

## Review of Various Instruments Used with an Adolescent Population

Michael J. Lambert

*Population.* This analysis will focus on a population of adolescent youth between the ages of 11 and 20 years old. This population is appropriate for school counselors working with adolescents in a middle or high school setting. Assessment instruments are a vital tool for school counselors. Psychological issues often initially arise during adolescence when youth experience significant physical and mental transitions. It is important that school counselors are aware of this vulnerability intervene when there is a possibility of mental and emotional struggles for students. Large school populations are a major issue however; counselors are spread thin and don't have much time to spend with each student. Assessments can be important tools, used to better understand the student's needs in an efficient manner. The assessments that this analysis will focus on are the Reynolds Adolescent Depression Scale – 2 (Reynolds, 2002), the Suicidal Ideation Questionnaire (Reynolds, 1988), and the Adolescent Anger Rating Scale (Burney, 2001).

### Reynolds Adolescent Depression Scale – 2<sup>nd</sup> Edition

*Purpose of Test.* The Reynolds Adolescent Depression Scale – 2<sup>nd</sup> Edition (RADS-2) was developed as a brief self-report instrument used to screen for depressive symptoms in adolescents. Rather than aiding as an instrument for the diagnosis of depression, the RADS-2 is designed investigate the severity of adolescent depressive symptoms. This instrument produces feedback on general depressive symptoms as well as on four subscales of depression: dysphoric mood (DM), negative affect (NA), negative self-evaluation (NS) and somatic symptoms (SS) brought on by depressive moods. Using the cutoff score of T-61, the RADS-2 is useful in identifying adolescents at a substantial risk for major depression.

*Use with Population.* The school-based standardization sample for the RADS-2 consisted of a total of 3,300 adolescents ranging in age from 11 to 20 years of age. The focus of this assessment tool on the 11 to 20 year old age group makes it appropriate for working with junior high and high school students. This instrument is also representative of gender norms (equal number of males and females) as well as ethnic diversity (following the 2000 US census data). Considering the makeup of the standardization sample, this instrument could appropriately be used with adolescents between ages 11 and 20.

*Quality of Test.* The internal consistency reliability of the RADS-2 was measured using Cronbach's alpha coefficient. The reliability of the total depression scale is .93 for the total school sample, .92 for the standardization sample and .94 for the clinical sample. The range of internal consistency reliability for the depression subscales of dysphoric mood, negative affect, negative self-evaluation, and somatic complaints were .80 to .87 for the total school sample, .79 to .89 for the standardization sample, and .81 to .87 for the clinical sample. Both Blair (2005) and Carlson (2005) support the strength of the internal consistency reliability for the RADS-2.

Similarly, the results of the test-retest reliability for the RADS-2 are also strong. There was a 2 week period between the initial test and retest administration. The test-retest reliability was measured using a Pearson correlation. Test-retest reliabilities for total depression were found to be  $r = .85$  for the total school sample,  $r = .86$  for the standardization sample, and  $r = .89$  for the clinical sample. Ranges for the subscales turned out to be  $r = .77$  to  $.84$  for the total school sample,  $r = .77$  to  $.85$  for the standardization sample, and  $r = .81$  to  $.87$  for the clinical sample. Although these reliability results are not extraordinarily strong, one can consider them quite strong due to the fact that they are measuring the construct of depression which is typically a less stable psychological domain. Additionally, low mean T-score differences were found

between each sample which further validates this instrument's test-retest reliability (Blair, 2005; Reynolds, 2002).

Item-with-total scale correlations were used to determine content validity. The median item-with-total scale correlation for the total depression scale was  $r = .53$  with a range of  $r = .27$  to  $r = .67$ . For each subscale of the RADS-2, the median item-with-total subscale correlations were  $r = .60$  for dysphoric mood,  $r = .66$  for negative affect,  $r = .64$  for negative self-evaluation, and  $r = .53$  for somatic complaints with ranges of  $r = .36$  to  $.70$ ,  $r = .54$  to  $.78$ ,  $r = .53$  to  $.69$ , and  $r = .46$  to  $.55$  for each of the subscales respectively (Reynolds, 2002).

The RADS-2 Professional Manual focused on concurrent validity rather than predictive validity to examine criterion validity due to the fact that this instrument is primarily concerned with an adolescent's current severity of depressive symptoms. Clinical interviews and self-report measures were used to determine criterion related validity. The Hamilton Depression Rating Scale, a 17-item clinical interview which measures depressive symptom severity, was used to interview 485 adolescents between 12 and 19 years old. From these interviews, the criterion related validity was found for the total depression scale ( $r = .82$ ) and the instrument subscales of dysphoric mood ( $r = .72$ ), negative affect ( $r = .54$ ), negative self-evaluation ( $r = .79$ ), and somatic complaints ( $r = .69$ ). To determine the construct validity, the RADS-2 examined both convergent and discriminant validity. Convergent validity was determined by examining the relationship between the RADS-2 and measures of self-esteem, anxiety, and suicidal behaviors among other established constructs related to depression. Correlation coefficients ranged from a low of  $r = .22$  up to  $r = .77$  for the depression total scale and RADS-2 subscales. Discriminant validity was examined by comparing the RADS-2 to other constructs unrelated to depression such as cognitive ability and social desirability. Correlation coefficients

ranged from  $r = -.38$  to  $r = -.14$  (Reynolds, 2002). The results support the strength of convergent and discriminant validity of the instrument as there is a strong relationship between the RADS-2 and other similar constructs (self-esteem, anxiety, and suicidal behaviors) and little relationship between the RADS-2 and other dissimilar constructs (academic performance and social desirability) (Blair, 2005; Carlson, 2005).

*Quality of Test Manual.* The test manual for the RADS-2 is written in a manner that makes it clear and easy to understand (Carlson, 2005). As an upgraded version of the original RADS manual, the RADS-2 is also very thorough. The research and evidence for its effectiveness is extensive (Blair, 2005).

*Administration and Scoring.* The RADS-2 is a brief instrument completed using pencil and paper. It relies on self-report responses. It may be administered aloud individually or in groups. Both administration and scoring are quick and easy (Carlson, 2005).

*Interpretation and Adequacy of Test Norms.* Standardization of the RADS-2 has accounted for gender ( $n=1,650$  for both males and females) and age ( $n=550$  for both males and females in ages 11 through 13, 14 through 16, and 17 through 20). Using the 2000 US Census data, the standardization sample is also representative of national ethnic (70.5% Caucasian, 12.1% African American, 11.8% Hispanic American, 4.3% Asian American, and 1.3% Native American) and socioeconomic status ( $M=8.37$ ) rates (Reynolds, 2002). The large sample size ( $N=3,300$ ) strengthen the norm group (Blair, 2005). The author references the 2000 census data as a close match to the RADS-2 norm sample, however no actual census data is provided in the manual (Blair, 2005; Carlson, 2005).

*Use in Counseling.* As the purpose of the test indicates, this instrument would be useful in determining the necessity for further investigation of depression (Blair, 2005; Carlson, 2005).

This instrument is most useful as a screening tool to detect levels of depressive symptomology either with large groups or with individuals (Carlson, 2005). Other means for diagnosing depression should be discussed as this instrument is not used for diagnosis.

#### Adolescent Anger Rating Scale

*Purpose of Test.* The Adolescent Anger Rating Scale (AARS) is designed to assist clinicians in identifying anger in adolescents between the ages of 11 and 19 years old. Specifically, it measures an adolescent's usual style of anger expression and anger control. It measures total anger as well as three domains of anger including reactive anger (RA), instrumental anger (IA) and anger control (AC) (Burney, 2001). The manual claims it is useful as a guide to treatment as well as a measure of treatment efficacy, however, due to less than ideal validity measures, it is recommended that its main purpose remain as a tool for treatment planning (Henington, 2003; Stephenson, 2003).

*Use with Population.* The AARS may most appropriately be used to screen adolescents between the ages of 11 and 19 years old. According to the manual, the norm sample contains adolescents (N=4,187) with girls representing 55% (n=2,329) and boys representing 45% (n=1,858) of the sample. The sample also contains several ethnic groups including African American, Asian, Caucasian, Hispanic, and other multi-ethnic adolescents. Adolescents from inner city, urban, and suburban environments are also represented (Burney, 2001).

*Quality of Test.* Internal consistency and test-retest reliability were examined to determine the reliability of the AARS. Using item analysis for the total anger and anger subscales of the AARS, the Cronbach's alpha results ranged from a low of .81 to a high of .92. Item-total correlations were also examined for each subscale. Correlations between items on the IA subscale ranged from  $r = .42$  to  $r = .69$ . Additionally, correlations ranged from  $r = .37$  to  $r =$

.64 for the RA subscale and from  $r = .34$  to  $r = .65$  for the AC subscale. Using 175 pairs of tests, the test-retest reliabilities were covered over an interval of 2 weeks between testing.

Correlations were calculated for the total anger scores ( $r = .79$ ) as well as the subscales of IA ( $r = .71$ ), RA ( $r = .71$ ), and AC ( $r = .74$ ) (Burney, 2001). The test developers suggest that IA is a stable characteristic while RA is a more fluid characteristic yet they both have a correlation of .71 (Stephenson, 2003).

Content and face validity was examined using an expert panel of professionals to agree on the item selection of the instrument. There were 101 initial potential items which were dwindled down to 41 items to comprise the instrument. The panel used the Item Development questionnaire to identify the behaviors that should be represented by the items on the AARS (Burney, 2001). This instrument has strong face validity and does not include a lie scale which weakens it to some degree (Stephenson, 2003). Criterion validity was found using correlations between the AARS scale and subscale scores and types of conduct referrals. There are instrumental anger type referrals related to cheating, skipping class and threatening behaviors and there are reactive anger type referrals related to fighting and argumentative behavior. Modest correlations were found for both IA and RA subscales in relation to frequency of conduct referrals (range of  $r = .10$  to  $r = .22$ ). Slightly higher correlations were found for the total anger scale each type of conduct referral (range of  $r = .27$  to  $r = .30$ ). A more significant negative relationship was found between the AC subscale and frequency of conduct referrals (range of  $r = -.29$  to  $r = -.36$ ) (Burney, 2001, Stephenson, 2003). Construct validity was found using convergent and discriminant validity measures. The AARS was compared to the similar anger control problems (ACP) and conduct problems (CP) subscales of the Conners-Wells Self-Report Scales. Results showed strong correlations between related constructs (RA and ACP,  $r = .61$  and

IA and CP,  $r = .57$ ) and negative correlations between unrelated constructs (AC and ACP,  $r = -.24$ , AC and PC,  $r = -.26$ ). Discriminant validity was found by comparing the AARS and the Multidimensional Anger Inventory (MAI), which measures different aspects of anger. The AARS reports relatively low correlations between the MAI and IA ( $r = .46$ ) and RA ( $r = .44$ ) subscales. The correlation between the MAI and AC was also low ( $r = -.11$ ). The shared variance between the MAI and the AARS subscales is small ( $r^2 = -.01$  to  $.21$ ) which supports the discrepancy between the two rating scales. (Burney, 2001).

*Quality of Test Manual.* The AARS test manual is clear and easy to read. The development of the instrument and its statistical information is highly praised (Henington, 2003)

*Ease of Administration and Scoring.* The AARS is easy to administer as well as score. Administration can take as little as five and up to 20 minutes to complete depending on whether it is being administered to an individual or a group (Henington, 2003; Stephenson, 2003). Scoring is also simple and can be done quickly (Henington, 2003).

*Interpretation and Adequacy of Test Norms.* The normative sample consists of 4,187 adolescents between the ages of 11 and 19. There is no differentiation given for regional demographic information of the norm group except to say that the sample was drawn from urban suburban and rural public schools in southeast, southwest and northern parts of the U.S. The sample was broken down by grade and gender (boys and girls; grades 6-8 and 9-12) since scores differ by age and gender. Multiple ethnicities were represented, however the African American population was possibly over represented at 31% of the sample (Stephenson, 2003).

*Use in Counseling.* This instrument would most effectively be used as a means to guide the treatment of anger in adolescents. It would be able to pinpoint those subscales of anger in a student. Due to weaker validity and strong face value, this instrument is not yet ready to be

relied on as a means to measure the effectiveness of treatment (Henington, 2003; Stephenson, 2003).

### Suicidal Ideation Questionnaire

*Purpose of Test.* Since suicidal ideation (SI) is an early warning for possible suicidal behavior, the Suicidal Ideation Questionnaire (SIQ) is designed to uncover the presence of suicidal thought patterns in adolescents. The SIQ manual clearly distinguishes the measurement of overall depression from the specific measurement of suicidal ideation. SIQ is designed to specifically measure the domain of suicidal ideation (Cramer, 1992; Reynolds, 1988). It may be used to screen for SI in large groups or in an individual, clinical setting.

*Use with Population.* This instrument has two forms: the SIQ, used for grades 10-12, and SIQ-JR, used for grades 7-9. The norm sample for the SIQ consists of 890 adolescents while the SIQ-JR includes 1,290 adolescents. These norm samples have been obtained from a high school and two junior high schools in urban and suburban areas of the Midwestern U.S. Although the demographic area is limited, the age range makes it appropriate to use with adolescents.

*Quality of Test.* Reliability was examined using measures of internal consistency and test-retest reliability. Cronbach's alpha was used to measure internal consistency. By grade level, the SIQ alpha was strong (range from .969 to .974) as was the SIQ-JR (range from .932 to .938). Only the SIQ was investigated for test-retest reliability. A sample of 801 adolescents were tested and retested over a four week interval. Results showed a test-retest reliability coefficient of  $r = .72$ . The instrument's validity was tested by looking at content and construct validity. Item-total scale correlations were used to examine content validity of the SIQ and SIQ-JR. For the SIQ with a total norm sample of 890 adolescents in grades 10, 11, and 12, item-total sample correlations ranged from  $r = .84$  to  $r = .41$ . For the SIQ-JR with a total norm sample of



1,290 adolescents in grades 7, 8 and 9, item-total sample correlations ranged from  $r = .86$  to  $r = .44$ . Construct validity was determined by examining convergent and discriminant validity. The SIQ and SIQ-JR were compared to related constructs such as depression, hopelessness and anxiety to examine convergent validity. The SIQ used the Beck Depression Inventory (BDI), the Reynolds Adolescent Depression Scale (RADS), the Learned Helplessness Scale, and the Children's Manifest Anxiety Scale – Revised to find convergent or positive correlation relationships ( $r = .69$  and  $.70$  for BD,  $r = .58, .61$  and  $.63$  for RADS,  $r = .36$  for learned helplessness and  $r = .56$  and  $.58$  for Manifest Anxiety). The SIQ-JR used also used the RADS and Manifest Anxiety Scale as well as the Children's Depression Inventory to find convergent relationships ( $r = .55$  and  $.59$  for RADS,  $r = .54$  and  $.56$  for Manifest Anxiety, and  $r = .65$  and  $.66$  for Child Depression Inventory). For discriminant validity, the SIQ and the SIQ-JR were compared to unrelated constructs such as self-esteem and academic performance. The SIQ used the Rosenberg Self-Esteem Scale and the Academic Self-Concept Scale – High School Version to find discriminant relationships ( $r = -.52$  and  $-.56$  for self-esteem scale and  $r = -.39$  and  $-.42$  for academic self-concept). The SIQ-JR also used those same comparative instruments ( $r = -.54$  and  $r = -.42$  respectively) (Reynolds, 1988). The SIQ does have strong face validity, but that may be acceptable because the adolescent might still honestly report, using it as a cry for help (Cramer, 1992). On the other hand, false positives might be an outcome if students answer in a way that brings adult attention to them when they are not actually thinking about suicide. In balancing the efficiency of the instrument with its thoroughness, the SIQ remains a strong option for suicidal ideation screening. It may produce false positives, but it is better to include more and invite precaution when dealing with something possibly fatal like suicide (Conoley, 1992).

*Quality of Test Manual.* The SIQ test manual written clearly and is easy to understand. Although there is ample discussion of how this instrument may be used for screening adolescents for suicidal ideation, there are still improvements that can be made in discussing its use in intervening when an adolescent is identified as at risk for suicidal thoughts or behavior (Conoley, 1992)

*Ease of Administration and Scoring.* The administration of the SIQ and SIQ-JR takes approximately 5 to 10 minutes and can be administered for individuals as well as in groups (Reynolds, 1988). The SIQ is 30 items long (15 items for the SIQ-JR) and can be hand scored. Scoring is simple as a template is used (Reynolds, 1988).

*Interpretation and Adequacy of Test Norms.* The SIQ and SIQ-JR norm sample is comprised of 890 and 1,290 adolescents respectfully. Gender is almost equally represented in the normative groups. In the SIQ norm sample white (78%), black (19%), Asian (1.5%), and Hispanic (0.4%) ethnicities are represented. Likewise, in the SIQ-JR norm sample the same ethnicities are represented (74%, 22%, 1.7%, and 0.8% respectively) (Reynolds, 1988). The gender balance as well as ethnic representation in the norm sample prove to be adequate for the population.

*Use in Counseling.* This instrument may be used effectively as a tool in a school crisis response situation or prevention program. It is important that the SIQ be used in a situation in which the administrator can follow up in a timely manner due to the nature of the content. With suicide as the focus, thoroughness and punctuality are of utmost importance (Cramer, 1992). Additionally, there is the cutoff score, although it is not entirely reliable. It was found that one third of those who attempted suicide actually scored below the cutoff score. Thus, administrators of this instrument should not base their judgment solely on this cutoff score (Cramer, 1992).

## References

- Blair, K. A. (2005). [Review of the Reynolds Adolescent Depression Scale – 2<sup>nd</sup> Edition]. In *The sixteenth mental measurements yearbook*. Available from:  
<http://web.ebscohost.com.proxy.lib.pdx.edu>
- Burney, D. M. (2001). Adolescent Anger Rating Scale: Professional manual. Lutz, FL:  
Psychological Assessment Resources, Inc.
- Carlson, J. F. (2005). [Review of the Reynolds Adolescent Depression Scale – 2<sup>nd</sup> Edition]. In *The sixteenth mental measurements yearbook*. Available from:  
<http://web.ebscohost.com.proxy.lib.pdx.edu>
- Cramer, J. C. (1992). [Review of the Suicidal Ideation Questionnaire]. In *The eleventh mental measurements yearbook*. Available from: <http://web.ebscohost.com.proxy.lib.pdx.edu>
- Conoley, C. (1992). [Review of the Suicidal Ideation Questionnaire]. In *The eleventh mental measurements yearbook*. Available from: <http://web.ebscohost.com.proxy.lib.pdx.edu>
- Henington, C. (2003). [Review of the Adolescent Anger Rating Scale]. In *The fifteenth mental measurements yearbook*. Available from: <http://web.ebscohost.com.proxy.lib.pdx.edu>
- Reynolds, W. M. (1988). Suicidal Ideation Questionnaire: Professional manual. Lutz, FL:  
Psychological Assessment Resources, Inc.
- Reynolds, W. M. (2002). Reynolds Adolescent Depression Scale – 2<sup>nd</sup> Edition: Professional manual. Lutz, FL: Psychological Assessment Resources, Inc.
- Stephenson, H. (2003). [Review of the Adolescent Anger Rating Scale]. In *The fifteenth mental measurements yearbook*. Available from: <http://web.ebscohost.com.proxy.lib.pdx.edu>