

RELIABILITY AND VALIDITY EVIDENCE OF BENCHMARK ASSESSMENT TOOL
(BAT) FOR SOCIAL AND EMOTIONAL LEARNING

Sara Harris
Bradley University

Adena Meyers and Mark E. Swerdlik
Illinois State University

Patrick Arity and Paul Cicciarelli
East Peoria Illinois School District

Abstract

The purpose of this study was to evaluate the reliability and validity of the Benchmark Assessment Tool (BAT), a recently developed tool designed to assess social and emotional learning in students from grades kindergarten through 8. The BAT is based on research-based, core competencies of social and emotional learning. The current study presents three separate studies that demonstrate moderate to high test-retest reliability, and support for the construct and criterion-related predictive validity of the BAT for the purpose for which it was developed.

RELIABILITY AND VALIDITY EVIDENCE OF BENCHMARK ASSESSMENT TOOL FOR SOCIAL AND EMOTIONAL LEARNING

Research consistently shows that academic and social/emotional skills are intimately related and jointly contribute to children's learning in school (e.g., Shriver & Weissberg, 2005). Recognizing this critical link, beginning in 2004, the Illinois State Board of Education mandated that every school district create and submit a policy for assessing and monitoring students' progress in acquiring developmentally appropriate social/emotional skills. As a result, the Benchmark Assessment Tool (BAT) was developed by School Psychologist Paul Cicciarelli and School Social Worker Patrick Arity in 2008 (<https://besstweb.com/?p=index>) to identify students who exhibit behavioral difficulties and/or have deficits in social/emotional skills that may interfere with learning in the classroom.

The BAT addresses 3 of 5 core competencies identified by CASEL researchers that provide the foundation for social and emotional learning: (1) Develop self-awareness and self-management skills to achieve school and life success. (2) Use social-awareness and interpersonal skills to establish and maintain positive relationships. (3) Demonstrate decision-making skills and responsible behaviors in personal, school, and community contexts. It is a brief assessment that is designed to take approximately one minute to complete for each student, or 30 minutes total for an average-sized classroom. The BAT is for use with students enrolled in grades K-8. Classroom teachers can quickly enter students' scores into an online data management program (BESTTWEB) that identifies those who may be at-risk for social/emotional learning difficulties, as well as monitoring their progress over time. The BAT is similar to other well-established data management programs, such as AIMSWEB, which provides an efficient system for assessing

and monitoring basic reading and mathematics skills (<http://www.aimsweb.com/>). See Table 1 for sample BAT goals by grade.

Table 1

Sample BAT Goals by Grade

Goal	3 rd Grade	4 th Grade	5 th Grade
Develop Self-Awareness and Self-Management Skills	Student can identify a range of emotions and recognize situations that cause them.	Student can describe personal skills and interests that one wants to develop.	Student can describe the steps in setting and working toward goal achievement.
Use Social Awareness and Interpersonal Skills to Maintain Positive Relationships	Student can explain the feelings and perspectives of others.	Student can analyze ways to work effectively in groups.	Student can describe causes and consequences of conflicts.
Demonstrate Decision-Making Skills	Student demonstrates an understanding of how social norms influence decision making.	Student can identify and apply the steps of systematic decision making.	Student identifies and performs roles that contribute to the school community.

Purpose of Study

As the BAT is the only instrument available for Tier I universal screening that is tied directly to the Illinois Social-Emotional Learning Standards, it has generated a great deal of interest and has been adopted by a number of school districts. The goal of the present set of investigations is to provide support for the reliability and validity of the Benchmark Assessment Tool (BAT). The researchers were specifically interested in the following two research questions:

- (1) Based on the 5 core competencies of the CASEL, is the BAT a reliable measure of social and emotional learning in school-age children?
- (2) Based on the 5 core competencies of the CASEL, is the BAT a valid measure of social and emotional learning in school-age children?

Methodology

Study 1: Test-Retest Reliability

One hundred and eighty-three students were recruited for this study—59 third graders, 66 fourth graders, and 58 fifth graders enrolled in a Midwestern school district. The sample consisted of 105 boys and 78 girls from a single elementary school. The majority of the participants, 89 percent, were Caucasian, 6.6 percent were multiracial, 2 percent were African American, 1.5 percent were Hispanic, and .5 percent were Native American. A total of 15 teachers (12 general education teachers and 3 special education teachers) rated students using the Benchmark Assessment Tool (BAT) in the Fall of 2008 and Winter of 2009, with approximately 12 weeks between pre- and posttest administration. No social/emotional interventions were implemented between the two data collection periods. Table 1 includes sample questions by grade level, which were used to analyze test-retest reliability. A Pearson Rank order correlation was computed using the BAT composite score. Results indicated that the overall correlational coefficient for test-retest reliability of the BAT was .85, $p < .05$ across all grade levels.

Study 2: Construct Validity

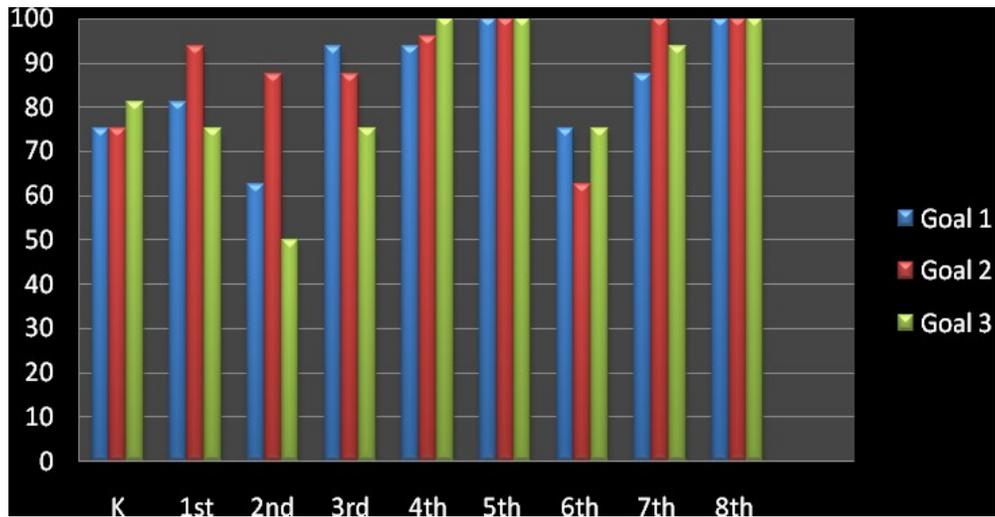
To assess the construct validity of the BAT, a study was designed utilizing an expert panel of 8 school-based professionals, ranging in years of experience from one to 29—three school psychologists, two school social workers, one principal, one special education coordinator, and one director of special education services. A questionnaire was designed to assess the developmental nature of the various BAT items and behaviors associated with the domains included as part of the Illinois State Standards of Social/Emotional Learning (SEL). The assessment covered the social/emotional learning standards for grades K through 8 and included the following 3 goals—(1) Develop self-awareness and self-management skills to achieve school

and life success; (2) Use social-awareness and interpersonal skills to establish and maintain positive relationships; and (3) Demonstrate decision-making skills and responsible behaviors in personal, school, and community context—and items included under each goal. Participants were asked to identify specific BAT items that should be included under each goal. Specifically, for each grade level and goal, respondents were presented with a list of items, or behaviors, and asked to match them with the appropriate SEL goal.

Data analyses involved computing percentage of respondent agreement for each goal and its corresponding behaviors (items), as included on the BAT. For example, what percentage of respondents accurately identified behavioral items associated with Goal 1: Develop self-awareness and self-management skills to achieve school and life success? Percentages were based on accuracy for all grade levels. Overall, findings indicated a high rate of accuracy for all 3 goals—85 percent for Goals 1 and 2 and 81 percent for Goal 3, suggesting evidence of construct validity for the BAT. Figure 1 illustrates percentage of correct answers by goal and grade level.

Figure 1

Expert Agreement: Percentage of Correct Answers by Grade Level



Study 3: Criterion-Related Predictive Validity of the Benchmark Assessment Tool

To determine if the Benchmark Assessment Tool (BAT) predicts actual behavior in the classroom, researchers evaluated its relationship to an existing Positive Behavioral Support (PBS) program for 3rd, 4th, and 5th graders enrolled in a Midwestern school district. All of the participants were from the same school. A total of 189 participants were included in this study, with approximately equal numbers of boys and girls (97 boys, 92 girls). The sample consisted of 178 Caucasian students, 7 African-American students, 3 Asian students, and 1 Hispanic student. All participants came from the same school.

BAT ratings were completed in the Fall, Winter, and Spring of the 2008-2009 school year. The PBS program, which served as the criterion measure, involved issuing “tickets” for inappropriate behavior that violates school expectations. As part of the school code of conduct,

students are expected to be respectful, responsible, and resourceful. Students who are observed violating one of these expectations are issued a ticket.

The researchers were interested in whether or not BAT scores predicted the number of tickets issued for misbehavior. Findings indicated that 60 tickets were issued for 3rd graders, 78 tickets were issued for 4th graders, and 59 tickets were issued for 5th graders over the course of the investigation.

Pearson correlations were computed to determine if a relationship existed between composite BAT scores and total number of tickets issued. Overall, there was an inverse relationship between BAT scores and tickets issued. That is, students with high BAT scores, which reflect average to above average social/emotional learning, were less likely to receive tickets for behavioral violations than students who were identified as at-risk for adjustment difficulties. Pearson correlations by grade level included $r = -0.39$, $p < .05$ for 3rd grade, $r = -0.63$, $p < .05$ for 4th grade, and $r = -0.49$, $p < .05$ for 5th grade. These correlations represent a moderate, inverse relationship suggesting that the BAT represents a good predictor of classroom behavior.

In addition to computing correlations between composite BAT scores and total number of tickets issued, correlations were computed for individual BAT goals and total number of tickets issued: (1) Develop Self-Awareness and Self-Management Skills, $r = -0.47$, $p < .05$; (2) Use Social Awareness and Interpersonal Skills to Maintain Positive Relationships, $r = -0.91$, $p < .05$; and (3) Demonstrate Decision-Making Skills, $r = -0.81$, $p < .05$. Thus, in addition to the moderate, inverse relationship between composite BAT scores and classroom behavior, there is a moderate- to high correlation between individual BAT goals and classroom behavior.

Discussion

Results of three reliability and validity studies presented as part of this research provides evidence that the BAT is both a reliable and valid measure of social/emotional learning, as identified by the Illinois Learning Standards, in elementary and middle school students. These findings are especially important and applicable in that the BAT is aligned with current Illinois Learning Standards, which are based on the empirically-supported CASEL core competencies. Because the BAT helps to identify social/emotional strengths and challenges, it can also be part of the universal tier I social-emotional screening within an RTI system. The BAT allows educators to quickly identify students who may require intervention and similar to curriculum-based measures like AIMSWEB (<http://www.aimsweb.com/>), the BAT is designed to monitor individual student progress over time. Thus, it is a highly efficient and applicable classroom tool and may serve as an effective tool of social/emotional screening for children in grades kindergarten through 8. This tool may also be an option for other states that are striving to develop SEL standards and screening measures consistent with those standards.

Although the preceding studies suggest that the BAT is a reliable and valid assessment instrument, the studies do have important limitations. First, although the BAT focuses on grades kindergarten through 8, the studies presented only include students from kindergarten through 5th grade. In addition, the students were not from a single school; rather, they were drawn from classrooms throughout a school district. Second, the majority of the participants were Caucasian, limiting the generalizability of findings to students from other racial/ethnic backgrounds. Third, the participants were from a single Midwestern school district that may not be representative of school districts in other parts of Illinois or the United States. Fourth, although the BAT is

designed for progress monitoring, no research currently exists on its ability to detect changes in social and emotional learning over time.

Based on the limitations of the current study, future research should focus on the following:

- (1) Include a larger sample of students from various grade levels, schools, racial/ethnic, and socioeconomic backgrounds.
- (2) Compare the BAT to existing, empirically-supported measures of social and emotional learning in school-age children, such as the Behavioral Assessment System for Children.
- (3) Evaluate the sensitivity of the BAT for progress monitoring.

REFERENCES

- Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111-127.
- Buchanan, R., Gueldner, B. A., Tran, O. K., & Merrell, K. W. (2008). Social and emotional learning in classrooms: A survey of teachers' knowledge, perceptions, and practices. *Journal of Applied School Psychology, 25*, 187 – 203
- Burgess, K. B., Wojslawowicz, J. C., Rubin, K. H., Rose-Krasnor, L., & Booth-LaForce, C. (2006). Social information processing and coping strategies of shy/withdrawn and aggressive children. Does friendship matter? *Child Development, 77*, 371-383.
- Cain, K. M., & Dweck, C. S. (1995). The relation between motivational patterns and achievement cognitions through the elementary school years. *Merrill Palmer Quarterly, 41*, 25-52.
- Clements, M. A., Reynolds, A. J., & Hickey, E. (2004). Site-level predictors of children's school and social competence in the Chicago Child-Parent Centers. *Early Childhood Research Quarterly, 19*, 273–296.
- Cohen, R. J., & Swerdlik, M. E. (2010). *Psychological testing and assessment (7th ed.)*. Boston, MA: McGraw-Hill.
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74-101.
- Denham, S. A. (2006). Social-emotional competence as support for school readiness: What is it and how do we assess it? *Early Education and Development, Special Issue: Measurement of School Readiness, 17*, 57-89.
- Denham, S. A., Workman, E., Cole, P. M., Weissbrod, C., Kendziora, K. T., & Zahn-Waxler, C. (2000). Prediction of externalizing behavior problems from early to middle childhood: The role of parental socialization and emotion expression. *Development and Psychopathology, 12*, 23-45.
- Dweck, C. S. (2002). Beliefs that make smart people dumb. In R.J. Sternberg (Ed.), *Why smart people can be so stupid*. New Haven : Yale University Press.
- Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional and personality development* (Vol. 3, pp. 646-718). Hoboken, NJ: John Wiley & Sons.

- Eisenberg, N. Guthrie, I. K., Murphy, B. C., Shepard, S. A., Cumberland, A., & Carlo, G. (1999). Consistency and development of prosocial dispositions: A longitudinal study. *Child Development, 70*, 1360-1372.
- FitzGerald, D. P., & White, K. J. (2003). Linking children's social worlds: Perspective-taking in parent-child and peer contexts. *Social Behavior and Personality, 31*, 509-522.
- Garner, P. (1996). The relations of emotional role-taking, affective/moral attributions and emotional display rule knowledge to low-income school-age children's social competence. *Journal of Applied Developmental Psychology, 17*, 19-36.
- Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. New York: Bantam.
- Goleman, D. (1998). *Working with emotional intelligence*. New York: Bantam.
- Gopnik, A., Meltzoff, A., & Kuhl, P. (1999). *The scientist in the crib: Minds, brains, and how children learn*. New York: William Morrow and Company.
- Guerra, N. G., Enron, L. D., Huesmann, L. R., Tolan, P. H., & VanAcker, R. (1997). A cognitive-ecological approach to the prevention and mitigation of violence and aggression in inner-city youth. In P. Fry & K. Bjorkqvist (Eds.), *Cultural variation in conflict resolution: Alternatives for reducing violence* (pp. 199-213). New York: Polonium.
- Hastings, P. D., Zahn-Waxler, C., Usher, B., Robinson, J., & Bridges, D. (2000). The development of concern for others in children with behavior problems. *Developmental Psychology, 36*, 531-546.
- Heyman, G. D. and Dweck, C. S. (1998). Children's thinking about traits: Implications for judgments of the self and others. *Child Development, 69*, 392-403.
- Ladd, G. W., Buhs, E. S., & Seid, M. (2000). Children's initial sentiments about kindergarten: Is school liking an antecedent of early participation and achievement? *Merrill-Palmer Quarterly, 46*, 255-279.
- Ladd, G. W., Herald, S. L., & Kochel, K. P. (2006). School readiness: Are there social prerequisites? *Early Education and Development, 17*, 115-150.
- Lamb-Parker, F., LeBuffe, P., Powell, G., & Halpern, E. (2008). A strength-based, systematic mental health approach to support children's social and emotional development. *Infants & Young Children, 21*, 45-55.
- Lopes, P. N., Brackett, M., Nezlek, J. B., Schutz, A., Sellin, A., & Salovey, P. (2004). Emotional

- intelligence and social interaction. *Personality and Social Psychology Bulletin*, 30, 1018-1034.
- Merrell, K. W. (2010). Linking prevention science and social and emotional learning: The Oregon resiliency project. *Psychology in the Schools*, 47, 55-70.
- National Institute of Child Health and Human Development Early Child Care Research Network (2008). Mothers' and fathers' support for child autonomy and early school achievement. *Developmental Psychology*, 44, 895-907.
- Rose-Krasnor, L. (1997). The nature of social competence: A theoretical review. *Social Development*, 6, 111-135.
- Saarni, C. (1999). *The development of emotional competence*. New York: Guilford.
- Salovey, P., & Pizzaro, D. A. (2003). The value of emotional intelligence. In R. J. Sternberg, J. Lautrey, & T. I. Lubart (Eds.), *Models of intelligence: International perspectives* (pp. 263-278). Washington, DC: American Psychological Association.
- Shonkoff, J., & Phillips (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.
- Shriver, T. P., & Weissberg, R. P. (2005). No emotional left behind. *New York Times*, August 16, 2005. <http://www.casel.org/downloads/NYToped.doc>
- Wojslawowicz Bowker, J. C., Rubin, K. H., Burgess, K. B., Booth-LaForce, C., & Rose-Krasnor, L. (2006). Behavioral characteristics associated with stable and fluid friendship patterns in middle childhood. *Merrill-Palmer Quarterly*, 52, 671-693.