoes and hopping), however, are replaced by more advanced physical skills for the first-second grade version (bouncing a ball and jumping rope).

Within the domain of peer acceptance, four of the items involving friends are common across the two versions. Two of the items on the preschool-kindergarten version (staying overnight and eating at friends' houses) are replaced at the first-second grade level by others sharing toys and others sitting next to you. These particular social overtures in the early primary grades would appear to be important indexes of popularity.

For the domain of maternal acceptance, there are four activities or maternal behaviors in common across the two age-graded versions (Mom takes you places you like, cooks your favorite foods, reads to you, and talks to you). Two preschool-kindergarten items drop out (Mom smiles and Mom talks to you) for the first-second grade version and are replaced by Mom lets you eat at friends' and stay overnight. These maternal acceptance items were generated from a list of the most commonly mentioned behaviors by young children in re-
TABLE 1

I ITEMS GROUPED ACCORDING TO SUBSCALE FOR EACH FORM

<table>
<thead>
<tr>
<th>Subscale and Item No.</th>
<th>Preschool-Kindergarten</th>
<th>First-Second Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive competence:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Good at puzzles</td>
<td>Good at numbers</td>
</tr>
<tr>
<td>5</td>
<td>Gets stars on paper</td>
<td>Knows a lot in school</td>
</tr>
<tr>
<td>9</td>
<td>Knows names of colors</td>
<td>Can read alone</td>
</tr>
<tr>
<td>13</td>
<td>Good at counting</td>
<td>Can write words</td>
</tr>
<tr>
<td>17</td>
<td>Knows alphabet</td>
<td>Good at spelling</td>
</tr>
<tr>
<td>21</td>
<td>Knows first letter of name</td>
<td>Good at adding</td>
</tr>
<tr>
<td>Physical competence:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Good at swinging</td>
<td>Good at swinging</td>
</tr>
<tr>
<td>7*</td>
<td>Good at climbing</td>
<td>Good at climbing</td>
</tr>
<tr>
<td>11</td>
<td>Can tie shoes</td>
<td>Good at bouncing ball</td>
</tr>
<tr>
<td>15*</td>
<td>Good at skipping</td>
<td>Good at skipping</td>
</tr>
<tr>
<td>19*</td>
<td>Good at running</td>
<td>Good at running</td>
</tr>
<tr>
<td>23</td>
<td>Good at hopping</td>
<td>Good at jump-roping</td>
</tr>
<tr>
<td>Peer acceptance:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2*</td>
<td>Has lots of friends</td>
<td>Has lots of friends</td>
</tr>
<tr>
<td>6</td>
<td>Stays overnight at friends’</td>
<td>Others share their toys</td>
</tr>
<tr>
<td>10*</td>
<td>Has friends to play with</td>
<td>Has friends to play with</td>
</tr>
<tr>
<td>14*</td>
<td>Has friends on playground</td>
<td>Has friends on playground</td>
</tr>
<tr>
<td>18*</td>
<td>Gets asked to play with others</td>
<td>Gets asked to play with others</td>
</tr>
<tr>
<td>22</td>
<td>Eats dinner at friends’ house</td>
<td>Others sit next to you</td>
</tr>
<tr>
<td>Maternal acceptance:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Mom smiles</td>
<td>Mom lets you eat at friends’</td>
</tr>
<tr>
<td>8*</td>
<td>Mom takes you places you like</td>
<td>Mom takes you places you like</td>
</tr>
<tr>
<td>12*</td>
<td>Mom cooks favorite foods</td>
<td>Mom cooks favorite foods</td>
</tr>
<tr>
<td>16*</td>
<td>Mom reads to you</td>
<td>Mom reads to you</td>
</tr>
<tr>
<td>20</td>
<td>Mom plays with you</td>
<td>Mom plays with you</td>
</tr>
<tr>
<td>24*</td>
<td>Mom talks to you</td>
<td>Mom talks to you</td>
</tr>
</tbody>
</table>

Note.—Item number refers to position of the item in the order administered to the child. Asterisk designates items common to both forms.

response to the question, "Tell me the things your mother does that let you know that she likes or loves you."

Picture plates.—The pictures accompanying each version are bound separately, as are sets for boys and girls. Thus, there are four books of plates, both a boys' and girls' set for the preschool-kindergarten and first-second grade versions. The activities depicted in each item are identical for girls and boys. Only the gender of the target child is different, so that a subject can respond to pictures depicting a same-gender child.

Items occur in the order of cognitive competence, social acceptance, physical competence, and maternal acceptance, and continue to repeat themselves in that order. Within each subscale, items are counterbalanced so that three of the pictures depict the most competent or accepted child on the left and three of them depict the more competent or accepted child on the right.

Sample item.—The scale is individually administered. A sample item is presented in Figure 1. The child is first read a brief statement about each child depicted. For the sample item, the female subject would be told that the girl on the child's left is good at puzzles but the child on the right is not very good at puzzles. The child's first task is to indicate which of the two girls she is most like. After making that decision, the child is then asked to think only about the picture on that side and indicate whether she is a lot like that girl (the big circle) or just a little bit like that girl (the smaller circle). For each item there are more specific descriptive questions that accompany each circle, such as "Are you just pretty good at puzzles [small circle] or really good [large circle]?" The book of plates is constructed so that, as the picture for a given item is presented to the child, the item description to be read by the examiner sitting opposite the child is printed on the back of the preceding picture (see manual [Harter & Pike, Note 2] for more specific instructions).

Scoring.—Each item is scored on a four-point scale, where a score of 4 would be the most competent or accepted and a score of 1 would designate the least competent or accepted. Thus, for the sample item, the child who indicates that she is a lot like the girl on the left who is good at puzzles would receive
a score of 4. If she chose the smaller circle on the left, she would get a 3. If she indicates that she is a little like the girl on the right who is not very good at puzzles, she would receive a 2. And if she is a lot like that girl, she would get a score of 1. (These scores are designated on a scoring key under the verbal descriptions provided for the examiner for each item in the picture plates.) Item scores are averaged across the six items for a given subscale, and these four means provide the child’s profile of perceived competence and social acceptance.

Teacher rating scale.—A teacher rating scale parallels the child’s instrument. Teachers rate the child in three of the four areas tapped on the child’s version: cognitive competence, physical competence, and peer acceptance. (We did not feel that it was appropriate to have teachers rate the maternal acceptance of the child.) On this scale, teachers are given a brief verbal description of each item (e.g., good at puzzles, has lots of friends, good at swinging) and then rate how true that statement is on a four-point scale (really true, pretty true, only sort of true, and not very true). Thus, these scores can be compared with the child’s scores, depending on the purposes of the study.

Results

The primary results bear upon the psychometric properties of the scale. To determine the factorial validity of the scale, the factor pattern will first be presented, along with item means and standard deviations.

Subscale reliabilities, in the form of internal consistency coefficients, will then be presented, followed by subscale means and standard deviations. Intercorrelations among subscales, as well as correlations between child and teacher ratings, will then be described.

Factor Pattern

Tables 2 and 3 present the factor pattern based on an oblique (promax) rotation, a solution that allows the factors to intercorrelate. This solution was considered the most appropriate given our expectation, based on previous findings, that there would be moderate and meaningful correlations among self-judgments in these domains. Cattell's 'scree' text, based on the magnitude of the eigenvalues, as well as interpretability, indicated that a two-factor solution best described the data from both the combined preschool-kindergarten samples as well as the combined first-second grade samples.¹

As can be seen in Tables 2 and 3, for both groups, items generally have moderate to high loadings on their designated factor, and with two exceptions for the preschool-kindergarten sample, items do not cross-load on the other factor. Loadings are somewhat higher for the first-second grade samples. (Loadings less than .19 are not presented, for the sake of clarity.) Factor 1 is defined by the two competence subscales, cognitive and physical; thus it is considered to reflect perceptions of general competence. Factor 2 is defined by the peer acceptance and maternal acceptance sub-

¹ Initially we performed the more traditional orthogonal rotation, which also revealed a two-factor solution. However, the oblique rotation not only seemed more appropriate conceptually but provided a somewhat better fit.
TABLE 2
FACTOR PATTERN AND ITEM MEANS AND STANDARD DEVIATIONS FOR THE PRESCHOOL AND KINDERGARTEN SAMPLES COMBINED

<table>
<thead>
<tr>
<th>Subscale, Item No., and Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive competence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Good at puzzles</td>
<td>.39</td>
<td></td>
<td>3.2</td>
<td>.77</td>
</tr>
<tr>
<td>5. Gets stars on paper</td>
<td>.37</td>
<td></td>
<td>3.1</td>
<td>.95</td>
</tr>
<tr>
<td>9. Knows names of colors</td>
<td>.57</td>
<td></td>
<td>3.6</td>
<td>.60</td>
</tr>
<tr>
<td>13. Good at counting</td>
<td>.43</td>
<td></td>
<td>3.6</td>
<td>.61</td>
</tr>
<tr>
<td>17. Knows alphabet</td>
<td>.48</td>
<td></td>
<td>3.6</td>
<td>.57</td>
</tr>
<tr>
<td>21. Knows first letter of name</td>
<td>.58</td>
<td></td>
<td>3.6</td>
<td>.62</td>
</tr>
<tr>
<td>Physical competence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Good at swinging</td>
<td>.19</td>
<td></td>
<td>3.6</td>
<td>.84</td>
</tr>
<tr>
<td>7. Good at climbing</td>
<td>.33</td>
<td></td>
<td>3.4</td>
<td>.77</td>
</tr>
<tr>
<td>11. Can tie shoes</td>
<td>.42</td>
<td></td>
<td>2.8</td>
<td>1.12</td>
</tr>
<tr>
<td>15. Good at skipping</td>
<td>.34</td>
<td></td>
<td>3.4</td>
<td>.84</td>
</tr>
<tr>
<td>19. Good at running</td>
<td>.23</td>
<td></td>
<td>3.4</td>
<td>.75</td>
</tr>
<tr>
<td>23. Good at hopping</td>
<td>.22</td>
<td></td>
<td>3.4</td>
<td>.75</td>
</tr>
<tr>
<td>Peer acceptance:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Has lots of friends</td>
<td>.36</td>
<td></td>
<td>3.2</td>
<td>.79</td>
</tr>
<tr>
<td>6. Stays overnight at friends’</td>
<td>.47</td>
<td></td>
<td>3.1</td>
<td>.92</td>
</tr>
<tr>
<td>10. Has friends to play games with</td>
<td>.23</td>
<td></td>
<td>3.1</td>
<td>.86</td>
</tr>
<tr>
<td>14. Has friends on the playground</td>
<td>.36</td>
<td></td>
<td>3.2</td>
<td>.79</td>
</tr>
<tr>
<td>18. Gets asked to play with others</td>
<td>.44</td>
<td></td>
<td>3.1</td>
<td>.81</td>
</tr>
<tr>
<td>22. Eats dinner at friends’ house</td>
<td>.61</td>
<td></td>
<td>2.7</td>
<td>1.01</td>
</tr>
<tr>
<td>Maternal acceptance:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Mom smiles</td>
<td>.52</td>
<td></td>
<td>3.3</td>
<td>.67</td>
</tr>
<tr>
<td>8. Mom takes you places you like</td>
<td>.52</td>
<td></td>
<td>3.1</td>
<td>.80</td>
</tr>
<tr>
<td>12. Mom cooks favorite foods</td>
<td>.53</td>
<td></td>
<td>3.0</td>
<td>.75</td>
</tr>
<tr>
<td>16. Mom reads to you</td>
<td>.61</td>
<td></td>
<td>3.0</td>
<td>.96</td>
</tr>
<tr>
<td>20. Mom plays with you</td>
<td>.70</td>
<td></td>
<td>2.5</td>
<td>1.04</td>
</tr>
<tr>
<td>24. Mom talks to you</td>
<td>.62</td>
<td></td>
<td>3.1</td>
<td>.91</td>
</tr>
</tbody>
</table>

Note.—N = 145.

scales; it is considered to reflect perceptions of social acceptance.

Item Means and Standard Deviations
As can be seen in Tables 2 and 3, the majority of means are in the range of 3.0–3.6, indicating that young children tend to report relatively positive feelings of competence and acceptance. Standard deviations indicate that there is still considerable variability, even though judgments are being made in the upper ranges of the scale. The use of the upper ranges is not thought to reflect social desirability response tendencies so much as the young child’s blurring of the boundaries between reality and the wish to be competent or accepted, as anticipated.

Reliability
Subscale reliabilities, presented in Table 4, were assessed by employing coefficient α that provides an index of internal consistency. If one looks at individual subscales, it can be seen that these values range from .50 to .85. When one combines subscales according to their designated factors, these reliabilities increase substantially, falling within a range of .75–.89. The reliability of the total scale, all 24 items, is in the mid- to high .80s.

It should be noted that, since the item means for the competence subscales, in particular, were skewed toward the upper end of the scale, the range of scores was restricted, which in turn attenuated the magnitude of these reliability estimates. That is, the overwhelming majority of children’s item scores were either 3 or 4. Paradoxically, therefore, although children responded consistently to these items in terms of scores at the upper end of the scale, the restricted range necessarily leads to lower statistical estimates of reliability.

Subscale Means and Standard Deviations
The subscale means and standard deviations are presented in Table 5. These subscale means, like the item means, are skewed in the direction of positive judgments, reflecting the tendency for young children to report relatively positive feelings of competence and social acceptance. Scores are somewhat higher for the two competence subscales, compared to the two social acceptance subscales. Consistent with this pattern, the standard deviations are somewhat lower for the competence than the social acceptance subscales.

A 4 × 4 (group × subscale) analysis of variance, with subscale as a repeated mea-
Harter and Pike 1975

**TABLE 3**

**FACTOR PATTERN AND ITEM MEANS AND STANDARD DEVIATIONS FOR THE FIRST- AND SECOND-GRADE SAMPLES COMBINED**

<table>
<thead>
<tr>
<th>Subscale, Item No., and Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive competence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Good at numbers</td>
<td>.51</td>
<td></td>
<td>3.2</td>
<td>.73</td>
</tr>
<tr>
<td>5. Knows a lot in school</td>
<td>.63</td>
<td></td>
<td>3.5</td>
<td>.64</td>
</tr>
<tr>
<td>9. Can read alone</td>
<td>.50</td>
<td></td>
<td>3.4</td>
<td>.73</td>
</tr>
<tr>
<td>15. Can write words</td>
<td>.65</td>
<td></td>
<td>3.6</td>
<td>.58</td>
</tr>
<tr>
<td>17. Good at spelling</td>
<td>.51</td>
<td></td>
<td>3.4</td>
<td>.65</td>
</tr>
<tr>
<td>21. Good at adding</td>
<td>.40</td>
<td></td>
<td>3.5</td>
<td>.62</td>
</tr>
<tr>
<td>Physical competence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Good at swinging</td>
<td>.22</td>
<td></td>
<td>3.7</td>
<td>.60</td>
</tr>
<tr>
<td>7. Good at climbing</td>
<td>.48</td>
<td></td>
<td>3.4</td>
<td>.60</td>
</tr>
<tr>
<td>11. Good at bouncing ball</td>
<td>.43</td>
<td></td>
<td>3.5</td>
<td>.71</td>
</tr>
<tr>
<td>15. Good at skipping</td>
<td>.33</td>
<td></td>
<td>3.7</td>
<td>.63</td>
</tr>
<tr>
<td>19. Good at running</td>
<td>.50</td>
<td></td>
<td>3.4</td>
<td>.70</td>
</tr>
<tr>
<td>23. Good at jumping rope</td>
<td>.40</td>
<td></td>
<td>3.1</td>
<td>1.02</td>
</tr>
<tr>
<td>Peer acceptance:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Has lots of friends</td>
<td>.67</td>
<td></td>
<td>3.1</td>
<td>.85</td>
</tr>
<tr>
<td>6. Others share their toys</td>
<td>.27</td>
<td></td>
<td>3.3</td>
<td>.78</td>
</tr>
<tr>
<td>10. Has friends to play games with</td>
<td>.69</td>
<td></td>
<td>3.0</td>
<td>.90</td>
</tr>
<tr>
<td>14. Has friends on the playground</td>
<td>.67</td>
<td></td>
<td>3.2</td>
<td>.89</td>
</tr>
<tr>
<td>18. Gets asked to play with others</td>
<td>.72</td>
<td></td>
<td>3.1</td>
<td>.85</td>
</tr>
<tr>
<td>22. Others sit next to you</td>
<td>.67</td>
<td></td>
<td>3.1</td>
<td>.81</td>
</tr>
<tr>
<td>Maternal acceptance:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Mom lets you eat at friends’ meal</td>
<td>.44</td>
<td></td>
<td>2.8</td>
<td>.89</td>
</tr>
<tr>
<td>8. Mom takes you places you like</td>
<td>.58</td>
<td></td>
<td>3.1</td>
<td>.95</td>
</tr>
<tr>
<td>12. Mom cooks favorite foods</td>
<td>.63</td>
<td></td>
<td>3.1</td>
<td>.77</td>
</tr>
<tr>
<td>16. Mom reads to you</td>
<td>.61</td>
<td></td>
<td>2.7</td>
<td>1.13</td>
</tr>
<tr>
<td>20. Mom lets you stay overnight</td>
<td>.51</td>
<td></td>
<td>2.9</td>
<td>1.01</td>
</tr>
<tr>
<td>24. Mom talks to you</td>
<td>.50</td>
<td></td>
<td>3.0</td>
<td>.94</td>
</tr>
</tbody>
</table>

Note.—N = 104.

that, since the item scores were restricted, the standard deviation of the mean was not the best estimate of reliability. The standard deviations for the subscales and standard deviation of the mean are presented in Table 5. These subscale means, reflecting the child's report of competence and social acceptance, were somewhat higher than the item scores. The subscale means reflected the group's perception of competence and social acceptance. The standard deviation for the competence subscale was higher than for the item scores, which was not the case for the younger groups. F(9, 693) = 5.85, p = .01.

**Intercorrelations among Subscales**

Table 6 presents the intercorrelations among the four subscales for each of the four groups. The correlations obtained were consistent with the factor analysis, for the two competence scales to intercorrelate moderately for each group and for the two social acceptance subscales to intercorrelate somewhat more highly. Among preschoolers, kindergartners, and first graders, peer acceptance also correlates moderately with both cognitive and physical competence, as does maternal acceptance. The correlations between physical competence and the two social acceptance subscales for the second graders, however, did not fit this pattern.

**Correlations between Child and Teacher Ratings**

The intercorrelations between child and teacher judgments were calculated across all subjects since differences between age groups were small. These values were .37 (p < .001) for cognitive competence, .30 (p < .005) for physical competence, and .06 for social acceptance. While these correlations are moderately weak, the pattern indicates that agreement between pupil and teacher is highest in the cognitive competence domain, next highest in the physical domain, and virtually negligible for peer acceptance.

**Validity Data**

*Convergent validity.*—As one index of the validity of children's judgments, we conducted an inquiry to determine the measure had been administered asking children the bases for their responses in the two competence domains. Children were asked, “How do you know you are good at/not good at [depending on the child's initial response] this [activity...
<table>
<thead>
<tr>
<th></th>
<th>Cognitive</th>
<th>Physical</th>
<th>Combined</th>
<th>Peer</th>
<th>Maternal</th>
<th>Combined</th>
<th>Total Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool (N = 90)</td>
<td>.71</td>
<td>.66</td>
<td>.79</td>
<td>.74</td>
<td>.85</td>
<td>.86</td>
<td>.89</td>
</tr>
<tr>
<td>Kindergarten (N = 56)</td>
<td>.52</td>
<td>.53</td>
<td>.66</td>
<td>.75</td>
<td>.81</td>
<td>.87</td>
<td>.86</td>
</tr>
<tr>
<td>Preschool and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kindergarten combined</td>
<td>.67</td>
<td>.62</td>
<td>.76</td>
<td>.74</td>
<td>.83</td>
<td>.87</td>
<td>.88</td>
</tr>
<tr>
<td>First grade (N = 65)</td>
<td>.71</td>
<td>.50</td>
<td>.75</td>
<td>.78</td>
<td>.72</td>
<td>.84</td>
<td>.87</td>
</tr>
<tr>
<td>Second grade (N = 44)</td>
<td>.79</td>
<td>.62</td>
<td>.80</td>
<td>.83</td>
<td>.78</td>
<td>.89</td>
<td>.85</td>
</tr>
<tr>
<td>First and second</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grades combined</td>
<td>.76</td>
<td>.53</td>
<td>.77</td>
<td>.79</td>
<td>.74</td>
<td>.86</td>
<td>.87</td>
</tr>
</tbody>
</table>

TABLE 4
INTERNAL CONSISTENCY RELIABILITIES FOR EACH GROUP
TABLE 5
Subscale Means and Standard Deviations for Each Group

<table>
<thead>
<tr>
<th>Competence Subscales</th>
<th>Acceptance Subscales</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cognitive</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Preschool (N = 90)</td>
<td>3.4</td>
</tr>
<tr>
<td>Kindergarten (N = 56)</td>
<td>3.6</td>
</tr>
<tr>
<td>Preschool and</td>
<td></td>
</tr>
<tr>
<td>kindergarten combined</td>
<td>3.5</td>
</tr>
<tr>
<td>First grade (N = 65)</td>
<td>3.4</td>
</tr>
<tr>
<td>Second grade (N = 44)</td>
<td>3.5</td>
</tr>
<tr>
<td>First and second</td>
<td></td>
</tr>
<tr>
<td>grades combined</td>
<td>3.4</td>
</tr>
</tbody>
</table>

TABLE 6
Intercorrelations Among Subscales for Each Group

<table>
<thead>
<tr>
<th></th>
<th>Cognitive Competence</th>
<th>Physical Competence</th>
<th>Peer Acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical competence:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preschool</td>
<td>.56***</td>
<td>.46***</td>
<td></td>
</tr>
<tr>
<td>Kindergarten</td>
<td>.43***</td>
<td>.42***</td>
<td></td>
</tr>
<tr>
<td>First grade</td>
<td>.55***</td>
<td>.50***</td>
<td></td>
</tr>
<tr>
<td>Second grade</td>
<td>.43**</td>
<td>.08</td>
<td></td>
</tr>
<tr>
<td>Peer acceptance:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preschool</td>
<td>.56***</td>
<td>.43***</td>
<td>.64***</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>.45***</td>
<td>.50***</td>
<td>.62***</td>
</tr>
<tr>
<td>First grade</td>
<td>.59***</td>
<td>.48***</td>
<td>.66***</td>
</tr>
<tr>
<td>Second grade</td>
<td>.32*</td>
<td>.00</td>
<td>.80***</td>
</tr>
</tbody>
</table>

*p < .05.
**p < .01.
***p < .001.

specified? How can you tell?” The purpose of this procedure was to determine (a) whether children could give reasons, and if so, (b) whether they were compelling in the sense that they bolstered or supported the particular self-judgment they had given previously.

Systematic data (Chao, Harter, Adams, & Strop, Note 3) were available for a sample of 43 first graders and 48 second graders who were asked about three cognitive skills (reading, spelling, and writing) and two physical skills (climbing and running). For the cognitive skills, 96% of the children readily gave specific reasons for why they felt that they were competent or not competent. These fell into the following five categories for which sample responses are provided: (a) performance feedback (19%): “I get the hardest words right on a test,” “the teacher tells me”; (b) possess/perform component skills (11%): “I sound out the letters,” “I memorize the words,” “I draw straight letters”; (c) specific demonstrations (22%): “I started reading when I was 3,” “I can write words like ‘cat’ and ‘dog,’” “I can read two whole books,” “I can write in handwriting,” “I don’t have to read out loud, I can think it up in my mind”; (d) routes to developing skills (20%): “I practice a lot,” “I can spell ‘cause I read a lot,” “I practice on my flash cards,” “My mom and dad helped me learn how”; and (e) habitual activity (14%): “I read a lot at home,” “I do writing every day,” “I’ve spelled a lot before.”

In the physical domain, the categories differed somewhat from the cognitive domain, and there were also within-domain differences in the percentage of responses for
climbing and running. The categories, including sample responses, were as follows: (a) social comparison (51% for running, 2% for climbing): "I got to race a lot and win," "I was first place in running in gym," "The boys say they can run faster but at races I can beat them," "I'm the best in climbing races"; (b) habitual activity (18% for running; 31% for climbing): "I run a lot," "I practice jogging," "I do a lot of feet work in the gym," "I climb a lot on bars," "I've been climbing for a long time"; (c) specific demonstrations (25% and 52%, respectively): "I can run around the block a couple of times," "I run a lot in football and tag 'em down," "I climb up to my treehouse," "I can do chin-ups"; (d) somebody teaches (1% and 40%, respectively): "Mom taught me," "Somebody taught me to climb the jungle gym"; and (e) don't injure self (4% and 3%, respectively): "I hardly ever trip and fall," "I don't fall off and scratch myself." (A total of 96% of the responses about running and 97% of the responses about climbing could be coded in these categories.)

Therefore, for both the cognitive and physical domains, the findings demonstrate that children can provide very definite reasons for their alleged competencies; moreover, they volunteer these readily. Although we do not have systematic data for the younger ages, many of these children spontaneously elaborated on (and sometimes demonstrated) their prowess during the course of the normal administration, and these comments reveal that they too have specific reasons for their judgments. Furthermore, although the sample responses presented were those offered by children judging themselves to be competent, children rating themselves as incompetent also gave plausible responses (e.g., "I can't spell words on tests," "I draw crooked letters," "I watch too much TV," "I can't do twirls on the jungle gym," "I'm the last when we run"). Therefore, the overall pattern is one of convergence between the initial perceived competence judgments and the reasons children offered for these perceptions.

**Discriminant validity.**—As one test of validity in the cognitive domain, we made the prediction that children held back in first grade for academic reasons should score lower on the cognitive competence subscale than those who were promoted to the second grade. Over a 2-year period we identified 12 children who had been held back, and we compared these children with a sample of 12 children, matched on age and gender, from the pool of those who had been promoted. The cognitive competence scores of those held back (mean = 2.4) were found to be significantly lower than the scores (mean = 3.3) of those who were promoted, t\(22\) = 3.5, \(p < .003\).

For the social domain, we examined the perceived peer acceptance scores for kindergartners and first and second graders who recently moved and who had attended this particular school for less than 2 months. We hypothesized that these children would have lower peer acceptance scores than children who had been in the school for a minimum of 1 year. The scores of the 10 "new" children were significantly lower (mean = 2.9) than a comparison group of children, matched for age and gender, whose scores (mean = 3.3) indicated greater peer acceptance, t(18) = 2.7, \(p < .01\).

In the physical domain, we have examined the validity issue with regard to the scores of children who were preterm infants. Prematurity is frequently associated with developmental lags in gross motor skills. From preschools that had participated in our studies, we were able to obtain information from teachers as to which of the children they were certain had been born preterm. These were compared with a sample of preschoolers who were known to have been full-term infants. Group differences in teacher ratings for the physical domain were considerably lower for the preterm group (mean = 2.3) compared with the full-term group (mean = 3.1), t(14) = 3.4, \(p < .005\). Correspondingly, the physical competence scores of the eight children who had been preterm infants were found to be lower (mean = 2.8) than the scores (mean = 3.3) of children who had been full-term infants, t(14) = 2.9, \(p < .01\).

We have begun to examine the validity of the maternal acceptance subscale in one study of childhood depression (Harter & Wright, Note 4). Our prediction was that depression in young children (defined in terms of dysphoric mood and lack of energy or interest) would be directly related to lack of maternal acceptance. In this study, we did not have a group of severely depressed children. However, within the normal range of scores for kindergartners and first and second graders, we found the correlation between our depression/cheerfulness measure and maternal acceptance to be .48, \(p < .001\).

Finally, although our new paternal acceptance scale has not yet been integrated into the versions reported on in this paper, an interesting study has just been completed on young children with abusive fathers (Kelby, Note 5). Examining 11 such children, it was found that their fathers' acceptance scores...
the scores (mean = 2.6) were significantly lower than the scores of children from the same preschool, t(22) = 3.4, p < .005.

Predictive validity.—In one study, Bierer (1981) examined the relationship between first and second graders' perceived cognitive competence and their preference for challenge on a behavioral task involving subjects' choice of puzzles, varying in difficulty level. It was initially hypothesized that perceived cognitive competence would predict difficulty-level preferences. This correlation (r = .42) was significant, p < .005. However, further examination indicated that it was attenuated because a subgroup of children was present whose perceptions of their competence appeared to be inflated—that is, whose scores were at least 1.2 higher on a four-point scale than their teacher's ratings of their cognitive competence on the same items. This subgroup tended to select puzzles that were much easier than one would expect, based on their perceived competence, although their choices were consistent with their actual competence, as judged by the teacher. In terms of the validity question, these findings revealed that, for pupils whose ratings are either congruent or lower than the teachers, their perceived cognitive competence is predictive of their actual behavior. That is, these perceptions appear to mediate their behavioral preference for challenge. However, the presence of overratings in the sample attenuates the predictive validity of this subscale.

Correlations between child and teacher ratings.—In the introduction, it was suggested that the tendency of young children to be somewhat inaccurate observers of their own competencies does not necessarily indict the validity of the instrument. The findings presented indicate that the correlations between self- and teacher ratings in the two competence domains are significant, although they are moderately weak, at best, consistent with our expectation. Nevertheless, we did find that, for the competence domains, teacher and child ratings were more highly correlated within the same domain (cognitive = .37, physical = .30) than they were across the two domains (teacher-cognitive/pupil-physical = .11; teacher-physical/pupil-cognitive = .16). Thus, while we have not relied heavily on this type of external validity, the pattern suggests that children's competence judgments are related to their actual competence.

Moreover, when one examines the perceived competence scores of children whom the teachers rate as in the top and bottom quartiles (3.6 and 2.3, respectively, for the cognitive domain), differences in the perceived cognitive competence of these two subgroups are highly significant (3.8 vs. 2.6), t(102) = 5.9, p < .001. Thus, for children who fall at either end of the competence continuum, there is much more convergence between teacher and child ratings than for those falling within the mid-ranges of the distribution.

Discussion

The attempt to devise a pictorial self-report measure of young children's perceptions of their competence and social acceptance would appear to have been successful. Children eagerly respond to the pictorial format, they comprehend the items, and the psychometric properties of the scale seem sound. The item scores and standard deviations revealed reasonable variability, indicating that the scale is sensitive to individual differences in perceived competence and acceptance among young children.

The reliability, as assessed through indexes of internal consistency, was found to be acceptable. Several forms of validity were also examined. In normative samples, the reasons children gave for their self-perceptions were consistent with their judgments on the items themselves and were quite plausible. This suggests that the ratings are valid, in the sense that young children's self-perceptions of their competencies appear to be based on specific behavioral referents.

The findings also indicated that scores on the various subscales did discriminate between groups of children predicted to differ in each domain. For example, children new to a school setting reported lower peer acceptance than those who had attended the school for a year or more. Children who have been held back a grade for academic reasons reported lower perceived cognitive competence than those experiencing normal promotion. Children who were preterm infants, with related delays in motor development, had lower physical competence scores than children who had been born full-term. Thus, the various subscales would appear to discriminate clearly between a given subgroup for whom there is reason to expect relatively low scores and children, matched for age and gender, from the normative sample. In addition, children judged by teachers to be very competent scored considerably higher than those whom teachers judged to be low in competence.

At a more theoretical level, the factor pattern obtained with this instrument is of inter-
1980 Child Development

...since it provides certain clues with regard to the structure of the young child's self-descriptions. The two-factor solution revealed one factor comprised of the cognitive and physical competence subscales and a second factor comprised of the peer and maternal acceptance subscales. We have labeled these factors "general competence" and "social acceptance." The single competence factor implies that young children do not make a clear distinction between what we identified as cognitive and physical domains. Competence at one type of skill is associated with competence at the other. One is either "good at doing things" or one is not. These skill domains, however, are distinguished from social acceptance by peers and by mother.

The structure of young children's self-perceptions across these domains is less differentiated than the structure obtained for older children (Harter, 1982) where we find that cognitive and physical skills clearly define separate factors. This developmental difference is consistent with findings indicating that the structure of the self becomes more differentiated with age (see Harter, Note 1). In addition, the self-structure would appear to be more highly related to mental age than chronological age. In one study we employed retarded children whose mental ages ranged from 5 to approximately 8, the same range one would expect for the young normal-IQ children in the present study. Consistent with the findings reported here, the retarded pupils did not make a distinction between the cognitive and physical competence domains, whereas social acceptance defined a separate factor (Silon & Harter, in press).

Inspection of the subscale means indicates a general tendency for scores to be skewed toward positive self-evaluations, although this tendency was greater for the two competence subscales than for the two social acceptance subscales. This pattern appears plausible since judgments about one's competencies may be more intimately related to one's appraisal of the self, in contrast to judgments about social acceptance, which may be influenced by one's view of the characteristics of these particular others. Since fantasies about the ideal self intrude upon judgments of the real self at this age level, the competence scores are likely to be somewhat inflated. To the extent that social acceptance items pull for judgments of others, these scores would be expected to be lower since findings have demonstrated that young children are likely to be more critical of others than of the self (Gesell & Ilg, 1946; Stipek, 1981).

With regard to competence judgments, the verbal interview data revealed that self-perceptions of skill were directly tied to specific behaviors emitted by the self. We have yet to inquire about the bases for judgments concerning social acceptance. However, we have collected interview data on the perceived routes to both competence and social acceptance. We have systematically asked children what the child depicted as incompetent or socially unaccepted would have to do to become like the child pictured as competent or accepted. The responses to competence involve self-improvement, primarily through instruction (I learn from the teacher), or personal effort (I practice a lot, try harder, etc.), with a gradual shift toward high levels of personal effort over the age span of 4 to 7 (Chao, Harter, Adams, & Stroup, Note 3).

In contrast, the routes to peer acceptance involve behaviors designed to influence others. Here we find a gradual shift from relatively naive solutions (find friends, just ask people to be your friends) to social strategies such as being nice, helpful, polite, and kind. The spontaneous mention of these strategies increases from a low of 18% among 4–5-year-olds to 46% in first graders and 65% in second graders.

These findings are interesting in light of the recent emphasis on social skills (see Asher, Oden, & Gottman, 1977; Bash & Camp, 1980; Gottman, Gonso, & Rasmussen, 1975; Hartup, 1979, 1983; Ladd & Oden, 1979; Spivack & Shure, 1974). Across these studies it has been revealed that a variety of social skills are associated with peer popularity and social acceptance. Moreover, they reveal that elementary school children possess an awareness of how social skills influence their acceptance by others. Our own findings suggest that, at the youngest ages, children have not yet acquired the knowledge concerning this relationship in the social domain, although they do seem to appreciate the need for skill development in the cognitive and physical domains. Gradually, over the early grades, they come to appreciate the need to employ social strategies in order to obtain friends.

This pattern raises numerous questions for further study. For example, in what ways are social skills different from those in the cognitive and physical domains? Why should the

---

2 In the revision of the original perceived competence scale, we have determined that children age 8 and older make distinctions among five domains (scholastic, athletic, appearance, social acceptance, and conduct), as revealed by a clean five-factor solution (Harter, 1983).
of the self (Gesell 1934) suggested that self-concept judgments, revealed that self-esteem is directly tied to the self. We would be surprised if the bases for judgments of self-esteem were found to be the same as those for judgments of social acceptance. However, we would be surprised if a measure of self-esteem were not treated solely as a measure of self-esteem. In view of these considerations, we would urge that the instrument not be treated solely as a measure of self-concept since the social acceptance construct tapped may involve a number of dimensions extending beyond one's perceptions of the self per se.

Differences between the two competence domains and the social domain of peer acceptance were also revealed in the examination of the accuracy of children's judgments. The correlations between teacher and child ratings for the two competence domains were significant, though moderately weak; however, they were negligible in the social domain. The greater congruence in the competence domain may be because of clearer sources of information on which to base one's judgments. Our findings suggest that, in the cognitive domain, performance feedback is beginning to emerge as a criterion for performance, whereas in the physical domain, social comparison is becoming the basis for judgments of competence. Of particular interest is the finding that social comparison is used more frequently for the activity of running, which appears to be a more competitive activity, than for climbing.

Our data are consistent with Ruble and Frey's (Note 6) findings that, in the domain of academic achievement, social comparison is not consistently employed in the early grades. Our findings also indicate that social comparison effects may be somewhat domain-specific since social comparison does form the basis for judgments of certain physical skills. We have yet to examine the bases on which children make judgments in the social domain. However, the lack of congruence between child and teacher ratings of social acceptance may result from several factors: performance feedback may be less salient, children may be less able to employ social comparison in this domain, and/or children and adults may employ different criteria. Further research in this area would be fruitful.

Finally, to what uses might such an instrument be put, particularly given the qualification that young children's judgments are not very accurate? First, among normative samples, scores may be useful in predicting behaviors, motivations, and/or emotional reactions of interest. Our own findings (Harter & Wright, Note 4) indicate that the social acceptance subscales, particularly the maternal subscale, are significantly correlated with the child's self-reported mood as well as energy and interest for age-appropriate activities. To the extent that the child's mood and energy level may be critical mediators of behaviors leading to the development of new skills, it would be important to assess these predictors as early as possible. Furthermore, a domain-specific measure allows one to determine which domains best predict the mediators and behaviors of interest.

Second, the degree to which a young child's judgments are inaccurate might be important to examine. Normative or age-appropriate distortions may not be cause for concern. However, it may be important to detect either extreme inflation of one's abilities or the unrealistic portrayal of oneself as extremely incompetent. Furthermore, a child may show these inaccuracies in some domains more than others, and these particular distortions may well have behavioral correlates. Findings with somewhat older children indicate that, by third and fourth grades, there are behavioral patterns associated with extreme tendencies to overrate or underrate one's cognitive competence. For example, both of these inaccurate subgroups tend to avoid behavioral preference for challenge compared with those children who accurately rate their competence (Bierer, 1981).

Third, there would appear to be a need for an instrument to assess the self-perceptions among special subgroups of children who may be under particular types of stress. Children of divorce, of abusive parents, and with learning disabilities or physical handicaps are all special groups that have come to the attention of basic researchers, clinicians, and those engaged in social policy. However, as has been pointed out, not all children necessarily suffer from events that have been categorically identified as stressful (Garney & Rutter, 1983). A variety of individual difference variables, including self-concept, have been implicated as factors influencing the child's ability to withstand stress and cope adaptively. Thus, a domain-specific measure might well be useful in predicting children's reactions with an eye toward determining which type of profile is associated with resiliency and adaptation, or its counterpart. In conclusion, although there are several theoretical issues requiring further research, we believe that there are a number of uses to which this instrument might well be put in order to illuminate our understanding of the young child.

Reference Notes

1. Harter, S. Supplementary description of the Self-Perception Profile for Children: Revision
1982  Child Development


References
The Relationship of Five-Year-Olds' Academic Readiness and Perceptions of Competence and Acceptance

PEGGY L. ANDERSON
University of New Orleans

PHYLLIS J. ADAMS
University of Denver

ABSTRACT Research with young children suggests that the relationship between academic readiness achievement and self-perceptions of competence and social acceptance is weak, if present at all. However, most researchers have used teacher judgment as the criterion for achievement and/or a global measure of self-concept as the criterion of competence. It was postulated that academic readiness measured by a standardized test and self-perception assessed by a multidimensional evaluation instrument would result in a stronger relationship than formerly reported. The subjects, 76 5-year-olds, were administered the Woodcock-Johnson Psycho-Educational Battery and The Pictorial Scale of Perceived Competence and Acceptance for Young Children. The significant relationship found between cognitive competence and academic readiness suggests that 5-year-olds are more realistic in assessing their own achievement than indicated by previous researchers.

The study of children's perceptions of their social and cognitive status has long been a research topic that has intrigued educators and psychologists. One particular facet of this type of research that has enjoyed considerable attention concerns the relationship of self-evaluation of personological variables and cognitive achievement. The discrepancies and consistencies that appear to exist between perceived status and actual status have been explored from a multiplicity of theoretical frameworks and research design approaches and exhaustively described by researchers such as Silvernail (1981) and Wylie (1979). However, in spite of the sizable body of literature that has been accumulated in this area, there is a rather remarkable paucity of research devoted to the study of the young child. This obvious gap in the literature has been commented on by many researchers including Hansford and Hattie's (1982) meta-analysis of 128 studies that explored the relationship between self-evaluation and achievement measures; their analysis included only 5 studies at the preschool level in contrast to the 64 at the elementary level. Likewise, Wylie (1979) in a massive review of self-evaluation literature noted that the study of the child age 5 and under has "barely been touched" (p. 11).

The limited number of studies that have been reported largely reveal that the link between perceptions of self and achievement is a tenuous one for young children. In fact, some researchers have reported that children's ratings of their own ability and their actual ability (as judged by the teacher) have no correlational significance until approximately the second or third grades (Nicholls, 1978; Stipek, 1981). Yet, other researchers have reported significant findings with regard to such a relationship for children in kindergarten (Harter & Pike, 1984; Ozehosky & Clark, 1970; Wattenberg & Clifford, 1964). It should be noted that, in general, the research that has established a significant relationship has done so via modest correlations that suggest a trend of increasing ability over time to make accurate self-evaluations. This trend was exemplified by Hansford and Hattie's (1982) meta-analysis that calculated coefficients of .12, .20 and .27, respectively, for preschool, elementary, and secondary school groups.

Germaine to the study of self-perceptions and abilities is the question of whether or not children as young as 5 years of age have the cognitive capacity to evaluate their strengths and weaknesses. From a Piagetian perspective, it has long been established that children at this preoperational stage of development tend to have rather exaggerated notions of their competence. This tendency to overestimate abilities is often referred to as a developmental characteristic of young children (Harter & Pike, 1984; Parsons & Ruble, 1977; Ruble, 1983). However, this spurious reporting of abilities is a phenomenon that

Address correspondence to Peggy L. Anderson, Department of Special Education, University of New Orleans, New Orleans, LA 70148.
is not the exclusive property of young children but rather appears as a common thread throughout self literature and in older subjects is referred to as self-favorability (Wylie, 1979). Perhaps one of the reasons why this characteristic appears to be so blatant in younger children is that they lack the requisite cognitive and/or linguistic sophistication for handling social nuances in a manner that would mask their perceptions.

In spite of this pronounced tendency for self-preference, there is evidence to indicate that self-evaluation abilities are intact at early ages. Weiner’s (1964) study of 4-year-olds demonstrated that when language was carefully controlled for, children at this age level exhibited an understanding of the evaluation process as well as the ability to evaluate certain abilities, such as drawing skill, in a manner consistent with the teacher’s evaluation. Further support for the emergence of self-evaluation abilities under the age of 5 can be found in the work of Stipek and Hoffman (1980) who discovered that children as young as 3 show evidence of adult-like evaluation skills. However, interestingly enough, these researchers discovered that although preschool children demonstrated this ability, they did not exercise it for evaluating their own skills but rather reserved it for evaluation of their peers.

If it is accepted that 5-year-olds have the cognitive ability for self-evaluation of performance, then the task challenging the researcher is one of methodologically designing an investigatory framework that would facilitate the expression of these abilities. To accomplish this, there are several variables that seem important to address. Firstly, the ambiguous notion of self as a unidimensional construct should be elaborated upon to include specific domains of competence and acceptance that are more comprehensible to children. In recommending this approach, Harter and Pike (1984) noted that children do not view themselves as equally proficient in all domains. For example, cognitive and physical competence may have developed disparately, but this discrepancy would not be revealed in a general self-worth inventory yielding a single score. Secondly, the interviewing technique should be supplemented by stimulus pictures that may facilitate understanding (Harter & Pike, 1984; Perkins & Shannon, 1965). And thirdly, learning achievement should be measured by a statistically reliable instrument instead of relying on teacher ratings, a research technique that has raised questions concerning reliability (Hansford & Hatlie, 1982).

The current study implemented the aforementioned techniques with the hypothesis that 5-year-olds’ academic readiness achievement would be related to perceptions of competence and social acceptance. A secondary purpose was to determine whether school level (preschool and kindergarten) or sex differences would emerge on measures of competence, acceptance, and academic achievement.

Method

Subjects. Subjects in this study were 76 5-year-olds (32 girls and 44 boys) enrolled in preschools or kindergartens in a large suburban area of Maryland. The 42 preschoolers (16 girls and 26 boys) attended school for 3 to 4 hours each day; likewise, the 34 kindergarteners (17 girls and 17 boys) attended daily half-day sessions. The subjects ranged in age from 5 years, 2 months to 5 years, 11 months; mean ages for the preschool and kindergarten groups were 5 years, 4 months and 5 years, 5 months respectively. Family socioeconomic level ranged from lower middle to upper middle class, with the majority tending to concentrate at the upper end. All subjects were Caucasian with the exception of 3 black children.

Procedure. All subjects were individually tested by one of the researchers. Total testing time for each subject averaged about 2 hours.

Academic readiness data were gathered from the three subtest clusters of the Woodcock-Johnson Psycho-Educational Battery that were appropriate for 5-year-olds: the Preschool Scale Cluster from the Tests of Cognitive Ability; the Skills Cluster from the Tests of Achievement; and the Knowledge Cluster, also from the Tests of Achievement. Descriptions of these three subtest clusters and information concerning scores can be found in the manual of the Woodcock-Johnson Psycho-Educational Battery (1977).

The second instrument administered was The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (Harter & Pike, 1980). Unlike other typical self-concept measures, this scale is based on the rationale that self-judgments are domain specific. Thus, self-perceptions are assessed within two separate constructs, perceived competence and social acceptance. The perceived competence construct is defined by two subscales: Cognitive Competence and Physical Competence. In like manner, the social acceptance construct is divided into two subscales: Peer Acceptance and Maternal Acceptance. Each of the four subscales is comprised of 6 items, making a total of 24 items. Normative information and evidence for the validity and reliability of the instrument can be found in Harter and Pike (1984).

The items are presented in a pictorial format, with separate stimulus booklets for males and females. The items in the two booklets are identical with only the gender of the child in the illustrations being different. Each item depicts a specific skill, action, or activity, for example, naming alphabet letters, running, mom reading. A sample item from the booklet for girls is presented in Figure 1.

In administering the items, the examiner first reads aloud a brief statement about each child in the picture, for example, for the sample item: This girl is pretty
good at saying the alphabet (pointing to the picture on the subject's left), and this girl isn't very good (pointing to the one on the right). The subject is then asked to pick the child who is most like him or her. After this decision has been made, the examiner then reads two qualifying statements that pertain to the child selected by the subject. For example, if the subject selected the child who was pretty good at saying the alphabet the examiner would ask: Are you really good at saying the alphabet or are you pretty good? The examiner uses voice inflection to facilitate discrimination as well as further elaboration when necessary. The subject then indicates his or her choice of statement. For each item, specific descriptive qualifying statements are provided.

Each item is scored on a 4-point scale. The scores range from 1 for an indication of the least competence to 4 for the maximum competence. The highest possible score for each of the four subscales is 24. The calculation of a total score by summing items across the four subscales is not appropriate since the rationale underlying the scale is that self-perception assessment is domain specific.

Analysis. To determine the relationship between children's academic readiness and their perceptions of competence/acceptance, Pearson product-moment correlation coefficients were computed between Competence/Acceptance Scale scores and Woodcock-Johnson Psycho-Educational Battery Cluster scores. Possible schooling and gender differences were explored with the two-way analysis of variance procedure. School level (preschool or kindergarten) and sex served as the constant independent variables for this series of ANOVAs, which were calculated with the dependent variables of Competence/Acceptance Scale scores and Woodcock-Johnson Cluster scores.

Results

The results lend partial support to the contention that 5-year-olds' academic readiness is related to their perceptions of competence/acceptance. Cognitive Competence subscale scores (M = 20.7) were found to be significantly related to all three measures of academic readiness (see Table 1): the Preschool Scale Cluster (M = 463.5) and Knowledge Cluster (M = 463.5) both yielded correlation coefficients of .28, p < .01, while the Skills Cluster (M = 412.0) coefficient was slightly higher at .35, p < .001. Physical Competence subscale scores (M = 19.7) were significantly related to only one measure of readiness achievement, the Skills Cluster (r = .29, p < .01). Neither Peer nor Maternal Acceptance subscale scores were significantly correlated with any of the achievement measures.

Table 2 presents a summary of the ANOVA results. The series of two-way ANOVAs yielded four significant findings with respect to school level (preschool or kindergarten). For the Competence/Acceptance subscales analyses, the only subscale that reached significance was that of the Cognitive Competence on which kindergarteners (M = 21.4) rated themselves higher than preschoolers (M = 20.1), F(1, 72) = 5.37, p < .05. Significant school-level differences were dis-

Figure 1.—Sample item from The Pictorial Scale of Perceived Competence and Social Acceptance
Table 1.—Correlations between Competence/Acceptance Subscales and Achievement Clusters

<table>
<thead>
<tr>
<th>Achievement Clusters</th>
<th>Competence/Acceptance Subscales</th>
<th>CC</th>
<th>PC</th>
<th>PA</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td></td>
<td>.28*</td>
<td>.18</td>
<td>-.09</td>
<td>.04</td>
</tr>
<tr>
<td>Skills</td>
<td></td>
<td>.35***</td>
<td>.20*</td>
<td>.04</td>
<td>-.06</td>
</tr>
<tr>
<td>Knowledge</td>
<td></td>
<td>.28*</td>
<td>.18</td>
<td>-.09</td>
<td>.04</td>
</tr>
</tbody>
</table>

*Note. CC = Cognitive Competence; PC = Physical Competence; PA = Peer Acceptance; MA = Maternal Acceptance.
*p < .01. **p < .001.

covered for all three measures of readiness achievement. Kindergarteners (M = 467.1) scored higher than preschoolers (M = 460.6) on the Preschool Scale Cluster, F(1, 72) = 6.10, p < .05; similar results were obtained on the Skills Cluster, F(1, 72) = 12.79, p < .001 (Preschool M = 404.6; Kindergarten M = 421.1), and the Knowledge Cluster, F(1, 72) = 6.10, p < .01 (Preschool M = 406.6; Kindergarten M = 467.15). Sex was found to be a variable that significantly discriminated on only one measure; boys (M = 419.1) scored higher than girls (M = 406.6) on the Skills Cluster, F(1, 72) = 6.22, p < .05. No significant interactions between sex and school level were observed on any of the analyses.

Discussion

The results of this study suggest that 5-year-olds are more realistic in the assessment of their own achievement than previously reported. Thus, the current findings are contra-indicative to the contention that indications of accuracy in self-evaluation of achievement cannot be empirically observed until much later in development. Although the relationship between self-perceptions of cognitive competence and achievement was found to be a modest one, it is substantially higher than would have been expected based on Hansford and Hattie’s (1982) data and the results of work that has determined the absence of such a relationship (Nicholls, 1979; Stipek, 1981).

One explanation for the 5-year-olds in this study exhibiting evidence of logically derived perceptions of their academic achievement probably lies in the nature of the self-perception instrument used. The cognitive items on The Pictorial Scale of Perceived Competence and Acceptance describe specific performance-based activities, for example, good at completing puzzles, good at counting, knows names of colors. It is likely that the cognitive demands placed on the children were developmentally appropriate and enhanced their ability to make realistic judgments. As previously reported by Stipek and Hoffman (1980), even very young children can make accurate performance-based judgments, and in this study there is indication that they exercised this ability for self-evaluation. Unlike typically used self-perception measures (e.g., Coopersmith, 1967; Piers & Harris, 1969) and adaptations of such instruments that address areas of general self-worth and usually consist of verbal yes-no statements, The Pictorial Scale of Perceived Competence and Acceptance specifically relates to multidimensional aspects of self-performance and utilizes a pictorial format suitable for young children. These features, no doubt, contribute to this study’s findings, which contradict previously reported research.

Although somewhat speculative, the results here may reflect a more academically oriented home environment and the influence of achievement-oriented early childhood school curricula. As Elkind (1981) has indicated, many parents and educators are emphasizing intellectual learning at an early age. This may be particularly true for this study since the subjects tended to cluster socioeconomically in the middle to upper middle range, and thus the children’s responses may have been more likely to reflect educational values representative of this group.

If children at the age of 5 do have the ability to realistically assess self-performance, and this study suggests that they do, then this emerging competence should be nurtured as an important aspect of self-awareness development. It would seem that individuals working with young children should provide honest performance feedback, not indiscriminate praise. This is not to suggest that young children do not need encouragement and strong support. However, the large body of self-concept literature that advocates the bolstering of positive perceptions of self irrespective of performance should probably be questioned. It could be that indicating to young children that they are “doing great” on an academic task when in fact they are not performing well could create cognitive dissonance, lead to faulty perceptions of ability, and foster unrealistic goals.

Since Harter and Pike (1984) indicated that a clear distinction between the cognitive and physical domains is not made by young children and that children asso-

Table 2.—F Statistics for Two-Way ANOVAs

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Interaction</th>
<th>Level</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competence/Acceptance Subscales</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive</td>
<td>&lt; 1</td>
<td>5.37*</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Physical</td>
<td>&lt; 1</td>
<td>1.95</td>
<td>2.45</td>
</tr>
<tr>
<td>Peer</td>
<td>1.51</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Maternal</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
<td>1.86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Achievement Clusters</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td>&lt; 1</td>
<td>6.10*</td>
<td>1.83</td>
</tr>
<tr>
<td>Skills</td>
<td>1.81</td>
<td>12.79***</td>
<td>6.22*</td>
</tr>
<tr>
<td>Knowledge</td>
<td>&lt; 1</td>
<td>6.10**</td>
<td>1.83</td>
</tr>
</tbody>
</table>

*Note. dfs for all Fs are 1 and 72.
*p < .05. **p < .01. ***p < .001.
ciate competence in one domain with competence in the other, it was anticipated that physical competence would be significantly correlated with all academic measures. Such was not the case; this study found that physical competence was significantly related to only one of the academic measures, the Skills Cluster. The lack of significant relationships between physical competence and the other two academic measures (i.e., the Preschool Scale and Knowledge Clusters) is somewhat puzzling. This result does not completely support Harter and Pike’s findings and suggests that children at this age may actually be doing some discrimination between the cognitive and physical competence domains—discrimination that becomes apparent when academic measures are taken into consideration.

Significant relationships between 5-year-olds’ perceptions of acceptance (both peer and maternal) and academic readiness were not found in this study. Harter and Pike’s data (1984) also indicate that correlations between teacher and child ratings in the social acceptance domain are negligible. It appears likely that peer social acceptance at this age level may not be highly influential upon academic readiness. Research that has previously examined the relationships between peer acceptance and achievement as well as maternal nurturing and achievement has produced inconclusive results (Wylie, 1979; Purkey, 1970). Of course, it is unquestionably assumed that learning is facilitated when home environment and peer relations are positive. Nevertheless, it is also well accepted that achievement can be attained even under less favorable conditions. On the surface it would appear that this study’s data tend to support the latter of the two possibilities, but it is also quite probable that the explanation may be developmental. Perhaps social acceptance is not related to achievement at the preschool and kindergarten levels but emerges with additional school experiences.

Regarding school level differences, kindergarteners rated themselves higher on cognitive competence than did their preschool counterparts, and they were indeed significantly higher on all three measures of academic readiness. This finding, although expected, cannot be explained as a function of age, since the mean age of both groups was nearly the same. It seems reasonable, therefore, to suggest that the difference is attributable to the school curriculum experienced by the subjects. No doubt, the kindergarten curriculum was more academically oriented than the preschool curriculum and thus the difference in academic achievement. It would appear that the nature of the school program makes a difference in children’s academic attainment and their perceptions of this attainment. Specific to the objective of this study is simply the fact that kindergarteners as a group perceived themselves as more cognitively able than preschooers, and these perceptions were substantiated by the achievement measures.

Consistent with results reported by Wylie (1979) and Hansford and Hattie (1982) was the absence of sex differences regarding self-perceptions. The finding that boys scored significantly higher than girls on only one measure of academic readiness, the Skills Cluster, does not suggest any pattern.

In conclusion, this study indicates that 5-year-olds can realistically assess their own academic achievement. The findings from this study should be used as a basis for the generation of additional research with young children, research needed to address the multitude of still unanswered questions pertaining to self-evaluation abilities of this age group.

REFERENCES

The Family Relationships Project
Phase 1
Form 101

Susan Harter’s The Self-Perception Scale for Children
(Shirley McGuire)

This measure was chosen because it is the best-known
and most widely used at present. Unlike many other
self-esteem scales, Harter’s Scale is designed for
school-aged children and has been used with children in
grades three through nine. Harter also developed a question
format which tries to overcome social desirability problems
associated with self-report inventories. The scale not only
asks about global self-worth, but domains which are
important to most school-aged children also (scholastics,
social life-peers, athletics, physical appearance, and
behavior).

See attached sheets for psychometric data.

References:

Harter, S. (in press). Causes, correlates and the
functional role of global self-worth: A life-span
perspective. J. Kolligian and R. Sternberg (Eds.),
Perceptions of Competence and Incompetence Across the
Life-Span. New Haven, CT: Yale University Press.

Harter, S. (1987). The determinants and mediational role of
global self-worth in children. In N. Eisenberg (Ed.),
Contemporary Issues in Developmental Psychology. New
York: Wiley.

Harter, S. (1985). Competence as a dimensions of
self-evaluation: Toward a Comprehensive model of
Self-worth. In R. Leahy (Ed.), The Development of the

M. Hetherington (Ed.), Handbook of Child Psychology:
Social and Personality Development. (Vol. 4) New York:
Wiley.
of these research areas appears to be related to the assessment instruments employed. That is, few investigators have employed comprehensive assessment batteries capable of assessing a full range of potential symptoms. In order to address this concern, the children completed three instruments designed to tap different areas of disorders.

The Self-Perception Profile for Children

The Self-Perception Profile for Children (Harter, 1983) is a self-report instrument completed by the older sibling, which is designed to assess a child's self-evaluation in terms of discrete domains as well as a general sense of self-worth (see Appendix H). The instrument contains six separate subscales: (1) Scholastic Competence; (2) Social Acceptance; (3) Athletic Competence; (4) Physical Appearance; 5) Behavior/Conduct; and 6) General Self-worth. The measure consists of 36 items, with six items per subscale. An example of an item from this scale is: "Some kids are happy with themselves, but other kids are often not happy with themselves." A "structured alternative format" is presented to the child, who is required to decide which kind of child he or she is most like, the description of a child on the left or on the right side of the page. The child is then asked to make a second decision as to whether the description chosen is "sort of true" or "really true" for him or her. Each item is scored from 1 to 4 where a score of 1 indicates low perceived competence and a score of 4 reflects
high perceived competence. These scores are then summed and averaged, resulting in six subscale means. The items within each subscale are counter-balanced such that half of the items of one subscale are worded with the most positive self-judgment presented first, and half are worded with the most positive statement presented second.

Information about the factor patterns, subscale reliabilities, and correlations among the subscales is currently available based on a sample of 748 sixth and seventh graders. The factor patterns described by Harter (1983) are based on the five specific self-perception domains only. The Self-worth domain was not included in the analyses of factor patterns. Harter (1983) argues that given the conceptualization of self-worth, that is, that a child's judgment of self-worth might be based on information related to the other subscales, one would not expect a clear self-worth factor to result for the factor analyses.

Factor loadings for each of the five sub-scales proved to be substantial. The range of loadings for each of the separate subscales were: (1) Scholastic Competence (range .52 to .67); (2) Social Acceptance (range=.45 to .78); (3) Athletic Competence (range=.59 to .78); (4) Physical Appearance (range=.46 to .72); and (5) Conduct/Behavior (range=.41 to .70). The average cross loadings were very low ranging from .04 to .08.

The subscale reliabilities for the sixth and seventh graders combined reached acceptable levels -- ranging from .75 for Conduct/Behavior to .84 for Self-worth and Athletic competence. Test-retest reliabilities (over a nine-month interval) for 810 New York students ranged from $r = .69$ to $r = .86$. Finally, substantial
correlations were observed for all of the specific domains and the Self-
worth domain ranging from \( r = .44 \) for the Athletic Competence domain to \( r = .64 \) for Physical Appearance.

**Children's Depression Inventory (CDI)**

The CDI (Kovacs, 1981) is a self-report questionnaire designed to assess depression in children between 8 and 16 years of age. The scale is a modification of the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), with most of the items covering content and symptom areas similar to those assessed on the BDI. The CDI differs, however, primarily in the phrasing of the items. In addition, the authors have added several items that pertain to areas of school (e.g., "I have to push myself all the time to do my schoolwork") and social/peer relations (e.g., "I do not have any friends") that are not included on the BDI (see Appendix H).

The CDI is a 27-item questionnaire; however, the suicide ideation item (9) was deleted when administered to the children in this sample, reducing the total number of items to 26. For each item, the child is asked to choose one of three statements which best describes themselves. An example of an item is: "I can never be as good as other kids; I can be as good as other kids if I want to; I am just as good as other kids." The statements are graded in severity in terms of the depressive feelings. Statements selected by the child are then assigned numerical values from 0 to 2. The higher the numerical value, the more clinically severe the symptom being rated. The items were read aloud by the
The Perceived Competence Scale for Children

Susan Harter
University of Denver

HARTER, SUSAN. The Perceived Competence Scale for Children. Child Development, 1982, 53, 87-97. A new self-report instrument, the Perceived Competence Scale for Children, is described. Emphasis is placed on the assessment of a child's sense of competence across different domains, instead of viewing perceived competence as a unitary construct. 3 domains of competence, each constituting a separate subscale, were identified: (a) cognitive, (b) social, and (c) physical. A fourth subscale, general self-worth, independent of any particular skill domain, was included. A new question format was devised which provides a broader range of responses and reduces the tendency to give socially desirable responses. The psychometric properties of the scale are presented for third through ninth grades. Emphasis is placed on its factorial validity. Each subscale defines a separate factor, indicating that children make clear differentiations among these domains. The factor structure is extremely stable across this grade range. The scale is viewed as an alternative to those existing self-concept measures of questionable validity and reliability.

Concepts such as "self-esteem," "self-concept," and "perceived competence" have become central to formulations emerging from social learning theory (Bandura 1977), self-perception theory (Bem 1972), social cognition (Lepper & Greene 1978), and theories of competence and intrinsic motivation (de Charms 1968; Deci 1975; Harter 1978, 1981a, 1981b; White 1959, 1963). At a more applied level, the assessment and enhancement of an individual's self-esteem is critical to clinicians, educators, and program-evaluation researchers.

However, there are serious problems with those instruments designed to tap these self-evaluative processes (see critiques by Harter, in press a; Robinson & Shaver 1973; Wylie 1974, 1979). Typically, constructs such as self-concept and self-esteem are vaguely defined at the conceptual level and therefore do not point to any clear operational definition. Nor is any rationale provided for the potpourri of empirical referents selected. Items on such popular scales as the Coopersmith Self-Esteem Inventory (1967) and the Piers-Harris Self-Concept Scale (Piers & Harris 1964) tap a range of diverse content including cognitive competencies, physical skills, popularity, acceptability by parent, morality, personality traits, physical characteristics, and affective reactions. Responses to these heterogeneous items are then summed, and the total score is interpreted as an index of global self-regard. In employing such a procedure, Cooper-Smith has assumed that children do not make distinctions among the domains in their lives. This assumption was seriously questioned in the scale-construction efforts reported here.

This research was supported by grant no. HD-00613, from the National Institute of Child Health and Human Development, Department of Health and Human Services, U.S. Public Health Service. I would like to acknowledge the extensive cooperation of both the school personnel and pupils from the following school systems, without whose assistance this scale could not have been constructed: Cherry Creek Public School System, Denver Public School System, Jefferson County Public School System, and Colorado Academy, all in Colorado; Ventura School District in Ventura, California; Killingworth Public School System and Wightwood School, in Connecticut; Rochester Public Schools, in conjunction with the research team led by Edward Deci at the University of Rochester. I would also like to thank the many research personnel in our group who worked extensively on the development of the scale. These included Carole Efron, Jim Connell, Karen John, Jeff Baarslag, Barbara Minton, Duane Buhmester, and Kris Chandler, each of whom contributed valuable input into the various revisions of the scale and assisted in administering the instrument to numerous children. Special thanks go to Claire Kott, net to her many hours of coding and keypunching the data. Finally, we are deeply indebted to Bob Engstrom, whose application of his statistical expertise to the data analysis was ultimately responsible for making the scale both empirically and conceptually sound. A manual is available for $9.95. Author's address: Department of Psychology, University of Denver, University Park, Denver, Colorado 80208.
One dimension of self-evaluation, perceived competence, was selected, given my theoretical interest in developing a model of competence motivation (see Harter 1978, 1981a). The emphasis was on the child's perception of his or her competence. The relationship between this perception and one's actual competence was viewed as an empirically interesting question in and of itself.

The model underlying the construction of the Perceived Competence Scale for Children hypothesizes that children do not feel equally competent in every skill domain. We first sought to specify the major competence domains which are relevant to elementary school children. The skills represented by items on existing scales included scholastic performance, peer relationships, and a variety of physical skills. These competencies were consistent with our own observations of the mastery behaviors most salient for elementary school children, and thus they provided an initial framework. Interviews with children were then designed to determine which activities within these domains were particularly important to children in making judgments of competence.

Three separate competence subscales emerged: (a) cognitive competence, with an emphasis on academic performance (doing well at schoolwork, being smart, feeling good about one's classroom performance); (b) social competence, vis-à-vis one's peers (having a lot of friends, being easy to like, and being an important member of one's class); and (c) physical competence, with a focus on sports and outdoor games (doing well at sports, learning new outdoor games readily, and preferring to play sports rather than merely watch others play).

It was also hypothesized that children (age 8 and older) not only make discrete judgments about their competence in different domains, but that by this age they have also constructed a view of their general self-worth as a person, over and above these specific competence judgments. This assumption highlights the hierarchical nature of the self-evaluative process whereby self-esteem or self-worth is viewed as a superordinate construct and competence judgments represent one type of lower-order evaluative dimension. (See Epstein [1973]; Harter [in press]; Rosenberg [1979], each of whom supports this general position; see also Harter [in press] for the developmental rationale which places the emergence of this type of self-theory around the age of 8.)

According to such a model, judgments concerning one's overall self-worth are not inferred from the summation of responses to items tapping a wide array of specific abilities and attributes; rather, they are tapped by items which directly inquire into how much the individual likes himself or herself as a person. Thus, in order to assess these perceptions in children, we included a fourth subscale, general self-worth. Items on this subscale made reference to being sure of oneself, being happy with the way one is, feeling good about the way one acts, and thinking that one is a good person.

In addition to these theoretical considerations, there are certain psychometric problems which have plagued existing measures. Many scale constructors initially include items or subscales designed to tap different domains of the child's life; however, the emerging factor pattern does not provide clear support for the subscale structure built into the measure. Some investigators have continued to treat the subscales as psychologically meaningful. Others have concluded that elementary school children simply do not make evaluative distinctions among domains. The former strategy is not justified on psychometric grounds. Nor is the latter conclusion warranted, since an inappropriate designation of domains and/or poor item selection may be responsible for the failure to demonstrate that children do indeed make differentiations. The scale construction efforts reported here were initially directed toward establishing an instrument which met the criteria for the factorial validity of the subscale structure and for acceptable subscale reliabilities. Furthermore, it was essential to demonstrate that the factor structure remained stable across the grade levels for which it was initially intended, third through sixth grades. Subsequently, we began to examine the convergent construct and discriminant validity of the instrument, as well as its appropriateness for a sample of junior high school pupils.

Finally, the typical response format for extant measures (e.g., true-false, like-me-unlike-me) also presents problems. Our exploratory work with Cooper-Smith's Self-Esteem Inventory reveals that this type of two-choice format pulls for socially desirable responding. Specifically, the self-esteem score from this instrument correlates positively and significantly with both the lie scale embedded in this measure and with scores on the Children's Social Desirability Scale (Crandall, Crandall, & Ko, 1965). Thus, altering the presentation format woul
to such a model, judgment overall self-worth are not a summation of responses, but rather, they are tapped directly into how one sees himself or herself as a person to assess these perceptions. Included in a fourth subscale of items on this subscale is the obvious of oneself, being happy, and thinking that one is a person. A secondary goal was to examine the relationship between pupils' perceived competence and their teachers' ratings of their actual competence. Thus a parallel teacher-rating scale was also constructed. This was an exploratory attempt to determine the extent to which the factor pattern and subscale intercorrelations for teacher ratings were comparable to the pupil ratings. Since teachers are most familiar with the child's scholastic performance, their ratings also represented a possible test of the convergent validity of the measure in the cognitive domain.

Method

Subjects

Earlier versions of the scale were individually administered to approximately 300 third- through sixth-grade children in Colorado. The factorial validity of the scale was then demonstrated with a sample of 133 9- to 12-year-old children from California, to whom the scale was group-administered. Replications were then conducted on the following samples: (a) a combined Connecticut-California sample of 341 third through sixth graders, (b) a New York sample of 714 third through sixth graders, (c) three separate Colorado samples in this same age range totaling 470 subjects, and (d) a California sample of 746 subjects from third through ninth grade. These samples have been drawn from primarily middle- and upper-middle-class populations. For every sample, there was approximately the same number of boys and girls at each grade level.

Experimenters

Scale administration in Colorado and California has been conducted by me and a team of research staff and graduate students, comprised of three men and six women. Data from Connecticut and New York were collected by colleagues and their research staff.

Question Format

A "structure alternative format" was designed to offset the tendency to give socially desirable responses. The child is presented with the type of question shown at the bottom of this page.

The child is first asked to decide which kind of kid he or she is most like—the kids described on the right or the left. Once having made this decision, the child decides whether the description on that side is sort of true or really true for him or her. Each item is scored from 1 to 4, where a score of 1 indicates low perceived competence and a score of 4 reflects high perceived competence. Scores are summed and then averaged for each subscale, resulting in four separate subscale means.

The effectiveness of this question format lies in the implication that half of the children in the world (or in one's reference group) view themselves in one way, whereas the other half view themselves in the opposite manner. That is, this type of question legitimizes either choice. The option of choosing either "sort of true for me" or "really true for me" also broadens the range of choices over the typical two-choice format. In addition, none of the choices involves the response "false." Rather, the child must decide which option is most true for him or her. Several sources of evidence bear on the effectiveness of this format. The correlation between perceived competence ratings and scores on the Children's Social Desirability Scale (Crandall et al. 1965) is .09, whereas the Coopersmith Self-Esteem Inventory correlates .33 with the Crandall scale. Inspection of individual protocols reveals that children utilize the entire range of scores, including the extreme designating the lowest perceived competence. The item standard deviations, presented in table 1, also reflect this variability. Furthermore, although the rating scale is ordi-
nal, both item and subscale score distributions are extremely normal, making parametric statistical analyses appropriate.

Item Construction and Scale Revision

From the outset, four domains—cognitive, social, physical, and general—were designated. Face validity and meaningfulness to children, as determined from individual interviews, served as initial guidelines. Some items were adapted from existing scales. The cognitive-competence domain included school as well as nonschool, social domains referred primarily to peers but also included items referring to “people,” and the physical domain included sports as well as skills at making and building things. After initial item revisions, based on feedback from individual children, a 40-item version, 10 items per subscale, was grouped and administered to a sample of 215 third through sixth graders. Factor analysis indicated that a four-factor solution was the most appropriate, in terms of both statistical criteria (Cattell’s scree test) and interpretability. However, only six to seven items on each subscale met all of the criteria: (a) moderate to high loadings on the designated factor, (b) no cross loadings of the same magnitude, (c) mean value near the midpoint, (d) sufficient variability (SD at or near 1), and (e) contribution to the internal consistency of the subscale.

The content of items not meeting these criteria included items that were good, worrying about thinking one’s ideas were good, worrying about what people think, knowing what to say to people, liking to build things, being good at fixing things, thinking one was good looking, and feeling that one would grow up to be important. These items seemed to be overly general and therefore vague, or they tapped skills that went beyond the content of the majority of items that defined a given factor. For example, they tapped nonschool cognitive competence, social interaction with adults, and physical skills involving crafts and fine-motor skills. Thus it was decided to define cognitive as school competence, social as peer related, and physical as skill at sports and outdoor games. Several new items were written, and an eight-item per subscale revision was then administered to a new sample of 133 children. Analyses revealed that seven items on each subscale met the criteria outlined above, resulting in the final 28-item scale.

In 14 of the items, the first part of the statement reflects high perceived competence, and in the remaining 14, low competence is described first. Within each subscale, three items are keyed in one direction, four in the other. With regard to item order, no two consecutive items are from the same subscale, and no more than two consecutive items are keyed in the same direction.

Teacher-Rating Scale

A 28-item teacher-rating form parallels the child scale. Items were reworded to obtain the teacher’s best judgment of the child’s actual competence. Thus a teacher item would read, “This kid often forgets what he or she learns, but this kid remembers things easily.” The same four-choice question format and scoring procedure was employed.

Results

Factorial Validity

Both orthogonal and oblique solutions were obtained, each revealing the same stable factor structure. The factor pattern1 from the oblique (Promax) rotation is presented. Although it was assumed that individuals would show differences across the four subscales, it was also anticipated that there would be some correlation among subscale scores. Thus an oblique solution, which allows the factors to intercorrelate, was considered to be the most appropriate. Cattell’s scree test, which employs criteria based on the magnitude of the eigenvalues, indicated that four factors should be extracted.

The factor pattern for the combined Connecticut-California sample of 341 pupils is presented in table 1. There it can be seen that items have moderate to high loadings on their designated factor and that with one exception in this sample they do not cross-load on other factors. This factor pattern has been found to be extremely stable and has been replicated in five additional samples. In any given sample, between one and four moderately low cross-loadings have been observed; however, particular items do not systematically cross-load.

To give an example of the replicability of the factor pattern, for the New York sample the average loadings of items on their designated factors were .67, .61, .64, and .50 for the cognitive, social, physical, and general subscales, respectively. (Three items, one physical and two general, cross-load on other factors with loadings of .32, .36, and .37. However, since items which cross-load vary from

---

1 For those more familiar with orthogonal rotations, these are available on request.
the remaining 14, low competence items first. Within each subscale, they are keyed in one direction, four for each subject. With regard to item order, no two or more items are from the same subscale, or else two consecutive items are from the same direction.

Rating Scale
28-item teacher-rating form parallel scale. Items were worded to obtain the teacher's best judgment of the child's competence. Thus a teacher item was given to the kid that he or she observed. This kid often forgets what he did, but this kid remembers things easily. A typical four-choice question format was employed.

Validity
The orthogonal and oblique solutions of each variable, each revealing the same factor structure. The factor pattern was obtained from the (Promax) rotation. It was assumed that individuals would score high across the four subscales. It is suggested that there would be a moderate amount of sub-scale scores. The solution, which allows the factors to correlate, was considered to be the best. Cattell's scores, which were based on the magnitude of the correlation, indicated that four factors should be included.

Factor Pattern
Factor pattern for the combined California sample of 341 pupils is given in table 1. There it can be seen that these factors moderate to high loadings on the factors. The scores were obtained across four subscales. For example, they do not cross-load on. This factor pattern has been extremely stable and has been replicated in additional samples. In any given sample, one and four moderately high scores have been obtained; however, items do not systematically cross-load. To give an example of the replicability of the factor pattern, for the New York sample, average loadings on items on their factors were .67, .61, .64, and .50, respectively, social, physical, and general, respectively. (Three items, one physical, two general, cross-loaded on other loadings of .32, .36, and .37. Hence, it which cross-load item availability is on request.

<table>
<thead>
<tr>
<th>Item Abbreviation</th>
<th>Cognitive</th>
<th>Social</th>
<th>Physical</th>
<th>General</th>
<th>Item</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive competence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good at schoolwork</td>
<td>.64</td>
<td>2.7</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Like school, doing well</td>
<td>.63</td>
<td>2.8</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Just as smart as others</td>
<td>.67</td>
<td>2.9</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can figure out answers</td>
<td>.45</td>
<td>2.8</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finish schoolwork quickly</td>
<td>.36</td>
<td>3.1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remember things easily</td>
<td>.42</td>
<td>2.9</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understand what read</td>
<td>.49</td>
<td>2.8</td>
<td>.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social competence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have a lot of friends</td>
<td>.61</td>
<td>3.0</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Popular with kids</td>
<td>.66</td>
<td>3.2</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy to like</td>
<td>.40</td>
<td>2.5</td>
<td>.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do things with kids</td>
<td>.40</td>
<td>2.9</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy to make friends</td>
<td>.42</td>
<td>(.43)</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important to classmates</td>
<td>.43</td>
<td>2.8</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most kids like me</td>
<td>.49</td>
<td>2.9</td>
<td>.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical competence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do well at all sports</td>
<td>.79</td>
<td>2.9</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Better at sports</td>
<td>.62</td>
<td>2.5</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do well at new activity</td>
<td>.45</td>
<td>2.8</td>
<td>.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good enough at sports</td>
<td>.75</td>
<td>2.7</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First chosen for games</td>
<td>.55</td>
<td>2.5</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Play rather than watch</td>
<td>.60</td>
<td>2.8</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good at new games</td>
<td>.46</td>
<td>2.5</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General self-worth:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sure of myself</td>
<td>.64</td>
<td>2.6</td>
<td>1.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Happy the way I am</td>
<td>.41</td>
<td>3.0</td>
<td>.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feel good way I act</td>
<td>.55</td>
<td>2.8</td>
<td>.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sure am doing right thing</td>
<td>.39</td>
<td>3.0</td>
<td>.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Am a good person</td>
<td>.69</td>
<td>3.0</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Want to stay the same</td>
<td>.48</td>
<td>2.9</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do things fine</td>
<td>.35</td>
<td>2.7</td>
<td>.95</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Loadings less than .37 are not presented for the sake of clarity.

2 Congruence coefficients (see Cattell 1966) to assess the correlation between factor loadings across samples were calculated between each pair of grades for the California sample, for each factor separately. For the fourth through ninth grades, these values ranged from .69 to .89. Third-grade values were slightly lower, from .81 to .83. Over all grade comparisons, the average congruence coefficients were .83 for cognitive, .74 for social, .80 for physical, and .76 for general.
sample of 810 pupils retested after 9 months. These correlations, corrected for attenuation, were .78, .80, .87, and .70 for the Colorado sample, and .78, .75, .80, and .69 for the New York sample, for the four subscales.

Subscale Means

Subscale means for four samples are presented in Table 2. These values are highly stable across subscale as well as grade, and this pattern is replicable, with slight fluctuations, across the four samples. Subscale standard deviations are also very consistent across grade and sample, ranging from .55 to .79. Grade × gender analyses of variance revealed no significant differences for the cognitive, social, and general subscales. However, on the physical subscale males consistently have shown significantly (p < .05) higher scores than have females (e.g., 3.0 vs. 2.5 for the sample in Table 1).

Intercorrelations among Subscales

Table 3 reveals that while there are sample variations, there are several common patterns. The correlations of the general self-worth subscale with each of the three competence subscales consistently tend to be among the highest, in the range between .40 and .58. The social and physical subscales are highly related across samples, within the range of .46 and .58. The relationship between the cognitive subscale and both the social and physical subscales tends to be lower, with some sample variation. Correlations among factors are of a similar magnitude, for example, general self-worth correlates with the three competence subscales .42–.53, social and physical correlate .56, and social and cognitive correlate .43.

Teacher Ratings

Teacher ratings were obtained from 28 teachers for the California sample and 16 teachers for the Colorado sample, which represented four teachers per grade. A factor pattern, virtually identical to the pupils’, resulted with the following average loadings on the designated factor: .84, .74, .83, and .86. One general item (this child is fine the way he [she] is) cross-loaded substantially on cognitive, and one social item (most kids like him [her]) also had a moderate loading on the general subscale.8 Internal consistency reliabilities for the teachers’ ratings were .96, .93, .94, and .93 for the cognitive, social, physical, and general subscales.

8 To examine the comparability in factor patterns, coefficients representing the congruence between teacher and pupil factor loadings were calculated for each factor. These values ranged from .90 to .97 for elementary school and .72 to .88 for the junior high.

The pattern of subscale relationships is slightly different from that of the pupils. Teachers view general self-worth as related to both cognitive (r = .40 and .54 for the two samples) and social (r = .63 and .50) competence, as do pupils. However, teachers consider general self-worth to be significantly (p < .05) less related to physical competence (r = .26 and .17) than do pupils. Teachers are consistent with pupils in viewing social and physical as highly related (.33 and .38 for teachers). However, teachers tend to see cognitive and social competence as more highly related (.43 and .45) than do pupils.

Convergent Validity

Cognitive domain.—Teachers have consistently said that they felt most confident about their judgments of cognitive competence. Across samples, the correlation between pupil ratings and teacher ratings is in the .40’s. For the California sample, extending into sixth grade, these correlations document a definite age trend. For third, fourth, fifth, and sixth grades, they were .28, .32, .50, and .55, in that order, showing a steady increase in magnitude. In the seventh grade the correlation drops to .31 and then rises to .66 in the eighth grade and to .73 in the ninth grade.

It is interesting that this same pattern is obtained when we examine the relationship between perceived cognitive competence and standardized achievement test scores from this same sample, taken from the Iowa Test of Basic Skills. Given the high correlation among the percentile scores for reading comprehension, language skills, and math skills, the mean of these three scores was correlated with the perceived cognitive competence score. For the third, fourth, fifth, and sixth grades, these values were .37, .40, .45, and .45, in that order. For the seventh grade this correlation dropped to .29 and then increased to .44 for the eighth grade and to .54 in the ninth grade.

Social domain.—In the domain of peer friendship, we have examined the relationship between the social subscale score and sociometric standing as defined by the Roster and Rating Scale developed by Rotaacher (1974). On this latter measure, each child is given a class roster and asked to rate each classmate in terms of how good a friend that person is, on a scale from 1 to 5. A child’s sociometric score is derived by averaging the same-sex
TABLE 2
MEAN SUBSCALE SCORES AND STANDARD DEVIATIONS BY GRADE FOR FOUR SAMPLES

<table>
<thead>
<tr>
<th>Competence and Sample</th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>Grade 6</th>
<th>Grade 7</th>
<th>Grade 8</th>
<th>Grade 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \bar{X} )</td>
<td>SD</td>
<td>( \bar{X} )</td>
<td>SD</td>
<td>( \bar{X} )</td>
<td>SD</td>
<td>( \bar{X} )</td>
</tr>
<tr>
<td>Cognitive:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>3.0</td>
<td>.58</td>
<td>2.9</td>
<td>.61</td>
<td>2.7</td>
<td>.70</td>
<td>2.7</td>
</tr>
<tr>
<td>2.</td>
<td>2.8</td>
<td>.59</td>
<td>3.0</td>
<td>.59</td>
<td>2.7</td>
<td>.53</td>
<td>2.8</td>
</tr>
<tr>
<td>3.</td>
<td>2.9</td>
<td>.64</td>
<td>2.7</td>
<td>.64</td>
<td>3.0</td>
<td>.55</td>
<td>2.9</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td>2.8</td>
<td>.56</td>
<td>2.9</td>
<td>.60</td>
<td>2.8</td>
</tr>
<tr>
<td>Social:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>2.9</td>
<td>.59</td>
<td>2.7</td>
<td>.66</td>
<td>3.0</td>
<td>.66</td>
<td>2.8</td>
</tr>
<tr>
<td>2.</td>
<td>2.9</td>
<td>.71</td>
<td>3.0</td>
<td>.72</td>
<td>2.7</td>
<td>.69</td>
<td>3.0</td>
</tr>
<tr>
<td>3.</td>
<td>3.0</td>
<td>.63</td>
<td>2.7</td>
<td>.69</td>
<td>3.0</td>
<td>.66</td>
<td>2.8</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td>2.9</td>
<td>.64</td>
<td>2.9</td>
<td>.65</td>
<td>2.8</td>
</tr>
<tr>
<td>Physical:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>2.8</td>
<td>.63</td>
<td>2.8</td>
<td>.63</td>
<td>2.8</td>
<td>.62</td>
<td>2.8</td>
</tr>
<tr>
<td>2.</td>
<td>2.8</td>
<td>.72</td>
<td>3.1</td>
<td>.61</td>
<td>2.7</td>
<td>.65</td>
<td>2.9</td>
</tr>
<tr>
<td>3.</td>
<td>2.8</td>
<td>.66</td>
<td>2.8</td>
<td>.71</td>
<td>2.9</td>
<td>.79</td>
<td>2.6</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td>2.7</td>
<td>.69</td>
<td>2.8</td>
<td>.69</td>
<td>2.7</td>
</tr>
<tr>
<td>General:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>3.0</td>
<td>.59</td>
<td>2.9</td>
<td>.60</td>
<td>2.8</td>
<td>.63</td>
<td>2.8</td>
</tr>
<tr>
<td>2.</td>
<td>3.1</td>
<td>.61</td>
<td>3.2</td>
<td>.47</td>
<td>2.8</td>
<td>.64</td>
<td>3.1</td>
</tr>
<tr>
<td>3.</td>
<td>3.0</td>
<td>.61</td>
<td>2.9</td>
<td>.55</td>
<td>3.0</td>
<td>.62</td>
<td>2.9</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td>2.9</td>
<td>.59</td>
<td>2.9</td>
<td>.61</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Note.—1 = California, \( N = 746 \); 2 = Colorado, \( N = 296 \); 3 = Connecticut and California, \( N = 341 \); 4 = New York, \( N = 714 \).
peers' ratings of the child. In a study with 85 fourth, fifth, and sixth graders, Thompson (Note 1) found the correlation between this sociometric index and the social subscale to be .59.

**Physical domain.**—Although ratings of physical competence from regular classroom teachers consistently correlate with pupil ratings between the narrow range of .41 and .43 across samples, these ratings have not been used as a primary index of validity since teachers report little confidence in these judgments. Thus, for one sample of 209 third through sixth graders in Colorado, the gym teacher rated each pupil on his or her physical competence. These ratings were made in February, after the teacher had at least 6 months to observe the children's skills. The correlations between the gym teacher's rating and pupil ratings was .62 and showed negligible variations across the four grades sampled.

**Construct Validity**

One purpose for devising this instrument was to test certain hypotheses in my model of competence motivation. One such hypothesis is that perceived competence should be positively related to one's intrinsic motivational orientation to prefer challenge, to be curious, and to engage in independent mastery attempts (Harter 1978). Clear support for this prediction comes from the correlations between perceived cognitive competence and the motivational subscales on our measure of intrinsic versus extrinsic orientation in the classroom (Harter 1981b). Perceived cognitive competence is strongly related to preference for challenge (r = .57) and to independent mastery (r = .54), and it is moderately related to curiosity (r = .33). Higher-order factoring reveals that perceived cognitive competence, challenge, independent mastery, and curiosity form a distinct factor with very high loadings of .76, .87, .80, and .79, respectively.

In a related behavioral study on preference for challenge (Harter, Note 2) sixth-grade children were divided into high (top third) and low (bottom third) perceived cognitive competence subgroups and given their choice of anagrams varying in difficulty level, that is, three-, four-, five-, and six-letter anagrams. Those with relatively high perceived competence scores chose significantly more (48% = .001, more difficult anagrams to solve (X number of letters = 5.1) than did those with low perceived competence scores (X = 3.8).

**Discriminant Validity**

**Cognitive domain.**—In one study (Barr, Note 3) with learning disabled children, support was found for the prediction that while such children should rate their competence lower than do normal children of the same age and grade, this difference will be most pronounced in the cognitive area. Results indicated a significant difference, t(38) = 2.9, p < .005, for the perceived cognitive competence ratings (X = 1.9 for the 20 learning-disabled children compared with 2.8 for the 20 normal children). While social, physical, and general self-worth ratings were also lower for the learning disabled, these differences did not reach acceptable levels of significance.

**Social and physical domain.**—For one sixth-grade sample in which athletic achievement was a prominent school value, the hypothesis was tested that those pupils selected for the sports teams (N = 53) would score higher on perceived physical and social competence than would their classmates (N = 57). Physical and social scores for the sports group were 3.4 and 3.2, compared with the means of the classmates, 2.5 and 2.7, respectively, t(78) = 3.4 and 2.5, p < .001 and p < .01, respectively.

**Discussion**

The findings indicate that the goal of constructing a self-report measure which provides a profile of the child's perceived competence and a general self-worth belief. The very structure reveals that the meaningful differences tapped by this instrument in school, social, physical competence, and self-worth. The g individual profiles, the subscale differences (e.g., the Cognitive domain) which yield meaningful information about their children. In addition, this instrument provides a fuller, more functional profile of the child's feelings of self-worth.

Although the perceivably designed by school children, our school pupils have a g factor pattern that is remarkably stable across situations. The items seem central throughout the day, doing well at school, in sports, at home, and in any situation. The items seem to be central throughout, doing well at school, in sports, at home, and in any situation. The items seem to be central throughout, doing well at school, in sports, at home, and in any situation. The items seem to be central throughout, doing well at school, in sports, at home, and in any situation.
Susan Harter

Our measurement in the perceived competence of the child is modest. Higher-ordered cognitive competence, for example, with a variable of .79, respectively. One study (Note 2) identified the high (top) 25 percent of children with a perceived competence index, t(48) = 4.68, p < .01. Some children did not score as high as the others (X = 3).

Although the perceived competence scale was originally designed for use with elementary school children, our experience with junior high school pupils has demonstrated that this scale can be employed at these grade levels. The factor pattern and subscale reliabilities are highly stable across the third through ninth grades. The items seem to tap concerns which are central throughout this broad age range, namely, doing well at school, having friends, competence in sports, and liking the kind of person one is. It is quite likely, however, that junior high school pupils make additional distinctions, for example, between same-gender and opposite-gender friendships. This would argue for the construction of a scale specifically designed to tap critical new domains which emerge during adolescence.

Caution should also be taken in the use of this scale with very different populations or special groups. For example, in a recent dissertation (Sill, Note 4) the factor structure of this scale for a group of 126 mentally retarded pupils between the ages of 9 and 12 was examined. The findings revealed a very interpretable two-factor solution. The first factor, which was labeled “competence,” drew primarily from items which were on the original cognitive and physical subscales, suggesting that educable retarded children do not make a distinction between these two skill domains. The second factor, which we labeled “popularity,” was defined by the more concrete social items making reference to the number of friends one had. Finally, there was no evidence for a general self-worth factor. This finding suggests that at this particular IQ and mental age level, a child does not make the type of abstract evaluation of self as is tapped by our general self-worth items. In an independent effort (Harter & Pike, Note 5) to construct a downward extension of perceived competence scale for 4-7-year-olds, a similar pattern emerged on this pictorial version in the form of a two-factor solution, competence and social acceptance. This finding raises the possibility that the perceived competence construct is qualitatively different at different developmental levels.

For normal samples, subscale intercorrelations among the three competence domains reveal that the largest correlation is between the social and physical subscales. This pattern suggests that one’s popularity during the elementary and junior high school years, as tapped by the social subscale, is dependent to a considerable degree on one’s prowess at athletics. One’s academic performance bears less of a relationship to one’s popularity or one’s physical skills. The teachers’ ratings are in agreement with the pupils’ in recognizing that one’s peer popularity is very directly related to the pupil’s skill in sports. We focus on “popularity” here since we have begun to question whether the social scale actually assesses competence in the sense of social skills or whether it simply taps a peer friendship dimension which may be more highly related to factors such as athletic prowess than social competence per se. We are currently researching this issue (see Thompson, Note 1).

Another issue being pursued is precisely how the general self-worth scale is interpreted by children. Minton (Note 6) has discovered that the reasons children give for their judgments on the self-worth items include references to personal characteristics such as being nice, friendly, helpful, “morally” good; to the control of anger; to one’s relationship with parents; and to expressed feelings of self-acceptance, being happy with the way one is. Children make few references to competence. In future work, we will need to address the issue of whether this subscale actually taps a superordinate sense of self-worth or whether it is merely tapping second-order self-evaluative dimensions other than competence (see Harter, in press).

While validity data have been reported for the competence subscales, the interpretation of concurrent measures such as teacher ratings and achievement test scores should be...
It has also been demonstrated that discrepancy scores of normal IQ pupils are predictive of behavior on cognitive tasks. Bierer (Note 8) found that two groups of children, those who overestimate and those who underestimate their cognitive competence, are more unrealistic than accurate raters in their difficulty level choices.

These findings may have implications for program-evaluation efforts, suggesting that a goal should be to foster a realistic sense of competence rather than enhancement per se. That is, reduction in the magnitude of discrepancy scores may be as critical an outcome variable as a mean increase in the competence score.

Finally, it should be noted that the present scale is restricted to cognitive competence in school, peer social relationships, and physical competence in sports. It does not tap cognitive competence outside of school, social relationships with adults, or the type of physical skills required to do crafts, build, and fix things, and so on. These are the focuses of a supplementary scale in progress, which may be useful in the evaluation of competencies outside the more traditional domains. This type of extension also guards against the temptation to treat the existing scale as synonymous with the perceived construct itself. Like any such construct, it becomes increasingly more differentiated, if one allows it the conceptual and empirical room to do so.

Reference Notes

so been demonstrated that a set of normal IQ pupils are able to evaluate their cognitive competencies, and that two groups of children who estimate and those who underestimate, tend to have more accurate raters in their decisions.

Ratings may have implications for action efforts, suggesting that to foster a realistic sense of than enhancement per se. The magnitude of discrepancy as critical an outcome variable in the competence score: should be noted that the person's estimate to cognitive competence is related to social relationships, and physical sports. It does not tap cognitive inside of school, social relations, or the type of physical skills, crafts, build and fix things, and are the focus of a supplement that may be useful at competencies outside the domain. This type of scale is against the temptation to a scale as synonymous with the trait itself. Like any such construct, more differentiated and conceptual and empirical note.

### References


White, R. Ego and reality in psychoanalytic theory. Psychological Issues, 1963, 3(9). (Monograph)


<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Some kids feel that they are very good at their school work</th>
<th>BUT</th>
<th>Other kids worry about whether they can do the school work assigned to them.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>Some kids find it hard to make friends</td>
<td>BUT</td>
<td>Other kids find it's pretty easy to make friends.</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>Some kids do very well at all kinds of sports</td>
<td>BUT</td>
<td>Other kids don't feel that they are very good when it comes to sports.</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>Some kids are happy with the way they look</td>
<td>BUT</td>
<td>Other kids are not happy with the way they look.</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Some kids often do not like the way they behave</td>
<td>BUT</td>
<td>Other kids usually like the way they behave.</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>Some kids are often unhappy with themselves</td>
<td>BUT</td>
<td>Other kids are pretty pleased with themselves.</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>Some kids feel like they are just as smart as as other kids their age</td>
<td>BUT</td>
<td>Other kids aren't so sure and wonder if they are as smart.</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>Some kids have a lot of friends</td>
<td>BUT</td>
<td>Other kids don't have very many friends.</td>
</tr>
<tr>
<td></td>
<td>Really True for me</td>
<td>Sort of True for me</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-------------------</td>
<td>---------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1 2</td>
<td>Some kids wish they could be a lot better at sports</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids feel they are good enough at sports.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>4 3</td>
<td>Some kids are happy with their height and weight</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids wish their height or weight were different.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>4 3</td>
<td>Some kids usually do the right thing</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids often don't do the right thing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1 2</td>
<td>Some kids don't like the way they are leading their life</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids do like the way they are leading their life.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1 2</td>
<td>Some kids are pretty slow in finishing their school work</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids can do their school work quickly.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1 2</td>
<td>Some kids would like to have a lot more friends</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids have as many friends as they want.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>4 3</td>
<td>Some kids think they could do well at just about any new sports activity they haven't tried before</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids are afraid they might not do well at sports they haven't ever tried.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1 2</td>
<td>Some kids wish their body was different</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids like their body the way it is.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4 3</td>
<td>Some kids usually act the way they know they are supposed to</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids often don't act the way they are supposed to.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>4 3</td>
<td>Some kids are happy with themselves as a person</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids are often not happy with themselves.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1 2</td>
<td>Some kids often forget what they learn</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids can remember things easily.</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4 3</td>
<td>Some kids are always doing things with a lot of kids</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BUT Other kids usually do things by themselves.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Really True for me</td>
<td>Sort of True for me</td>
<td>BUT</td>
<td>Other kids don’t feel they can play as well.</td>
<td>Really True for me</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
<td>-----</td>
<td>-----------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>21. 4 3</td>
<td>Some kids feel that they are better than others their age at sports</td>
<td></td>
<td>Other kids don’t feel they can play as well.</td>
<td>2 1</td>
</tr>
<tr>
<td>22. 1 2</td>
<td>Some kids wish their physical appearance (how they look) was different</td>
<td>BUT</td>
<td>Other kids like their physical appearance the way it is.</td>
<td>3 4</td>
</tr>
<tr>
<td>23. 1 2</td>
<td>Some kids usually get in trouble because of things they do</td>
<td>BUT</td>
<td>Other kids usually don’t do things that get them in trouble.</td>
<td>3 4</td>
</tr>
<tr>
<td>24. 4 3</td>
<td>Some kids like the kind of person they are</td>
<td>BUT</td>
<td>Other kids often wish they were someone else.</td>
<td>2 1</td>
</tr>
<tr>
<td>25. 4 3</td>
<td>Some kids do very well at their classwork</td>
<td>BUT</td>
<td>Other kids don’t do very well at their classwork.</td>
<td>2 1</td>
</tr>
<tr>
<td>26. 1 2</td>
<td>Some kids wish that more people their age liked them</td>
<td>BUT</td>
<td>Other kids feel that most people their age do like them.</td>
<td>3 4</td>
</tr>
<tr>
<td>27. 1 2</td>
<td>In games and sports some kids usually watch instead of play</td>
<td>BUT</td>
<td>Other kids usually play rather than just watch.</td>
<td>3 4</td>
</tr>
<tr>
<td>28. 1 2</td>
<td>Some kids wish something about their face or hair looked different</td>
<td>BUT</td>
<td>Other kids like their face and hair the way they are.</td>
<td>3 4</td>
</tr>
<tr>
<td>29. 1 2</td>
<td>Some kids do things they know they shouldn’t do</td>
<td>BUT</td>
<td>Other kids hardly ever do things they know they shouldn’t do.</td>
<td>3 4</td>
</tr>
<tr>
<td>30. 4 3</td>
<td>Some kids are very happy being the way they are</td>
<td>BUT</td>
<td>Other kids wish they were different.</td>
<td>2 1</td>
</tr>
<tr>
<td>31. 1 2</td>
<td>Some kids have trouble figuring out the answers in school</td>
<td>BUT</td>
<td>Other kids almost always can figure out the answers.</td>
<td>3 4</td>
</tr>
<tr>
<td>32. 4 3</td>
<td>Some kids are popular with others their age</td>
<td>BUT</td>
<td>Other kids are not very popular.</td>
<td>2 1</td>
</tr>
<tr>
<td></td>
<td>Really True for me</td>
<td>Sort of True for me</td>
<td></td>
<td>Sort of True for me</td>
</tr>
<tr>
<td>---</td>
<td>-------------------</td>
<td>---------------------</td>
<td>---</td>
<td>---------------------</td>
</tr>
<tr>
<td>33.</td>
<td>1 2</td>
<td>Some kids <em>don't</em> do well at new outdoor games</td>
<td>BUT</td>
<td>Other kids are <em>good</em> at new games right away.</td>
</tr>
<tr>
<td>34.</td>
<td>4 3</td>
<td>Some kids think that they are good looking</td>
<td>BUT</td>
<td>Other kids think that they are not very good looking.</td>
</tr>
<tr>
<td>35.</td>
<td>4 3</td>
<td>Some kids behave themselves very well</td>
<td>BUT</td>
<td>Other kids often find it hard to behave themselves.</td>
</tr>
<tr>
<td>36.</td>
<td>1 2</td>
<td>Some kids <em>are not very</em> happy with the way they do a lot of things</td>
<td>BUT</td>
<td>Other kids think the way they do things is <em>fine</em>.</td>
</tr>
</tbody>
</table>
The Self-Perception Profile for College Students

This is an adaptation of the previous measure, designed for use with college students. It measures college students' self-concept in 13 different areas, as opposed to the four areas on the children's version. This measure was selected because of its low cost (the authors encourage people administering this test to copy it straight from the manual reducing cost significantly), preliminary evidence of reliability and validity, and because of the specific population it was designed to reach. It is administered in a group setting, adding to the ease of its use. The authors suggest that information gathered from this measure could be useful in helping students determine areas of self-development that they would like to focus on.


This paper describes a study in which the Self-Perception Profile for College Students was administered. It describes the instrument briefly, as well as the results of their study.

Murphy, L.L., Conoley, J.C., & Impara, J.C. (Eds.). (1994). Tests in print IV: An index to tests, test reviews, and the literature on specific tests. Lincoln, Nebraska: The University of Nebraska Press.

Provides a brief description of the test, references for additional information, and a write up evaluating the strengths and weaknesses of this measure.
2.10, p < .05). Except for Heterosexuality, the direction and degree of these sex differences were similar to the differences between men and women of the American sample (Murgatroyd & Gavurin, 1975). However, American college men were significantly higher than American college women on Achievement and Aggression while no sex differences were found on these two variables for the Chinese sample. By the same token, American college women were significantly higher than American college men in Affiliation and Abasement while no sex differences were found on these two variables in the Chinese group.

Hwang (1976) argued that, as a result of the effects of Western culture upon the Chinese, the cross-cultural differences on the Edwards Schedule for the Chinese and American college students had decreased. However, significant cross-cultural differences (p < .01) were found here for 10 of the 15 variables for both sexes. Furthermore, sex differences in the two cultural samples were only partially consistent, indicating there were some interaction effects between sex and culture.

REFERENCES


Accepted December 10, 1990.

THE HARTER SELF-PERCEPTION PROFILE: SOME NORMATIVE AND PSYCHOMETRIC DATA

SONIA W. MASCUIUC
Mount Saint Vincent University

LISSA S. E. McRAE
Mount Saint Vincent University

JEFFREY D. YOUNG
Mount Saint Vincent University

Summary.—The Harter Self-perception Profile was administered to 102 undergraduate students in business courses at two Canadian universities. Men's and women's ratings differed over-all for global self-worth and the “What I am like” domains of Creativity, Athletic Competence, and Appearance. Men and women did not differ significantly on the importance ratings.

In response to revived interest in mental structures, specifically self-concept, self-concept scales for children (Harter, 1982, 1985), adolescents (Harter, 1988, in press), college students (Neemann & Harter, 1986), and adults (Messer & Harter, 1986) have been developed. These scales utilize a developmental life-span perspective wherein self-concept is viewed as an over-all mental structure composed of domains which vary in number and type throughout the life span. For example, during childhood there are five domains of self-perception and at the college level there are 12 domains. In addition, a separate measure, global self-worth, is obtained by direct questions pertaining to the individual's global perception of self as a person.

College students represent a major group of interest for whose assessment Neemann and Harter (1986) published the Manual for Self-perception Profile for College Students. On this an individual rates 34 items on a scale of 1 to 4 to describe “What I am like” for 12 domains (creativity, intellectual ability, scholastic competence, job competence, athletic competence, appearance, romantic relationships, social acceptance, close friendships, parent relationships, humor, and morality) and global self-worth. In addition, 24 items are rated on a scale of 1 to 4 in terms of the importance of the above 12 domains to that individual. The individual's self-esteem in each domain is the difference between the rating on the “What I am like” scale (i.e., the competence score) and the importance ratings. This difference between ratings is termed the discrepancy score however; discrepancy scores are only considered if the domain has an importance rating of 4. For example, indi-
individuals who give themselves a low job-competence rating of 2 and an importance rating of 4 for job competence will have low self-esteem for job competence. The manual provides normative and psychometric data for 300 students at two American universities, which data support the reliability and validity of the instrument (Neeman & Harter, 1986).

Using three-way analysis of variance with sex, social class, and college attended as independent variables, Neeman and Harter (1986) reported that college women rated themselves significantly higher on the Close Friendship domain than did college men. Conversely, men rated themselves significantly higher on the Appearance and Athletic Competence domains than did women.

The present study was undertaken to assess whether Canadian men and women business college students differed from the college sample reported by Neeman and Harter (1986), differed on global self-worth, differed on any of the 12 domains proposed by Neeman and Harter (1986) or differed on the importance attributed to the 12 domains proposed by Neeman and Harter (1986). This would provide normative and psychometric data for a sample of Canadian undergraduate business students as well as further evaluate sex differences and self-concept.

METHOD

The Self-perception Profile for College Students (Neeman & Harter, 1986) and a demographic survey were administered to 45 men and 30 women from a university in Quebec, and 11 men and 16 women from a university in Atlantic Canada. These respondents were 56 (55%) men, ranging in age from 19 to 30 yr, and 46 (45%) were women, ranging in age from 18 to 29 yr. In terms of mother tongue, 79 (77%) spoke English, 20 (20%) spoke French, and 3 (3%) spoke other languages. In terms of program of study, 96 (94%) were enrolled in business studies and 6 (6%) were enrolled in other programs.

RESULTS AND DISCUSSION

The two university groups were combined because results of a multivariate analysis of variance indicated that the groups did not significantly differ on the What I am like scales ($F_{1.124} = 1.8, p > .05$) or on the importance scales ($F_{1.124} = 1.38, p > .05$).

Table 1 presents by sex the means and standard deviations for the What I am like scales, the importance scales, and the discrepancies for the present sample. A multivariate analysis of variance suggested that men's and women's ratings differed over-all for global self-worth and the What I am like domains ($F_{1.124} = 2.22, p < .02$). Specifically, significant differences were found on the global self-worth rating ($F_{1.124} = 6.31, p < .02$), Creativity ($F_{1.124} = 5.39, p < .03$), Athletic Competence ($F_{1.124} = 13.02, p < .001$), and

<table>
<thead>
<tr>
<th>Domain</th>
<th>What I Am Like</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women</td>
<td>Men</td>
</tr>
<tr>
<td>Self-worth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.2</td>
<td>3.4</td>
</tr>
<tr>
<td>SD</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Creativity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.8</td>
<td>3.1</td>
</tr>
<tr>
<td>SD</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>Intellectual Ability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.3</td>
<td>3.4</td>
</tr>
<tr>
<td>SD</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Scholastic Competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.9</td>
<td>2.9</td>
</tr>
<tr>
<td>SD</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Job Competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.3</td>
<td>3.4</td>
</tr>
<tr>
<td>SD</td>
<td>0.6</td>
<td>0.5</td>
</tr>
<tr>
<td>Athletic Competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.6</td>
<td>3.2</td>
</tr>
<tr>
<td>SD</td>
<td>1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>Appearance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.3</td>
<td>3.0</td>
</tr>
<tr>
<td>SD</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Romantic Relationships</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.7</td>
<td>2.8</td>
</tr>
<tr>
<td>SD</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Social Acceptance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.1</td>
<td>3.3</td>
</tr>
<tr>
<td>SD</td>
<td>0.8</td>
<td>0.6</td>
</tr>
<tr>
<td>Close Friends</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.3</td>
<td>3.4</td>
</tr>
<tr>
<td>SD</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Parent Relationships</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.4</td>
<td>3.5</td>
</tr>
<tr>
<td>SD</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>Humor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.2</td>
<td>3.4</td>
</tr>
<tr>
<td>SD</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Morality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.5</td>
<td>3.3</td>
</tr>
<tr>
<td>SD</td>
<td>0.5</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Appearance domains ($F_{1.124} = 13.2, p < .001$). Men and women did not differ significantly on the importance ratings ($F_{1.124} = 1.69, p > .05$).

Table 2 presents measures of internal consistency by sex and for the total sample.

The reliability of the instrument as assessed by coefficient alpha, a measure of internal consistency, was deemed adequate. The coefficients for the importance scales are somewhat lower than those for the What I am like scales, and this might be expected given that the importance scales contain only two items while the What I am like scales consist of four items.

The current sample is similar to Neeman and Harter's (1986) in that women in both samples rated themselves significantly lower than the men on Athletic Competence and Appearance domains. In the Canadian sample, women also rated the Creativity and Global Self-worth domains significantly lower than men. In Neeman and Harter's (1986) sample, women rated the
Close Friendship domain significantly higher than men; however, the Canadian sample showed no significant sex difference on the Close Friendship domain. Also, Canadian men and women who were business students did not differ significantly on the importance attributed to Harter’s 12 self-concept domains. In the Harter sample women rated intellectual ability, scholastic competence, and close friendships as more important than did men.

### TABLE 2

<table>
<thead>
<tr>
<th>Coefficients Alpha For “What I Am Like” and Importance Scales</th>
<th>What I Am Like</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Self-worth</td>
<td>.86</td>
<td></td>
</tr>
<tr>
<td>Job Competence</td>
<td>.74</td>
<td>.42</td>
</tr>
<tr>
<td>Scholastic Competence</td>
<td>.79</td>
<td>.71</td>
</tr>
<tr>
<td>Social Acceptance</td>
<td>.84</td>
<td>.36</td>
</tr>
<tr>
<td>Appearance</td>
<td>.87</td>
<td>.66</td>
</tr>
<tr>
<td>Parent Relationships</td>
<td>.85</td>
<td>.67</td>
</tr>
<tr>
<td>Close Friends</td>
<td>.85</td>
<td>.63</td>
</tr>
<tr>
<td>Intellectual Ability</td>
<td>.81</td>
<td>.68</td>
</tr>
<tr>
<td>Morality</td>
<td>.79</td>
<td>.61</td>
</tr>
<tr>
<td>Romantic Relationships</td>
<td>.86</td>
<td>.73</td>
</tr>
<tr>
<td>Humor</td>
<td>.85</td>
<td>.70</td>
</tr>
<tr>
<td>Creativity</td>
<td>.90</td>
<td>.54</td>
</tr>
<tr>
<td>Athletic Competence</td>
<td>.95</td>
<td>.87</td>
</tr>
</tbody>
</table>

The differences may be cultural, associated with university programs or requirements or students in business programs versus other programs. Regardless of the differences and accompanying reasons for them, the fact remains that in the Canadian sample, the women rated themselves significantly lower on Creativity, Athletic Competence, and Appearance domains even though they rated these domains to be as important as did the men. In addition, the women business students rated their Global Self-worth significantly lower than did the men.

Women business students who place less over-all value on themselves and who see themselves as being less inventive or creative, less athletic, and less attractive than their male peers may place themselves at a distinct disadvantage. It remains to be determined whether ratings on Global Self-worth and Creativity affect or influence academic success (Bandura, 1978; Kanfer, 1980; Harter, 1989). Further research must assess whether the way in which individuals rate themselves on Creativity, Athletic Competence, and Appearance affects academic and job performance. If self-worth and self-esteem affect the way individuals function academically and socially within a college setting, such findings would have implications for women and men business students as well as for the faculty who teach them.

### REFERENCES


*Accepted December 10, 1990.*
Self-Perception Profile for College Students

Purpose: To measure college students' self-concept.
Population: College students.
Publication Date: 1986.


Administration: Group.

Price Data, 1989: $9 per manual (84 pages).
Time: 30 (40) minutes.

Authors: Jennifer Neemann and Susan Harter.
Publisher: University of Denver.

TEST REFERENCES
Self-Perception Profile for College Students

Distinguished Professor of Educational Psychology, University of Nebraska-Lincoln, Lincoln, NE:

Three inventories are included in this package of self-perception instruments: “What I Am Like,” “Importance Ratings,” and a "Social Support" scale. The instruments can be used separately or as part of a package. The "What I Am Like" inventory is designed to assess college students' self-rating of their global self-worth and in 12 specific domains: Creativity, Intellectual Ability, Scholastic Competence, Job Competence, Athletic Competence, Appearance, Romantic Relationships, Social Acceptance, Close Friendships, Parent Relationships, Humor, Morality, and a Global Self-Worth Rating. Each of the content domain subscales has four items and the Global Self-Worth subscale has six items. Estimated completion time is 30 minutes. The "Importance Ratings" inventory parallels the Self-Worth inventory with two items for each of the 12 content domain scales. Finally, the "Social Support" scale has four items each for subscales labelled: Close Friends, Mother, Father, People in Campus Organizations, and Instructors. The scales and scoring information are provided in the manual and the authors inform readers that they should feel free to copy the instruments directly from the manual.

The "What I Am Like" inventory items ask students to select which one of two self-descriptions (e.g., "Some students are not satisfied with their social skill," and "Other students think their social skills are just fine") are Really true for me or Somewhat true for me. The self-descriptions are on the same lines with the word "But" between them. There is no middle response of "undecided" or "neither." Responses are recorded either on the left margin or on the right margin, depending upon which statement the student believes describes them. Items are scored on a 1 to 4 scale where 4 represents the most competent or adequate self-judgment. Half the items have the high competency phrasing on the left and the other half have the high competency phrasing on the right. Response formats for the "Importance Ratings" and "Social Support" scales are similar and the items are also similar in general content.

The "What I Am Like" inventory parallels a similar instrument designed by Hartter and associates for use with children, adolescents, and adults. The college student form was developed to fill the developmental gap between the adolescent and adult populations.

The current form of the instruments was developed using data obtained from 300 students, mostly single and Caucasian. Internal consistency assessed by coefficient alphas for the "What I Am Like" inventory were reasonably high for the Self-Worth scales with only one scale below .80. Factor analysis with an oblique rotation and using a scree test resulted in 12 scales perfectly matching the designed scales. (The Global Self-Worth scale was not included.) Correlations among the scales range from low (e.g., Athletic to Scholastic Competence, $r = .02$), to modest (Social Acceptance to Morality, $r = .24$), to moderate (Intellectual Ability to Scholastic Competence, $r = .65$). Reliabilities were less strong for the "Importance Ratings" scales ranging from .53 to .84 but reasonably high for the "Social Support" scale with the range from .76 to .90.

A variety of preliminary validity data are reported which provide interesting as well as supportive information. Gender differences are noted with males having higher self-worth ratings on Appearance and Athletic Competence, but lower importance ratings on Intellectual Ability, Scholastic Competence, and Close Friendships. Unexpectedly, the authors found no gender differences on the social support scales. Overall, Appearance, Social Acceptance, Job Competence, Scholastic Competence, and Intellectual Ability correlated most strongly with the Global Self-Worth scale on the "What I Am Like" inventory and the same general pattern held for the "Importance Ratings".

The authors suggest ways the instruments can be used for research purposes such as examining the relationship of resilience to stress and self-esteem and for individual therapy. In the latter use, they provide examples of how profiles of the self-worth and importance ratings can provide useful insights as well as how to use discrepancy scores. Such information can be useful for helping students determine what aspects of their self they wish to develop.

Personal measures designed specifically for use with a college population are too rare. Too many instruments were developed using clinical populations or focus extensively on problems. One of the most prominent instruments available for use with a normal college population for developmental purposes is the Student Developmental Task and Lifestyle Inventory (1984), which has proven useful for research purposes and for working with individual students. Researchers, counselors, and trained mentors will find it appropriate to consider use of these self-perception scales as well. Indeed, a worthwhile research topic would be to explore the relationship among the scales on these instruments and how receptive they are to intervention strategies.

If the self-perception inventories are deficient, beyond the fact that only preliminary validity data are available, it is because of what is omitted rather than what is present. It is unfortunate that the designers did not go further beyond the domains they already had covered in the adolescent and adult version than they did. Several domains immediately come to mind such as leisure-recreational skills, self-management (e.g., stress management) skills, etc.
Self-Perception Profile for College Students

It is noteworthy that the students ($N = 300$; 70 males, 230 females) employed for the development of this instrument were predominantly freshmen (47.33%) from the University of Denver and Colorado State University. Very few juniors (13.66%) and seniors (7.66%) were represented in the sample. Additionally, a majority (84%) were commuters. The mean age of this predominantly Caucasian (93%), predominantly single (94%) sample was 19.8 years. Such demographics suggest potential areas of concern for the representativeness of the sample. As “the target population of this measure is the traditional full-time college student, ages 17–23,” the potential usefulness of the instrument may be limited somewhat.

The reliabilities of the four-item subscales, as measured by coefficient alpha, ranged from .76 to .92 with only one domain (Job Competence) falling below .80. Suggesting the absence of ceiling and floor effects, subscale means clustered around 3.00 ($sd = \text{approximately } .80$). Three reliable gender effects were reported. Females scored higher on Close Friendship, whereas males scored higher on Athletic Competence and Appearance. The results of Cattell’s scree test “indicated that twelve factors should be extracted, and these corresponded perfectly to the intended twelve subscales.” A subsequent principle components factor analysis yielded high factor loadings (average = .78, range = .52–.92) with no cross-loadings.

In an attempt to place this scale within the theoretical model of William James who proposed that global self-esteem resulted from an “evaluation of the ratio of one’s successes to one’s pretensions,” an Importance Scale also was developed. By determining the discrepancy between the competence score for a domain and the importance score of that domain, the authors suggest that they have operationalized James’ theory. The Importance Scale consists of 24 questions (2 from each domain).

The items for the Importance Rating Scale are balanced and scored in the same manner as those for the Self-Perception Profile. The internal reliabilities (coefficient alpha) range from .53 to .84 with only three subscales falling below .72. For two-item subscales such figures are quite acceptable. Other descriptive measures show that the importance means are rather high (range = 2.68–3.75) with one-half of the measures falling above 3.50. The standard deviations fall within an acceptable range of .45 to .83.

The most sensitive use of the importance scores appears to be in conjunction with the global measure of self-worth. When importance and competency scores are plotted for each domain for high, medium, and low self-worth subjects, it is clear that although the importance ratings did not vary, the competence scores decreased dramatically from high...
to low self-worth. Although this is an excellent and informative use of these scores, the rationale behind the formation of the three self-worth categories is not explained or substantiated. They were formed “for demonstration purposes only.” Hence, other configurations and results are possible.

Finally, recognizing that self-worth may be related to the process of socialization, the authors have included a Social Support Scale evaluating the importance of Close Friends, Mother, Campus Organizations, Father, and Instructors. This scale, which has a total of 20 items (4 per subscale), is administered and scored in the same manner as the Self-Perception Profile. The subscale reliabilities (coefficient alpha) range from .76 to .90. Although the subscale means are rather high (3.10–3.64) there is sufficient variability (adj range = .53–.70).

Reflecting some lack of organization in preparing the manual, it is not until page 45 that one finds information concerning validity and the target population for the Self-Perception Profile. (Such information is not given for the Importance Ratings and Social Support scales.) As these sections are rather short, they should have been incorporated directly into an earlier section. As validity information is presented for only three subscales (Social Acceptance, Close Friendships, and Parent Relationships), it is clear that additional work is needed in this area. The fact, however, that all three reported measures are quite reliable (p = .001) is encouraging. Similarly, the authors occasionally refer to appendices that may have been planned but which are not to be found in the completed manual.

The descriptive portion of the manual concludes with a section on applications. The authors suggest the Self-Perception Profile can be used in research, as a predictor of the resilience to stress, and as a therapeutic adjunct. The remainder of the manual is devoted to complete versions of each of the three scales, scoring keys, and an Individual Profile Form.

As the user is given permission to reproduce the scales and the profile form, all of the necessary elements for administering, scoring, and interpreting this scale are contained in this manual. Based upon the information provided here, the Self-Perception Profile and its two related scales should prove to be quite informative and helpful to those interested in evaluation of the self-concept of college-age subjects.

Sentence Comprehension Test, Revised Edition

Purpose: “To assess children’s comprehension of English and/or Panjabi.”
Acronym: SCT.
Administration: Individual.
Price Data, 1988: £30.55 per complete set including 25 record forms (English version), 25 record forms (Panjabi bilingual version), picture book (‘87, 45 pages), and manual (‘87, 43 pages); £5.20 per 25 record forms (English version); £6.90 per 25 record forms (Panjabi bilingual version); £16.70 per picture book; £12.65 per manual.
Time: (10–20) minutes.
Authors: Kevin Wheldall, Peter Mittler, Angela Hobsbaum, Dorothy Gibbs (Panjabi bilingual version), Deirdre Duncan (Panjabi bilingual version), and Surinder Saund (Panjabi bilingual version).
Publisher: NFER-Nelson Publishing Co., Ltd. [England].
Cross References: For reviews by Francis X. Archambault and Mavis Donahue of the experimental edition, see 9:1107.

TEST REFERENCES

Review of the Sentence Comprehension Test, Revised Edition by GABRIEL DELLA-PIANA, Professor of Educational Psychology, University of Utah, Salt Lake City, UT:
The Sentence Comprehension Test (SCT), Revised Edition, has the same purpose and structure as the earlier experimental edition. The test is designed “to measure a child’s ability to comprehend sentences of varying length and grammatical complexity by requiring him to select appropriately from sets of four pictures” the one representing an orally presented stimulus sentence. The major changes introduced in the revised version are reduced testing time (due to fewer subtests) and redrawing of all pictures (for clarity, elimination of sex stereotyping, and including multicultural representation of persons illustrated). In addition, new studies of reliability and validity and some normative data support the revised edition.

The intended use of this measure of comprehension of orally presented sentences is primarily diagnostic, “[to] help . . . pinpoint the sentence constructions” that are causing the child some difficulty and to “lead to more systematic and specific approaches to remediation.” It is also intended as a “quick screening device” to identify children whose performance deviates markedly from other children. Thus, the key questions to be addressed are: adequacy of reliability, validity, and norms for these purposes; ease of administration and scoring by testers; and usefulness of the manual to support the intended interpretations and uses of test results.

The administration procedure is simple and the record form is easy to use. An example illustrates the
The Unconditional Self-Regard Scale

This instrument was designed to measure unconditional self-regard. This is defined by the authors of the test as "self-liking or self-acceptance is not contingent on one's own performances in various behavioral domains (e.g. athletic, academic, or social), although the individual may strive to improve her or his levels of performance in one or more of those areas" (Betz, et al., 1995, pg. 77). It was selected based on the authors' suggestion that this may be a useful tool for evaluating counseling outcomes. The authors demonstrate adequate evidence of reliability and validity, although they do acknowledge that further support is needed. This test is short, easy to administer and score, and has the potential to provide valuable outcome information in counseling settings.


This paper is the test authors' write up of the description, validity, and reliability information on the test. They also honestly discuss limitations of the instrument.
Evaluation of a Measure of Self-Esteem Based on the Concept of Unconditional Self-Regard

Nancy E. Betz, Elaine Wohlgemuth, Deborah Serling, John Harshbarger, and Karla Klein

A measure of unconditional self-regard (USR) derived from Rogerian (client-centered) theory was evaluated in four studies. Results indicated that USR was highly reliable and was strongly related to other measures of self-esteem and negatively related to depression, anxiety, and symptoms of psychological distress. The measure was significantly related to the control and commitment dimensions of psychological hardness but not to its challenge dimensions. Evidence for discriminant validity was provided by findings that USR was unrelated to social desirability or private self-consciousness. Finally, female students' self-regard was significantly related to the unconditional positive regard they reported receiving from their mothers. Other correlations between self-regard and perceived regard from others, although nonsignificant, were often large in absolute value. Implications for counseling interventions are discussed.

Few of those who are engaged or interested in counseling would deny the importance of the client's self-concept or self-esteem in either the initial occurrence or the successful treatment of client problems (e.g., Bednar, Wells, & Peterson, 1989; Coopersmith, 1967; Nurius, 1986). Several major counseling theories, among them Rogers's (1951) client-centered therapy and May's (1983) approach to existential therapy, rely on self-esteem as a major explanatory variable.

Despite the widely recognized importance of this variable to mental health practitioners, however, its use in both counseling and research has been limited somewhat by the lack of theory-based measures of self-esteem, that is, measures whose development have been based on theories of either the etiology or the treatment of problems of low self-esteem in particular and psychological problems in general. Although there are many extant measures of self-esteem, most were developed by personality theorists rather than by counseling theorists.

In terms of a general definition of self-esteem, Blascovich and Tomaka (1991) summarized available theory and research by noting that self-esteem, as opposed to self-concept, is an overall affective (as opposed to cognitive) evaluation of one's own worth and that it is also referred to as self-regard or self-acceptance (Blascovich & Tomaka, 1991). Self-concept, on the other hand, is a broader and more inclusive term and refers to more specific cognitive or behavioral self-perceptions and beliefs. Self-efficacy expectations, the belief that one can successfully complete a given behavior (Bandura, 1977, 1986), are an example of a component of the self-concept.

Although a general definition of self-esteem can be provided, most major reviewers of the area, beginning with those of Wylie (1961, 1974, 1979) and continuing with those of Blascovich and Tomaka (1991), Coopersmith (1967), Demo (1985), and Nurius (1986), have concluded that research on self-esteem has been plagued by problems in measurement. For example, Blascovich and Tomaka (1991), and Demo (1985) cited such problems as failure to validate measures against similar measures, failure to examine discriminant validity (i.e., correlations of a measure with conceptually dissimilar variables), and, most important, failure to provide a theoretical basis for measurement and validation of measures.

One theoretical framework that could be more highly developed in terms of both conceptual and applied relevance to self-esteem is that of Rogerian theory of healthy functioning and of client-centered therapy. Rogers's (1957, 1961) discussions of the healthy personality and, thus, the goals of counseling in effect defined self-esteem (which he called self-acceptance or unconditional self-regard) as a congruence between one's ideal and real selves. A discrepancy between real and ideal selves was thought to be the basis for psychological distress and a central target of counseling interventions. Furthermore, the means of helping the client move toward unconditional self-regard included the therapist's provision to the client of unconditional positive regard.

Despite the theoretical centrality of unconditional self-regard, there were few attempts to measure it directly. Some client-centered research focused on self-ideal discrepancy as an indicator of self-acceptance (e.g., Butler & Haigh, 1954; Dymond, 1954; Seeman, 1954). This research was based on Stephenson's Q-sort technology, which consisted of self-referent statements written on cards sorted into "like me," "not like me," and "not sure" piles. The discrepancy between self (What are you like today?) and ideal self (What would you like to be like?) sorts constituted the dependent variable (e.g., Butler & Haigh, 1954).

Psychometrically, difference (discrepancy) scores are less reliable than are direct measures of the construct (Nunnally, 1978). Consequently, we suggest that a measure of self-esteem based directly on the Rogerian (Rogers, 1957) concept of unconditional self-regard would address current concerns in the literature about the lack of theoretical bases for self-esteem measures. Furthermore, it could be useful in evaluating the results of counseling, because it would be an indicator of what Rogers (1961) considered the sine qua non of effective therapy, that is, unconditional self-regard in the client. Finally, it would have considerable utility as an outcome variable in counseling research because of the conceptual closeness between recommended therapist behavior (unconditional positive regard for the client) and a desired therapeutic outcome (unconditional self-regard by the client).

Accordingly, we designed the present series of studies to develop and evaluate a measure of unconditional self-regard based on Rogerian theory. Following accepted standards for the construction of rational or construct-based measures (Dawis, 1987; Nunnally, 1978; Spector, 1992), we based scale construction on a comprehensive definition of the construct of interest. The first two studies were designed to ex-
Measure of Self-Esteem Based on Unconditional Self-Regard

METHOD

Participants

Participants were students enrolled in introductory psychology classes at a large midwestern university over two successive academic quarters. Students were volunteers in the sense that they had a wide variety of psychological experiments from which to choose; students received course credit for their participation. Students were recruited by using a sign-up sheet that described the study and that provided the time and place where the study would be conducted. Of the 157 students who signed up for Study 1, 152 (97%) appeared for the study; of the 112 who signed up for Study 2, 107 (96%) appeared. All of the students appearing for the study agreed to participate. The mean age of Study 1 students was 19.1 years (SD = 2.4). The large majority (93%) were of European American origin, with 5% African American and 1% each of Hispanic and Asian American ethnicity. The mean age of Study 2 participants was 18.9 years (SD = 2.0); 90%, 6%, 3%, and 1% were of European American, African American, Asian American, and Hispanic ethnicity, respectively.

Instruments

Unconditional Self-Regard Scale (USRS). To develop a measure of unconditional self-regard, we defined the construct using the writings of Rogers. As discussed especially in Rogers’s early work (1951, 1961), one of the objectices of client-centered therapy is client self-acceptance (Rogers, 1951) or unconditional self-regard (Rogers, 1957). The salient features of these are that the client perceives himself or herself as a person of worth and places standards of evaluation within self. He or she achieves a more realistic appraisal of self, his or her relationships, and his or her environment (Rogers, 1951, pp. 138–139).

The other salient feature is noncontingent valuing and acceptance of oneself. In other words, self-liking or self-acceptance is not contingent on one’s own performances in various behavioral domains (e.g., athletic, academic, or social), although the individual may strive to improve her or his levels of performance in one or more of those areas.

In accordance with rational or construct-based approaches to scale construction (e.g., see Davis, 1987; Nunnally, 1978), the definition was used as a guide for Nancy E. Betz and Deborah Serling, two of the authors who are subject-matter experts in that both are counseling psychologists who use client-centered theory as a primary basis of their counseling and therapy. Each wrote up to 25 items based on this definition and included both positively and negatively worded items among those they wrote. Accordingly, 50 items reflecting this definition were written. The 50-item initial pool was reduced to 20 items by obtaining two independent ratings of how closely each item represented the definition given. Those items obtaining an average rating of 3.5 or above on a 5-point scale of relevance to the construct as defined were retained. In addition, because the original item writers had produced several instances of nearly identical items, only one of the pair was retained, either by selecting the one that read most clearly or by using a random numbers table to select one.

The 20 items selected included 9 positively and 11 negatively worded items. An example of a positively worded item is “Even though I make mistakes, I feel good about myself as a person,” and an example of a negatively worded item is “I can never quite measure up to my own standards.” Responses to the items were obtained on a 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5). Scores were computed by, first, reverse-scoring all negatively

STUDIES 1 AND 2: SCALE CONSTRUCTION, RELIABILITY, AND VALIDITY

The purposes of the first two studies were (a) to develop a measure of unconditional self-regard based on the Rogersian concept of noncontingent self-evaluation; (b) to evaluate the internal consistency reliability and factor structure of the measure; and (c) to collect evidence regarding its convergent and construct validity.
worded items and, second, adding the scores for all of the self-esteem items. Scores on the scale have a theoretical range of 20 to 100, with higher scores associated with higher levels of self-esteem.

Convergent Validity Measures

Rosenberg Self-Esteem Scale. The Rosenberg Self-Esteem Scale (Rosenberg, 1965) was designed to measure global feelings of self-worth. It consists of 10 items to which the examinee responds using a 4-point Likert scale (ranging from 1 for strongly disagree to 4 for strongly agree). Five of the items are positively worded (e.g., “I take a positive attitude toward myself” and “I feel that I have a number of good qualities”), and 5 are negatively worded (e.g., “I feel I do not have much to be proud of”). Scoring can be done either with Guttman scaling or simple additive scoring of item responses. Additive scoring was used herein. Silber and Tippett (1965) reported a 2-week test–retest reliability of .85. Internal consistency reliabilities (coefficient alphas) in Samples 1 and 2 were .89 and .87, respectively. Evidence for convergent validity is suggested by moderately strong correlations (r = .56 to .83) with numerous similar measures (see Blascovich & Tomaka, 1991).

Coopersmith Self-Esteem Inventory. The Coopersmith Self-Esteem Inventory (Coopersmith, 1967), using 50 items, was originally designed to measure self-regard in four areas: peers, parents, school, and personal interests. Coopersmith (1975) revised the original form by selecting only the 25 items with highest item–total correlations and claimed that it was then a unidimensional measure of self-regard. Factor analyses have not, however, supported the hypothesis of unidimensionality (Blascovich & Tomaka, 1991). The Coopersmith Self-Esteem Inventory consists of 25 items to which the examinee responds “like me” or “unlike me.” Items include “I have a low opinion of myself,” “I’m popular with persons my own age,” and “I find it very hard to talk in front of a group.” Items are scored like me (1) and unlike me (2) and are reverse-scored if necessary such that higher scores indicate higher levels of self-esteem. The Coopersmith Self-Esteem Inventory has been used extensively in research, summarized in the manual (Coopersmith, 1975). Internal consistency reliability in a sample of 226 college students was .81; in the present sample, an alpha of .86 was obtained.

Construct Validity Measures

State–Trait Anxiety Inventory. Trait anxiety was measured using the State–Trait Anxiety Inventory, Form Y (Spielberger, Gorsuch, Lushene, & Jacobs, 1977). This inventory comprises 40 items, 20 of which measure state anxiety and 20 of which measure trait anxiety. Trait anxiety items ask respondents to describe how they generally feel and include “I feel nervous and restless” and “I make decisions easily.” Students rate their reactions to each item on a 4-point Likert scale. Total scores on the scale range from 80 (high trait anxiety) to 20 (low trait anxiety).

Test–retest correlations for the trait anxiety scale range from .73 to .86 for undergraduate college students (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), and the median alpha coefficient (Cronbach, 1951) for trait anxiety is .90 (Spielberger et al., 1983). Values of coefficient alpha were .88 and .90 in Samples 1 and 2, respectively. Support for criterion-related and construct validity for the trait anxiety scale is provided in the Manual for the State–Trait Inventory (Spielberger et al., 1983) and includes high correlations between the trait anxiety scale, the Institute for Personality and Ability Testing Anxiety Scale (Cattell & Scheier, 1963), and the Taylor Manifest Anxiety Scale (Taylor, 1953).

Beck Depression Inventory. The revised Beck Depression Inventory (BDI; Beck, Rush, Shaw, & Emery, 1979) is a 21-item scale designed to assess the level of depression in adolescents and adults. The 21 items assess depressive symptoms and attitudes, including depressed mood, pessimism, sense of failure, self-dissatisfaction, guilt, punishment, and self-dislike. Each item has four response alternatives ranging from not depressed to severely depressed. The BDI is scored by summing the ratings of the examinee across the 21 items. Extensive information regarding the reliability and validity of the BDI are reported in Beck and Steer (1984). Reliability in Sample 2 was .82.

Mental Health. The 20-item version of Goldberg’s (1972) General Mental Health Questionnaire was used to measure the tendency toward psychiatric disturbance. Goldberg described this measure as one that forms a lowest common multiple of symptoms that can be encountered in the various differentiated syndromes of mental disorders. The item content focuses on the here and now rather than on how the individual has felt or behaved in the past and includes both physical (e.g., “I feel too tired and exhausted even to eat”) and psychologically oriented (e.g., “I feel unhappy and depressed”) items.

Hardiness. One personality variable receiving attention as a stress buffer is that of psychological hardness, as introduced in the late 1970s by Suzanne Kobasa and her colleagues at the University of Chicago (Kobasa, 1979, 1982; Kobasa, Maddi, & Courting, 1981). Kobasa defined hardness using concepts from existential psychology, that is, the three “constituent traits” of commitment, challenge, and control. Commitment refers to “a tendency to involve oneself in (rather than experience alienation from) whatever one is doing or encounters” (Kobasa, Maddi, & Kahn, 1982, p. 169). Challenge refers to “the belief that change rather than stability is normal in life and the anticipation of changes as interesting incentives to growth rather than threats to security” (Kobasa et al., 1982, pp. 169–170). Control refers to “a tendency to feel and act as if one is influential (rather than helpless) in the face of the varied contingencies of life” (Kobasa et al., 1982, p. 169).

Hardiness has been measured using several variations of the Personal Views Survey (Kobasa, 1985), which has ranged from 20 to 90 items. In this study, we used the 50-item version most recently used in research by Kobasa and her colleagues. The 50-item Personal Views Survey assesses challenge, commitment, and control using 17, 16, and 16 items, respectively. Examples of items include, “Most days, life just isn’t very exciting for me” (challenge); “I want to be sure someone will take care of me when I get old” (challenge); “I really look forward to my work” (commitment); and “Planning ahead can help avoid most future problems” (control). Subscale reliabilities in this sample were .57, .74, and .69, and the total scale reliability was .84.

Procedure

As mentioned, two samples of students participated in the study, in two successive academic quarters. The first sample consisted of 152 students, 78 male and 74 female. The second sample consisted of 107 students, 56 male and 51 female. In Study 1, participants received the following instruments in this order: (a) the Unconditional Self-Regard Scale; (b) Rosenberg Self-Esteem Scale; (c) Coopersmith Self-Esteem Inventory; (d) State–Trait Anxiety Inventory; and (e) Mental Health Questionnaire. In Study 2, participants were administered (a) the Unconditional Self-Regard Scale; (b) Rosenberg Self-Esteem Scale; (c) State–Trait Anxiety Inventory; (d) Beck Depression Inventory, (e) Psychological Hardiness Scale, and (f) Mental Health Questionnaire. Because of a desire not to burden the par-
Measure of Self-Esteem Based on Unconditional Self-Regard

RESULTS

First, the measure of unconditional self-regard was found to possess adequate internal consistency reliability, with alphas of .87 and .90 in Samples 1 and 2, respectively. Values of coefficient alpha usually differ somewhat across samples (Nunnally, 1978), and the difference found herein is within the range of normal sample variation. A principal-components factor analysis with oblique rotation indicated that the USRS was constituted by a large general factor accounting for 36% of the total variance on which 16 of the 20 items loaded significantly. The remaining 4 items loaded on a second factor (constituting 6% of the total variance), which reflected a tendency toward unrealistic/perfectionistic expectations regarding oneself (“I know that I am often too hard on myself”).

Table 1 provides the correlations of unconditional self-regard with other measures of self-esteem, providing evidence for concurrent validity, and with various indices of psychological adjustment, providing evidence for construct validity. Because of the large number of correlation coefficients computed both within and across the four studies described, we set the critical value of statistical significance for rejection of the null hypothesis that the correlation is zero in the population at p < .01. This procedure is one approach to controlling Type I error (e.g., Glass & Hopkins, 1984). As shown in Table 1, scores on the USRS are strongly correlated (p < .001) with other measures of global self-esteem. Correlations with the Rosenberg Self-Esteem Inventory were .77 and .78 in Samples 1 and 2, respectively.

Relationships with other indices of adjustment were also statistically significant (p < .001). Greater levels of unconditional self-regard were significantly related to lower levels of anxiety and depression, with fewer psychological symptoms, and with higher levels of two of the three components of the “hardy personality.” Of the hardness components, unconditional self-regard was most strongly related to control (r = .52) and somewhat less strongly related to commitment (r = .35, p < .001). With .01 as the critical value, the relationship of unconditional self-regard to challenge was not significant.

STUDY 3: CONSTRUCT AND DISCRIMINANT VALIDITY

This study was designed to evaluate discriminant validity by relating unconditional self-regard to a measure of social desirability and to a measure of private self-consciousness. Construct validity was examined by including measures of public self-consciousness and social anxiety.

METHOD

Participants

Participants were 96 undergraduates enrolled in introductory psychology classes at a large midwestern university. They were volunteers in the sense that they had a wide variety of psychological experiments from which to choose and they received course credit for their participation. Of 97 students originally signing up for the study, only 2 neglected to appear for the testing. All of the students appearing participated in the study. The students’ mean age was 19.2 years (SD = 2.4). The majority (94%) were of European American descent, with 4% African American and 2% Asian American.

Measures

Unconditional self-regard. This was measured using the 20-item USRS, as described previously. Reliability in this sample was .89.

Discriminant Validity Measures

Self-Consciousness Scale. The Self-Consciousness Scale developed by Fenigstein et al. (1975) was designed to measure two aspects of self-consciousness—public self-consciousness and private self-consciousness—and social anxiety. Public self-consciousness was intended to reflect a general awareness of the self as a social object being perceived as having an effect on others, and the 7 items used to measure the variable have a strong evaluative component, for example, “I’m self-conscious about the way I look.” Private self-consciousness (10 items) relates to attentiveness to self, for example, “I reflect about myself a lot.” Social anxiety is assessed by 6 items such as “I get embarrassed very easily” and “I feel anxious when I speak in front of a group.” Although these items may have some relationship to social self-esteem, they primarily refer to an attentional focus regarding the self.

In this study, a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree) was used. Fenigstein et al. (1975) reported test-retest correlations of .84 for public self-consciousness, .73 for private self-consciousness, .79 for social anxiety, and .80 for the total score. In this study, coefficient alphas were as follows: public self-consciousness (.76), private self-consciousness (.52), social anxiety (.71), and the total score (.73).

Marlowe–Crowne Social Desirability Scale—Short Form. The revised 20-item version of the Marlowe–Crowne Social Desirability Scale (MC-20; Crowne & Marlowe, 1960) was developed by Strahan and Gerbasi (1972). The original scale consisted of 33 true/false items, had an internal consistency of .88, and had a test-retest reliability of .89 (Crowne & Marlowe, 1960). In an unpublished study of the MC-20, Strahan (1986) reported that the short form had an alpha coefficient of .79 and was related at the level of r = .81 to the original.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Sample 1 r (n = 152)</th>
<th>Sample 2 r (n = 107)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Convergent validity correlations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenberg Self-Esteem Inventory</td>
<td>.77*</td>
<td>.78*</td>
</tr>
<tr>
<td>Coopersmith Self-Esteem Scale</td>
<td>.64*</td>
<td></td>
</tr>
<tr>
<td><strong>Construct validity correlations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State–Trait Anxiety Inventory (A-Trait)</td>
<td>−.70*</td>
<td>−.75*</td>
</tr>
<tr>
<td>Mental Health Symptoms</td>
<td>.67*</td>
<td>.69*</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>−.65*</td>
<td></td>
</tr>
<tr>
<td>Psychological Hardiness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Challenge</td>
<td>.22</td>
<td></td>
</tr>
<tr>
<td>Commitment</td>
<td>.35*</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>.52</td>
<td></td>
</tr>
</tbody>
</table>

*p < .001.
Marlowe–Crowne Social Desirability Scale. Strahan (1986) suggested that it provided an accurate measure of social desirability and could be used when test administration time was a consideration. In this study, the value of alpha was .67.

RESULTS

The correlation between unconditional self-regard and the scores on the Marlowe–Crowne Social Desirability Scale was 0.0, reflecting no tendency for socially desirable responding as measured in this way to influence reports of self-regard. The correlations of unconditional self-regard with the Self-Consciousness Scale were smaller than those found for the mental health indices discussed previously. The correlations of -.22 and -.21 with private self-consciousness and public self-consciousness were not statistically significant using p < .01 as the critical value. The correlation with social anxiety (-.36) was statistically significant and indicated that higher self-regard is associated with less social anxiety.

STUDY 4: CONSTRUCT VALIDITY: RELATIONSHIPS OF SELF-REGARD TO PERCEIVED REGARD FROM OTHERS

Once a measure has been found to possess adequate levels of reliability and validity, examination of the extent to which hypotheses about the constructs are supported by research using the measure adds further support not only for construct validity but also to theoretical understanding of the construct (Betz & Weiss, 1987; Cronbach & Meehl, 1955). One of the major hypotheses concerning unconditional self-regard is that it develops in response to perceptions of unconditional self-regard from significant others (Rogers, 1961). This study was designed to examine this hypothesis.

METHOD

Participants

Participants were 164 undergraduate students, 83 female and 81 male, enrolled in an introductory psychology course during the spring of 1990. Students participated in this research as partial fulfillment of course requirements, but participation in this particular study was voluntary. Of 169 students who signed up, 164 reported as scheduled. The 164 students had a mean age of 18.8 years (SD = 1.6), with a range from 17 to 26 years. Students were primarily of European American ethnicity (89%), with 6% African American, 3% Asian American, and 2% Hispanic.

Instruments

The USRS was used to measure self-regard. As mentioned previously, the USRS contains 20 items assessing unconditional self-regard. The previous three studies provided evidence for internal consistency reliability and concurrent, construct, and discriminant validity.

Bartlett-Lennard Relationship Inventory. The Bartlett-Lennard Relationship Inventory (BLRI; Bartlett-Lennard, 1962) was developed by Rogers' colleagues to measure client perceptions of the provision of the Rogerian core conditions (empathy, genuineness, and regard) by the therapist. The scale was used extensively in early client-centered research. In the BLRI, separate scales were used to assess level of regard and unconditionality of regard. In general, the instrument can be used to assess the interpersonal perceptions of either or both participants in a dyadic relationship.

The two 16-item scales measuring level and unconditionality of regard were used in this study; perceptions of empathy and genuine-

<table>
<thead>
<tr>
<th>Important Person</th>
<th>Frequency of Mentions*</th>
<th>r With Level of Regard</th>
<th>r With Unconditionality of Regard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
</tr>
<tr>
<td>Overall</td>
<td>81</td>
<td>83</td>
<td>.24</td>
</tr>
<tr>
<td>Mother</td>
<td>70</td>
<td>73</td>
<td>.20</td>
</tr>
<tr>
<td>(86%)</td>
<td>(88%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>69</td>
<td>68</td>
<td>.23</td>
</tr>
<tr>
<td>(85%)</td>
<td>(82%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brother</td>
<td>17</td>
<td>15</td>
<td>.16</td>
</tr>
<tr>
<td>(21%)</td>
<td>(18%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grandmother</td>
<td>13</td>
<td>19</td>
<td>.02</td>
</tr>
<tr>
<td>(16%)</td>
<td>(23%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher</td>
<td>13</td>
<td>15</td>
<td>-.41</td>
</tr>
<tr>
<td>(16%)</td>
<td>(18%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sister</td>
<td>9</td>
<td>15</td>
<td>.55</td>
</tr>
<tr>
<td>(11%)</td>
<td>(18%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grandfather</td>
<td>14</td>
<td>9</td>
<td>.08</td>
</tr>
<tr>
<td>(17%)</td>
<td>(11%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coach</td>
<td>15</td>
<td>3</td>
<td>.58</td>
</tr>
<tr>
<td>(19%)</td>
<td>(4%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. n = 81 for male students; n = 83 for female students. *Because respondents were allowed to list up to three "significant others" from their childhood and adolescence, percentages add up to more than 100. **p < .01, ***p < .001.
ness were not measured. Responses to the items (e.g., "She/he has been truly interested in me," assessing level of regard, and "Sometimes I’ve been more worthwhile in her/his eyes than I have been at other times," assessing unconditionality) were obtained on a 6-point Likert scale ranging from not at all true (1) to very true (6). Scoring was cumulative. Gurman (1977) reported that the average test-retest reliabilities across 10 studies reviewed were .85 for level of regard and .80 for unconditionality of regard. Cramer (1987) reported internal consistency reliability coefficients of .91 and .83 for level and unconditionality, respectively.

Before completing the BLRI, the student was asked the following statement: "Please select who you consider to be the three most important older people in your life as you were growing up." Options included parents, grandparents, aunts and uncles, older siblings, teacher, coach, and other. Then, the student was asked to complete the 32-item BLRI using each of those three people in turn as the "target" person of the inventory.

RESULTS

Table 2 shows the frequency with which the top eight categories of people were mentioned as among the "most important three people" for male and female students separately. It may be noted that parents are by far the most frequently mentioned, followed by other relatives, a teacher, and, among the male students, a coach. Although not shown because of low overall frequency, uncles were more important to male students (7 to 1) and aunts to female students (6 to 0). An overall chi-square analysis did not indicate significant gender differences in frequency of occurrence, probably because both genders predominately chose both parents.

Correlations of the USRS with the level and conditionalility of regard from significant others also are shown in Table 2, again only for those categories of others mentioned with some frequency. Again, the critical level of p was set at .01 to control for Type I error. This, in combination with very small sample sizes for important person categories other than parents, resulted in few statistically significant relationships. Among female students, overall perceived regard and mother’s level and conditionalility of regard were significantly related to daughter’s unconditional self-regard. Correlations of .25 with father’s level of regard, .51 with that of teachers, .68 and .69 with level and unconditionalility of grandfather’s regard, and .69 and .85 with that of coaches were not significant below p < .01. Among male students, no correlations were significant at p < .01, although those between the student’s self-regard and that of brothers (.51 for unconditionalility of regard) and coaches (.58 and .45 for level and unconditionalility) were relatively high in absolute magnitude.

GENERAL DISCUSSION

The results of these studies supported the convergent, discriminant, and construct validity of a measure of self-regard linked directly to the Rogersian concepts of unconditional positive regard and self-valuing.

First, values of internal consistency reliability ranged from .87 to .90 across samples; for a 20-item measure, these values indicate high reliability (Nunnally, 1978). The factor structure also supported the existence of a large general factor, with a small second factor suggesting unrealistic or perfectionistic expectations of the self.

Regarding convergent validity, the measure of unconditional self-regard was highly correlated with two other measures of self-esteem. Evidence for construct validity was provided by relationships to other adjustment indices as postulated; unconditional self-regard was associated with lower levels of anxiety, depression, and fewer symp-

joms and with higher levels of the control and commitment components of psychological hardiness, as conceptualized by Kobasa (1979). Its correlations with adjustment indices, ranging from absolute values of .65 to .75, were as high as those with other self-esteem measures. A correlation of -.36 with the Social Anxiety subscale of the Self-Consciousness Scale also supported construct validity, although a nonsignificant correlation with public self-consciousness was contrary to prediction.

Correlations with psychological hardiness varied according to the subscale in question. Unconditional self-regard was most strongly related (r = .52) to the control subscale, which Kobasa defined as the tendency to act as if one is influential (rather than helpless) in the face of the varied contingencies of life. The correlation with commitment (a tendency to be involved in life) was statistically significant but moderate (.35) in magnitude. With a critical value of .01 to minimize Type I error, the correlation of .22 between self-regard and challenge (seeking rather than fearing change) was not statistically significant.

Evidence for discriminant validity also was provided. First, as predicted, there was no significant relationship between unconditional self-regard and private self-consciousness. Second, evidence for discriminant validity was provided by the lack of relationship between self-regard and a measure of social desirability.

Some additional evidence for construct validity came from the results concerning the relationship between others’ positive regard and one’s own self-regard. Using the Barrett-Lennard (1962) Relationship Inventory to assess the perceived regard from significant others, results indicated significant relationships between the college women’s unconditional self-regard and both the level and unconditionalility of regard perceived from significant others in general and from mothers in particular. Unfortunately, breaking down the sample of 164 by gender and analyzing by category of most important significant other (small sample sizes for all categories but parents) resulted in some sample sizes so small that even large correlations (ranging from .45 to .85) failed to meet a .01 critical value of the p value associated with the correlation.

LIMITATIONS

Although the instrument examined herein seems to have promise for both research and applied purposes, several cautions should be kept in mind. First, the studies described herein represent the first efforts to evaluate the new scale. Further research in other samples, not only from other colleges and universities but also from other age groups, and in samples more varied in ethnicity should be undertaken to determine the generalizability of these findings.

Second, although many of the correlations between one’s own self-regard and that perceived from significant others while growing up were large in magnitude, small sample sizes due to breakdown by both gender and category of significant other (e.g., brother, grand-mother) limited findings of statistical significance. Further research focused on obtaining larger sample sizes would help to clarify the nature of these relationships.

It is suggested that, following additional research to support its psychometric properties, the USRS may be a useful instrument to include as a dependent variable in studies of counseling outcome. Because many theoretical orientations posit an increase in self-esteem as either a primary or secondary goal of counseling, a relatively short measure should be useful even if the counselor is not basing his or her work on client-centered theory. More important, interventions based at least in part on Rogersian or humanistic principles related to noncontingent self-valuing could be evaluated readily and in a theoretically consistent manner.
Furthermore, because most counseling training includes some emphasis on the Rogerian core counseling skills (e.g., see Hammond, Hepworth, & Smith, 1977), it is possible that, regardless of other therapeutic interventions, the unconditional regard of the therapist for the client leads to greater unconditional self-regard on the part of the client. Although this proposition received some attention in early client-centered research (e.g., Rogers & Dymond, 1954), research attention using analogous concepts, that is, unconditional positive regard from the therapist and unconditional self-regard in the client, is possible using the USRS.

The present patterns of intercorrelations also supported conclusions from clinical and personality research (Bednar et al., 1989) that there is a cluster of closely related variables connected to psychological health, including the concept of unconditional self-regard. Because Rogers’s initial theorizing posited self-ideal congruence as the key focus of change in counseling, it is possible that changes in unconditional self-regard would lead to reductions in such aversive conditions as anxiety, depression, and other symptoms of distress. Unconditional self-regard also may be useful to examine as a stress buffer, possibly functioning similarly to such stress buffers as social support and psychological hardiness. Intervention studies focusing explicitly on increasing the unconditional self-regard of the client as well as other studies using designs that would permit causal inferences also are needed.

In summary, we hope that the measure and results of this study will stimulate continued or renewed interest in a concept fundamental to much psychological distress and, conversely, to improvement through therapy.

REFERENCES


Measure of Self-Esteem Based on Unconditional Self-Regard


---

Nancy E. Betz is a professor in the Department of Psychology at Ohio State University. Elaine Wohlgemuth is a professor in the Department of Psychology at the University of Oregon. Deborah Seling is a psychologist in independent practice in Columbus, Ohio. John Harshbarger and Karla Klein are doctoral students in the Department of Psychology at Ohio State University. Correspondence regarding this article should be sent to Nancy E. Betz, Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, OH 43210.
**USR - SHORT FORM**

**ATTITUDE QUESTIONNAIRE**

**DIRECTIONS:** The 20 questions below deal with the attitudes of college students towards themselves and others. Please read each statement carefully. Then decide how strongly you agree or disagree with each statement. Mark your answers on this answer sheet, in the row of circles corresponding to each item number. Your response number indicates how closely each statement describes you and your feelings at the present time. There are no right or wrong answers. Please use the following key and DO NOT make up any of your own numbers:

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Aren't Sure or Neutral</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

1. I feel good about myself as a person.  
2. I make time for relaxation activities.  
3. I like who I am.  
4. It is hard for me to remember the positive things people say about me.  
5. I am very critical of myself.  
6. I think I am a worthwhile person.  
7. I argue a lot with my parents.  
8. I enjoy spending time with my friends.  
9. Even though I make mistakes, I still feel good about myself as a person.  
10. I think of myself in negative terms (e.g., stupid, lazy).  
11. It is easy for me to list 5 things I like about myself.  
12. I like to spend the holidays with my family.  
13. I can never quite measure up to my own standards.  
15. I like to be involved with team sports.  
16. Even when I goof up, I basically like myself.  
17. There are times when I doubt my worth as a person.  
18. I tend to look at what I do badly rather than what I do well.  
20. When I look in the mirror I like who I see.

F = Filler (not scored)  
R = Reverse keyed
Rosenberg Self-Esteem Scale

This measure is widely used and immensely popular. It is short, easy to score and administer, and widely accepted as reliable and valid. It was designed to measure global feelings of self-worth. It was selected based on its reputation and widespread support. It could be useful to agencies interested in their clients’ general feeling of self-worth. It may be helpful in a pre/post-test setting to evaluate the results of certain interventions.

# HOW I FEEL

For each of the following statements, indicate if you strongly agree, somewhat agree, somewhat disagree, or strongly disagree.

<table>
<thead>
<tr>
<th></th>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Somewhat Agree</th>
<th>Somewhat Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>There is really no way I can solve some of the problems I have.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2.</td>
<td>Sometimes I feel that I'm being pushed around in life.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3.</td>
<td>I have little control over the things that happen to me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4.</td>
<td>I can do just about anything I really set my mind to.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5.</td>
<td>I often feel helpless in dealing with the problems of life.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6.</td>
<td>What happens to me in the future depends mostly on me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7.</td>
<td>There is little I can do to change many of the important things in my life.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8.</td>
<td>I feel that I am a person of worth, at least on an equal plane with others</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9.</td>
<td>I feel that I have a number of good qualities.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10.</td>
<td>All in all, I am inclined to feel that I am a failure.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11.</td>
<td>I am able to do things as well as most other people.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12.</td>
<td>I feel I do not have much to be proud of.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13.</td>
<td>I take a positive attitude toward myself.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14.</td>
<td>On the whole, I am satisfied with myself.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15.</td>
<td>I wish I could have more respect for myself</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16.</td>
<td>I certainly feel useless at times.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17.</td>
<td>At times I think I am no good at all.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Appendix A

Annotated Bibliography

If you ever decide that you would like to look up a measure yourself, there are many handbooks and other resources that can be of help to you. Some resources to look through include:

- **The Handbook of Psychological Assessment**, by Gary Broth-Marnat, published by J. Wiley, New York, 1997. This can be found in the Psychological Sciences Library at Purdue University in the Reserve Section. Simply ask a librarian to get it for you and have your driver’s license or Purdue Visitor Identification Card (application available at any campus library) ready to give to him/her while you look through the book. This book provides background information, definitions of constructs (attitude or behavior you would like to measure), history, pros, cons, reliability, validity, and meaning information for the assessment interview, behavioral assessment, Wechsler Intelligence Scales, Bender Visual Motor Gestalt Test, Minnesota Multiphasic Personality Inventory, California Psychological Inventory, The Rorschach, Millon Clinical Multiaxial Inventory, Thematic Apperception Test, Screening and Assessment for Neuropsychological Impairment, Psychological Assessment Treatment Plan, and projective drawings. Additional readings are also recommended. The index of this book is very helpful in choosing which tests or measures measure which constructs.

  Call No. 150.287 G915h 1997 (also available on shelves for different years)


  Call No. 155.28 M623 1997 Psych. Library

- **Handbook of Multicultural Assessment: Clinical, Psychological, and Educational Applications**, published by Jossey-Bass, San Francisco, CA, 1996. This includes subject headings such as social aspects of psychological tests and is helpful in using tests and measures in culturally appropriate ways.

  Call No. 153.9308693 H191 1996 Psych. Library

- **Handbook of Assessment Methods for Eating Behaviors and Weight Related Problems: Measures, Theory, and Research**, published by Sage Publications, Thousand Oaks, CA, 1995. Contents include information on instruments to use to assess general personality and psychopathology among people with eating and weight related concerns, quality of life, subjective well-being, satisfaction with life, attitudes and beliefs about obese people, body image, eating behavior and style, and eating disordered thoughts, feelings, and behaviors. Some chapters are extremely clinical and focused on physiological aspects of eating behaviors and weight, but the psychologically focused chapters may be helpful.

  Call No. 616.8526075 H191 1995 Psych. Library

- **Buros Desk Reference: Psychological Assessment in the Schools**, published by the Buros Institute of Mental Measurements, University of Nebraska-Lincoln, distributed by the University of Nebraska Press, Lincoln, Nebraska, 1994. Provides information including title, purpose, target population, publication dates, acronym information, scores, administration, price data, time, comments, and a description of tests measuring general achievement, specific achievement, adaptive behaviors, behavior ratings, intelligence, perceptual functioning, personality, preschool
Appendix A

functioning, vocational interests, dropout prediction and prevention, and targeting educational risk.

Call No. 155.4180287 B937 1994 c.2 Psych. Library Reference Desk

- **Assessment of Children and Youth**, by Libby Cohen. Provides information on psychological tests for children, child development testing, youth psychological testing, behavioral assessment of children, behavioral assessment of teenagers, and educational tests and measurements.

Temporary Control No. AEX 7330 Consumer and Family Sciences

- **Tests in Print IV**, edited by L.L. Murphy, J.C. Conoley, and J.C. Impara, published by The University of Nebraska Press. A compilation of information about existing instruments to measure a broad range of traits. Includes brief descriptions of each test, price information, publishing information, and other articles to look at to get more information. This is in the reference section of HSSE Library.

- **The Eleventh Mental Measurements Yearbook**, edited by J.J. Kramer and J.C. Conoley, published by The University of Nebraska Press. Provides the same type of information as the Tests in Print, but tends to include instruments that the Tests in Print does not. This is in the reference section of HSSE Library.
Appendix A (cont.)

The following databases can be used to look up all kinds of pertinent information that agency personnel may be interested in. They are designed to do searches by subject, keyword, author, title, or journal title. For example, if someone were interested in obtaining information about how to measure self-esteem, he or she could go to one of these databases and enter “self-esteem” as their keyword and the system would perform a search. The computer would provide that individual with articles and books that contain information about self-esteem.

THOR

- THOR is the Purdue University cataloging program. You can access it at any of the three libraries that will be most helpful to you: HSSE (pronounced “Hissey”), Psychological Sciences (known as the “Psych Library”), and the Consumer and Family Sciences Library.

- THOR can be used to look up journal articles, books, and to check the Purdue catalog. To check the catalog, simply press enter from the initial screen and follow the directions on the screen. To look up journal articles or books, use the arrow key to scroll to the appropriate description and press enter. Follow the directions on the screen from there.

PSYCLIT

- PsychLit is a database of psychology related topics. It allows you to search both journal articles and books and chapters. This database is only available in the Psychological Sciences Library and Consumer and Family Sciences Library.

- Pick an unoccupied computer with the label “PsychLit” on it. The screen should be displaying the database selection screen. If it is not, ask a librarian for help. If it is, then use the arrow keys to scroll to the option you want: Journal articles 1991-12/97; Chapters and Books 1/98-1990; Journal articles 1974-1990. You probably want the first option, since this holds the most recent journal publications. When you have highlighted the option you want, press the space bar. This will put a check by the highlighted option, signifying that this is the one you want to search. If the check is in the right spot, hit enter. If not, repeat the above steps again.

- Now you should have a screen with instructions and the word “FIND” at the bottom. Type in a key word or phrase and press enter. The computer will search the database to find a match.

- Once PsychLit has found a match, it will let you know by telling you how many records it has found. Follow the instructions on the bottom of the screen to continue. If no match is found, you will see a message stating “0 records found.” Type in another key word or phrase and repeat the process until you find a match. Next to each terminal is a thesaurus to help you think of different words or phrases to use.

- The staff of the Psych library is willing to help if you need it. There are also instruction manuals and handouts located by each computer.

- The journals are arranged in alphabetical order, so once you have found some records, they are easy to look up in the stacks. Just look for the journal title, the volume, and number and go to the shelf to retrieve it. Copy machines are located in the library and copies can be made for $0.10.