

# Challenges for Behavioral Assessment in Brazilian Schools

## *Desafios para avaliação comportamental nas escolas brasileiras*

Nicholas F. **BENSON**<sup>1</sup>  0000-0001-8180-4243

Solange M. **WECHSLER**<sup>2</sup>  0000-0002-9757-9113

Brandon **PARKER**<sup>1</sup>  0000-0002-8386-1951

### Abstract

In this article we begin by providing an overview of the need to increase the frequency with which behavioral assessment is used in Brazilian schools. We then describe various methods of behavioral assessment. Next, we discuss various purposes for assessing behavior in school settings. We then discuss general challenges that impact the assessment of behavior. Finally, we conclude by discussing the importance of identifying psychological and social concerns early in childhood and call for the development of additional evidence-based assessment tools that are appropriate for use in Brazilian schools.

**Keywords:** Assessment; Behavior; Brazilian Schools; Rating Scales.

### Resumo

*Neste artigo, iniciamos com uma visão geral da necessidade de aumentar a frequência com qual avaliações de comportamento são utilizadas nas escolas brasileiras. Daí, descrevemos vários métodos de avaliações de comportamento. Próximo, discutimos os vários propósitos da avaliação do comportamento no ambiente escolar. Nós discutimos os desafios gerais que influenciam a avaliação do comportamento. Finalmente, concluímos com uma discussão sobre a importância da identificação de problemas psicológicos e sociais cedo durante a infância, e realçamos a necessidade do desenvolvimento de avaliações e ferramentas adicionais baseadas em evidências, que seriam apropriadas para uso nas escolas brasileiras.*

**Palavras-chave:** *Avaliação; Comportamento; Escolas Brasileiras; Escalas de Classificação.*

Behavioral assessment assumes an ecological approach to evaluating the behaviors that students display in schools (Shapiro & Kratochwill, 2000). Although within-student traits and characteristics may be assessed, this approach emphasizes the importance of environmental factors and student-environment interactions. While diagnostic considerations are sometimes the focus of behavioral assessment, it also encompasses broader purposes such as informing instructional or behavioral intervention and improving

▼ ▼ ▼ ▼ ▼

<sup>1</sup> Baylor University, School of Education, Department of Educational Psychology, School Psychology Program, One Bear Place #97301, Waco, TX, United States. *Correspondência para/Correspondence to:* N.F. BENSON: *Correspondência para/Correspondence to:* <Nicholas\_Benson@baylor.edu>.

<sup>2</sup> Pontifícia Universidade Católica de Campinas, Centro de Ciências da Vida, Programa de Pós-Graduação em Psicologia. Campinas, SP, Brasil.

▼ ▼ ▼ ▼ ▼

*Como citar este artigo/How to cite this article*

Benson, N. F., Wechsler, S. M., & Parker, B. S. Challenges for behavioral assessment in Brazilian schools. *Estudos de Psicologia* (Campinas), 36, e190010. <http://dx.doi.org/10.1590/1982-0275201936e190010>



student outcomes. Although it is required for some purposes (e.g., determining eligibility for special services), behavioral assessment is best viewed as a process for solving problems rather than as a procedural requirement (Chafouleas, Riley-Tillman, & Sugai, 2007).

In this article we begin by providing an overview of the need to increase the frequency with which behavioral assessment is used in Brazilian schools. We then describe various methods of behavioral assessment. Next, we discuss various purposes for assessing behavior in school settings. We then discuss general challenges that impact the assessment of behavior. Finally, we conclude by discussing the importance of identifying psychological and social concerns early in childhood and call for the development of additional evidence-based assessment tools that are appropriate for use in Brazilian schools.

## Context of Brazilian Schools

Psychological and social difficulties often manifest early in childhood and can jeopardize the well-being of both current and future generations. A systematic review of studies from middle- and low-income countries suggests that in the past decade there have been alarming increases in the prevalence of mental health problems such as anxiety, depression, and conduct disorder among children and adolescents (Kieling et al., 2011). In Brazil, for instance, significant increases in internalizing and externalizing problems were identified among children during an 11-year period, with most of these increases occurring for children and adolescents living in poverty and raised by parents with little formal education (Matijasevich et al., 2014). This is a public health issue with devastating consequences.

School psychologists are trained to directly observe students' behavior and to obtain indirect information regarding students' psychological and social functioning from parents and teachers. The information obtained is useful when developing intervention plans to reduce psychological and social problems and increase well-being. However, efforts to improve students' psychological and social functioning require that staff are adequately trained to recognize signs of psychopathology and elicit self-reports of symptoms from students who are experiencing distress. Additionally, these efforts require behavioral assessment tools with evidence of validity to support their use for purposes such as screening, diagnosis, and treatment planning.

In Brazil, teacher preparation and training are subpar. According to a survey conducted by the Organization for Economic Co-operation and Development (2015), about 25% of teachers in secondary settings did not graduate from a teacher-education program. Federal regulations mandate that teachers hold at least a college degree, but it is clear that these mandates have not been enforced with fidelity (Barretto, 2015). Regional disparities also apply to teacher training: Many teachers in rural areas and poorer regions have inadequate levels of education and training (Oakland & Wechsler, 1990). In many schools in the North and Northeast, teachers do not hold a college degree or do not have any training in education (Marcondes, Leite, & Ramos, 2017). In some cases, teachers did not complete their own public education (Marcondes, 1999). In addition to regional inequalities, many educators teach courses in disciplines they did not study. According to national surveys, only 54% of Portuguese teachers were trained to teach Portuguese courses, and only 38.6% of mathematics teachers were trained in the respective discipline (Gatti, 2014; Marcondes et al., 2017). Teacher education programs also do not adequately prepare future teachers for the challenges and realities encountered in schools and classrooms including poor behavior (Marcondes et al., 2017). Online teacher education courses have become more popular, but the coursework does not sufficiently prepare future teachers to be impactful and successful (Gatti, 2014; Marcondes et al., 2017).

These challenges highlight the need for school psychologists to be a present and active agent in Brazil's education system. Historically, school psychologists have been nearly absent from public schools and have had a limited presence in private schools due to a lack of governmental mandates (Valle, 2003).

Consequently, students have not had adequate academic, social, emotional, and behavioral supports to be successful in school; additionally, they have experienced a less stimulating educational career due external factors such as poor funding and inadequate teacher training.

School psychologists are uniquely trained to consult and collaborate with teachers and can assist with the inclusion of children with a range of disorders (Fonseca, Frutas, & Negreiros, 2018; Sant'Ana, Euzébio, Lacerda, & Guzzo, 2009). School psychologists also can work jointly with administrators, teachers, and parents on systemic and individual preventative programs, which can lead to better academic and psychological outcomes for children who experience obstacles to learning (Valle, 2003; Mendes, 2010). As school psychologists become more integrated into public and private school systems and take an active role in advocating for students along with other school personnel, obstacles to better academic performance (e.g., behavioral/emotional disorders, intellectual disability, learning disabilities, other disorders, poor instruction, other external factors, etc.) have greater potential to be reduced.

There are several methods of behavioral assessment that can be utilized to help accomplish goals such as improving behavioral and academic outcomes. It is common for practitioners to utilize multiple methods of behavioral assessment. An overview of these methods is provided to highlight the options available to mental health and educational professionals.

## Method

Behavioral assessment methods can be placed on a continuum ranging from direct to indirect (Cone, 1978). At the direct end of the continuum, behavior is observed in real time under naturally occurring conditions (Shapiro & Browder, 1990). Assuming observations are made with integrity, direct observations in natural settings will provide the most accurate information about behavior. However, this ideal is often impractical due to time and financial constraints. For example, it would not be practical to devote numerous hours to the direct observation of students who display problematic behaviors (e.g., self-harm, physical assault of others) that occur infrequently (e.g., an average of two times per week) at times that are not readily predictable. Observation of infrequent behaviors can be facilitated by creating analogue environments that simulate occasions during which the target behavior occurs naturally. Alternatively, students can self-monitor by collecting data on their own behavior while it occurs, although there is the potential for biasing factors such as self-interest and image management to impact the validity of data obtained using this method.

Indirect assessment methods are commonly used in schools (Benson et al., 2019), as the benefits of the data obtained tend to be high compared to the relatively low costs associated with these methods. Indirect approaches include those methods in which data are collected after the behaviors have already occurred (Shapiro & Browder, 1990). Indirect methods include self-reports provided during interviews or obtained using checklists or a rating scale format. The most indirect information is gathered from collateral informants such as teachers and parents.

### *Direct assessment*

Direct behavioral observation “requires careful attention to specifying what and how long the behaviors are observed, where and how observations are made, and how they are recorded” (Benson, 2010, p.74). In other words, it is important for the observations to be systematic. Skilled observers often develop coding systems tailored to the specific purposes of their observations, although structured observation forms have been developed for use with observations of various types. For example, the Student Observation System

was developed as a direct observation component of the comprehensive Behavior Assessment System for Children-Third Edition (BASC-3); (Reynolds & Kamphaus, 2015) that also includes indirect components.

Observational data can be qualitative or quantitative. While qualitative information can be useful for descriptive purposes, the utility of observational data is largely dependent on quantification. Quantification is needed so that patterns and relationships can be identified using visual analysis and/or statistical analysis. Behaviors can be quantified based on four observable characteristics: frequency, duration, latency, and intensity (Salvia, Ysseldyke, & Witmer, 2009). Frequency is measured by counting the number of occurrences of a behavior during a fixed period. The duration of a behavior is measured by recording the length of time it occurs (e.g., how long a student's tantrum lasts from onset to termination). Latency is measured by assessing the length of time that it takes for an individual to exert a behavior (e.g., how long it takes a student to start working on a worksheet in class). Finally, the intensity of a behavior refers to the strength or force with which it is expressed.

A variety of recording techniques can be used during behavioral observations. Narrative recording (i.e., a continuous description of behavior in progress) provides qualitative data that can be organized in an antecedent-behavior-consequence sequence and used to identify and operationally define target behaviors, identify antecedents and consequences of target behaviors, develop data collection procedures, and formulate behavioral goals (Sulzer-Azaroff & Mayer, 1991). Permanent product recording can be used to count real or concrete objects or outcomes (e.g., recording the number of books read) that result from a behavior. Event recording (e.g., tallying the number of times a specific behavior occurs during a specified time interval), duration recording (i.e., recording the length of time a behavior occurs), and interval time sampling (i.e., recording the presence or absence of a specific behavior during a specified time interval) are methods used to measure transitory behaviors, those behaviors that do not leave an enduring product (Sulzer-Azaroff & Mayer, 1991).

Systematic methods of observation and analysis have proven useful for supporting positive behaviors and constraining problematic behaviors (Cooper, Heron, & Heward, 2006; Johnston, Foxx, Jacobson, Green, & Mulick, 2006). Desired behaviors can be promoted, and problematic behaviors constrained, by identifying and manipulating contingencies between a targeted behavior and a stimulus or event that precedes it (i.e., an antecedent) and/or a stimulus or event that occurs immediately after it (i.e., a consequence). Notably, not all stimuli or events that are temporally adjacent to target behaviors are functionally related to the target behavior. Functional relations exist when the presence of a stimulus or event alters the probability that a target behavior occurs (Skinner, 1953). Systematic, performance-based methods of behavioral observation can be used to identify functional relations and determine which stimuli or events should be manipulated to cause behavioral change. The application of these methods to identify and manage environmental contingencies in real-world settings such as school classrooms is known as Applied Behavior Analysis; (Baer, Wolf, & Risley, 1968).

### *Indirect assessment*

As previously noted, information regarding behaviors can be gathered indirectly using interviews and rating scales. Interviews yield qualitative information, while rating scales yield quantitative data. A recent survey of school-based practitioners in the United States indicates that behavior rating scales, particularly the BASC-3 (Reynolds & Kamphaus, 2015), are the assessment instruments used most frequently by psychologists who practice in school settings (Benson et al., 2019) The popularity of rating scales undoubtedly relates to the fact that they are an efficient way to acquire information.

Behavior rating scales contain a limited set of items that measure the frequency with which relevant behaviors are displayed. Ratings provide an indirect measure of behaviors that summarize behaviors an

informant observes over an extended period. The professional administering the rating scale typically discusses the directions for completion with the respondent and explains why the information is needed as well as how it will be used. Benefits of rating scales include efficiency, ease of administration and scoring, and relatively low cost. Moreover, although it is not feasible to observe more than a few students when conducting direct behavioral observations, rating scales feature normative samples that allow comparisons with representative samples of same-age peers.

Behavior rating scales also have some important limitations, largely arising from error variance (Merrell, 1999). Important sources of error include setting variance (i.e., variance attributable to behavioral differences across settings), temporal variance (i.e., variance attributable to behavioral differences across time), instrument variance (i.e., variance attributable to the rating scale used), and source variance (i.e., variance attributable to rater subjectivity and bias). Setting variance reflects the phenomenon that the environment has a unique interaction with behavior, and temporal variance notes that behaviors vary and are inconsistent over time and across settings. Instrument variance acknowledges that different scales measure related but slightly different constructs and represent distinct normative samples, and thus, should not be expected to completely align with each other (Merrell, 1999). With respect to source variance, well-known response biases include (a) halo effects (i.e., an impression formed based on specific behaviors is over-generalized), (b) leniency or severity effects whereas a rater tends to rate all students either too leniently or too harshly, (c) central tendency effects whereas raters avoid extreme ratings at either end of the scale, and (d) self-serving bias whereas inaccurate responses are provided to gain or avoid services (Dowdy, Twyford, & Sharkey, 2013). Given these limitations, data obtained from rating scales should be scrutinized and compared with objective data obtained using other methods such as direct observations (Witt, Heffer, & Pfeiffer, 1990). Although threats to validity are a limitation of behavior rating scales, serious errors can be avoided by constructing scales that encourage accurate responding.

Numerous rating scales have been developed. Some rating scales measure a single construct (e.g., depression, attention, and anxiety) while others are broad and measure multiple constructs. The BASC-3 is an example of a broad-band rating scale developed to measure multiple constructs. The BASC-3 system includes self-report forms as well as parents and teacher rating forms, each measuring several areas (e.g., adaptive skills, externalizing problems, internalizing problems, school problems, clinical scales measuring specific constructs such as anxiety and aggression). A common narrow band rating scale is the Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 2008) which measures various signs and symptoms associated with anxiety (e.g., physiological anxiety, worry, social anxiety).

Notably, rating scales such as the BASC-3 and RCMAS are not utilized Brazil. While numerous rating scales have been developed for use in other countries, there is a pressing need for behavioral assessment tools with sufficient evidence of validity to support their use in Brazilian schools. According to the Federal Council of Psychology (Conselho Federal de Psicologia [CFP], 2003), a test can only be used by Brazilian psychologists when there is sufficient evidence of reliability and validity to support the interpretation and use of its scores. Tests approved by a national commission of experts are listed on the CFP website (Sistema de Avaliação de Testes Psicológicos, 2019). These regulations have significantly improved the quality of psychological tests in Brazil and have impacted professional's use of evidence-based assessment practices (Reppold & Noronha, 2018).

At present, there are only two behavior rating scales on the list of approved tests, the *Inventário de Habilidades Sociais- Inventory of Social Abilities* (Del Prette Z.A.P. & Dell Prette, 2001) and the *Escala de Transtorno de Deficit de Atenção/Hiperatividade- Scale for Assessing Hyperactivity and Attention Deficiencies* (Benczik, 2003). The former is a social skills inventory that can be completed by teachers, parents, and/or students while the latter is a measure of attention and hyperactivity completed by teachers. Notably, the norms for both rating scales are presently outdated per the CFP's (2018) regulation that test norms be updated at

least once every 15 years. Bordin et al. (2013) developed an adapted version of the Child Behavior Checklist for use in Brazil, but at present this rating scale has not been approved by the CFP.

Fortunately, the rating scales available for use in Brazil are designed to collect information from multiple informants (e.g., students, teachers, and parents), as this allows practitioners to gain multiple perspectives on behavior. As mentioned earlier, there can be issues regarding the lack of agreement between raters. Likely causes of disagreement between raters include systematic differences in: (a) what is observed, (b) access to information other than observations of performance, (c) expertise in interpreting what is observed, and (d) evaluating what is observed (Achenbach, McConaughy, & Howell, 1987). When using information from multiple informants (i.e. parents, teachers, and student) it is important to examine the data for informant discrepancies. However, it remains beneficial to administer assessment tools to multiple informants to gain valuable information on the student in varying settings (i.e. within school and outside of school) that can ultimately assist in student placement and treatment planning (De Los Reyes et al., 2015).

## Assessment of Behavior in School Settings

The purpose of assessing students' behavior is to obtain information that can be used to understand and improve their behavioral functioning and ultimately improve outcomes (e.g., academic achievement, health and well-being, employment). Assessment of activities (i.e., the ability to perform tasks) and assessment of participatory behaviors (i.e., the use of tasks to be engaged and productive in important life domains) are useful for clarifying the extent to which students display behaviors that are desirable and adaptive. Assessment also can focus on problematic behavior. The extent to which the presence or absence of behavior is viewed as problematic depends largely on societal expectations (Deno & Mirkin, 1977). In school settings, problematic behavior can be defined as those behaviors that are discrepant from expectations held by authority figures who are responsible for the management of students' behavior. Problems may occur at the individual level or group level. A student whose behavior is markedly different from the normative expectations for peers is likely to be viewed as manifesting a problem at the individual level. Sometimes what is viewed as a problem that is limited to individual students is actually a group level concern. For example, research has established that the quality of instruction and classroom management affects student engagement and learning outcomes (e.g., Brophy & Good, 1986; Kounin, 1970). While people constantly engage in the informal observation of others, the process can be systematized to improve its objectivity and usefulness. Systematic assessment processes that have been developed and implemented in schools include universal screening, diagnostic and eligibility assessments, and problem-solving models.

## Universal screening

Universal screening targets behaviors or *warning signs* associated with mental health problems and academic difficulties (O'Connell, Boat, & Warner, 2009). Thus, screening is used to provide information about behavioral and emotional risk, and integral targets for measurement include at least the following: externalizing behaviors, internalizing behaviors, inattention/school problems, and adaptive/prosocial behavior (Kamphaus, 2012). The aim of screening is early identification, or more specifically, to differentiate between students who do and do not have problems that warrant prevention/early intervention services. The accuracy with which these distinctions can be made are dependent on the standard set for determining what is indicative of a problem, then setting a criterion (e.g., a cut score) to maximize the identification of students truly in need of services and minimize the identification of students who are not in need of services.

The universal screening process typically involves use of indirect ratings of behavior or referrals (e.g., teacher referrals, office discipline referrals) rather than data obtained from the direct observation of behavior. Important factors to consider when developing a screening instrument or selecting an existing screening

instrument include cost, administrative ease, social acceptability, appropriateness for universal use across the entire school population, theoretical support, technical adequacy, predictive validity, classification accuracy, usefulness for guiding intervention, and evidence of improved student outcomes (Glover & Albers, 2007).

## Diagnostic and eligibility assessment

In educational settings, assessment aimed at identifying disabilities and determining eligibility for programs that assist those with special needs is commonly referred to as psychoeducational. Such assessments typically involve the collection and use of multiple types of data, including data regarding traits that are not directly observable (e.g., intelligence, personality). Diagnostic and eligibility assessment often includes indirect ratings of behavior provided by collateral informants such as teachers and parents. Data from direct observations are often used to "...provide evidence of either the level of impairment or the degree of discrepancy from peer performance" (Briesch, Volpe, & Floyd, 2018, p.10). In addition to informing diagnosis and eligibility decisions, information obtained during the assessment process is used to develop recommendations regarding instructional planning, the selection of interventions, and use of accommodations.

## Problem-solving models

Behavioral assessment data can be used to facilitate decision-making with problem-solving models. For example, behavioral assessment can be used to identify the specific behaviors displayed by students and determine if these behaviors differ markedly from that of peers and are associated with deleterious consequences such as classroom disruptions, disengagement from classroom instruction, and disciplinary actions. Thus, these data are useful for clarifying the nature of concerns expressed by parents, teachers, or others who interact with students, resulting in the definition of a target behavior that is operationalized to enable reliable coding and measurement. When a problem is confirmed, behavioral assessment can be used to generate hypotheses about environmental factors that influence the target behavior. Intervention can then be planned and implemented, which typically includes the manipulation of antecedents and consequences that are functionally related to the target behavior. Formative assessment involves frequent, ongoing collection of data to monitor progress and guide the intervention process. Finally, summative assessment is completed at the end of an intervention to summarize outcomes and determine if the discrepancy between behavioral expectations and behavioral reality has been mitigated.

## General Challenges to the Assessment of Behavior

There are several technical and practical challenges that impact behavioral assessment. First, there are limitations associated with both direct and indirect methods of assessment. Direct behavioral observations have limited generalizability. The scope of time and settings in which observations occur is limited by practical constraints. Consequently, it is important to take an ecological approach to the assessment of student behavior (Dowdy et al., 2013). Thus, student behavior must be considered across multiple occasions and multiple settings. Problems that occur only in a specific setting may result from class-wide issues or may simply result from random error. While indirect information can be obtained regarding behavior displayed over several months and across a wide range of settings, little insight is provided regarding the antecedents and consequences that maintain behaviors. Given these limitations, it is often necessary to utilize multiple behavioral assessment techniques (Chafouleas, Volpe, Gresham, & Cook, 2010).

Second, as direct behavioral observation is time and labor intensive, it does not lend itself to the frequent and repeated measurements needed to monitor progress during formative assessment (Briesch et al.,

2018). Progress monitoring for academic interventions has been facilitated by the development of general outcome measures. For example, oral reading fluency is an example of a general outcome measure commonly used to facilitate progress monitoring. Unfortunately, it has been difficult to identify a general outcome measure that is sensitive to the wide range of mental, emotional, and behavioral issues that may be considered problematic (Chafouleas et al., 2010). Direct Behavior Rating (DBR) is a hybrid method that shows promise as a progress monitoring tool (Chafouleas, Christ, & Riley-Tillman, 2009). It is considered a hybrid model because, like the direct observation method, data are collected during specified observation periods under specific conditions and, like indirect assessment, data consist of ratings provided by collateral informants (e.g., teachers). DBR data consist of single item ratings reflecting broad classes of behavior (e.g., academic engagement, disruptive behavior). Research suggests that DBR are sensitive to change in academic engagement and disruptive behavior (Chafouleas, Sanetti, Kilgus, & Maggin, 2012; Smith, Eklund, & Kilgus, 2018).

Finally, behavioral assessment has tended to focus on problem behaviors (Dowdy et al., 2013). Assessment of academic engagement and other classroom variables is needed to gauge the quality of instruction and differentiate classroom or even school-wide issues from issues occurring at the level of individual students. Additionally, comprehensive understanding of behavioral functioning and general well-being requires assessment of behavioral strengths and weaknesses (Huebner & Gilman, 2003). In addition to improving understanding, information regarding strengths may improve both the acceptability (Walrath, Mandell, Holden, & Santiago, 2004) and effectiveness (Jimerson, Sharkey, Nyborg, & Furlong, 2004) of behavioral interventions.

## Future Directions

Behavioral assessment can offer important information about students' psychological problems and can be used to identify students in need of intervention. Direct observation measures require trained personnel, and considerable training is often needed to obtain ratings with adequate inter-rater reliability. Given these limitations, there is a need for rating scales that can be used to gather indirect information from multiple informants. Few existing measures have sufficient evidence of reliability and validity to support their use in Brazilian schools, and those that do have outdated norms. The process of examining reliability and validity evidence is largely funded by grants obtained by researchers, and this information is difficult to collect if researchers are unable to obtain grants to support this research. This issue tends to constrain the number of tests that are approved by the CFP.

The psychological assessment movement in Brazil has grown considerably in the last decade. This growth has had a positive impact in many respects, including improvements in the professional preparation of psychologists' and the availability of more instruments with a sound scientific basis (Wechsler et al., 2014). Certainly, these conditions have helped to improve behavioral assessment in school settings. Additionally, there has been a small increase in the number of school psychologists working in public schools at the municipal level (Guzzo, Mezzalira, & Moreira, 2012). The increase in school psychologists provides additional screening opportunities for children with emotional difficulties.

In conclusion, there is an urgent need for mental health and educational professionals to identify psychological difficulties earlier in children's lives. Better conditions for early detection, as well as for behavioral assessment in general, seemingly are emerging in Brazilian schools. However, the use of behavioral assessment in Brazilian schools must be increased to help translate these favorable conditions into earlier identification and earlier intervention for psychological and social difficulties.



## Contributors

N. F. BENSON Responsible for conception of manuscript, development of manuscript, critical revision.  
S. M. WECHSLER Responsible for manuscript review and critique. B. S. PARKER Responsible for literature review regarding the manuscript subject.

## References

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213-232.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91-97.
- Barretto, E. S. S. (2015). Políticas de formação docente para a educação básica no Brasil: embates contemporâneos. *Revista Brasileira de Educação*, *20*(62), 679-701. <http://dx.doi.org/10.1590/s1413-24782015206207>
- Benson, N. F. (2010). Types of tests and assessments. In E. Mpofu & T. Oakland (Eds.), *Assessment in rehabilitation and health* (pp.72-90). Boston: Allyn & Bacon.
- Benson, N. F., Floyd, R. G., Kranzler, J. H., Eckert, T. L., Fefer, S. A., & Morgan, G. B. (2019). Test use and assessment practices of school psychologists in the United States: Findings from the 2017 national survey. *Journal of School Psychology*, *72*, 29-48.
- Benczik, E. B. P. (2003). *Escala de Transtorno de Deficit de Atenção/Hiperatividade*. São Paulo: Casa do Psicólogo.
- Bordin, I. A., Rocha, M. R., Paula, C. S., Teixeira, M. C. T. V., Achenbach, T. M., Rescorla, L. A., & Silveiras, E. F. M. (2013). Child Behavior Checklist (CBCL), Youth Self-Report (YSR) and Teacher's Report Form (TRF): An overview of the development of the original and Brazilian versions. *Cadernos Saúde Pública*, *29*(1), 13-28.
- Briesch, A. M., Volpe, R. J., & Floyd, R. G. (2018). *School-based observation: A practical guide to assessing student behavior*. New York: Guilford Press.
- Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp.328-375). New York: MacMillan.
- Chafouleas, S. M., Christ, T. J., & Riley-Tillman, T. C. (2009). Generalizability of scaling gradients on direct behavior ratings. *Educational and Psychological Measurement*, *69*(1), 157-173.
- Chafouleas, S.M., Riley-Tillman, T. C., & Sugai, G. (2007). *School-based behavioral assessment: Informing intervention and instruction*. New York: Guilford Press.
- Chafouleas, S. M., Sanetti, L. M. H., Kilgus, S. P., & Maggin, D. M. (2012). Evaluating sensitivity to behavioral change using direct behavior rating single-item scales. *Exceptional Children*, *78*(4), 491-505.
- Chafouleas, S. M., Volpe, R. J., Gresham, F. M., & Cook, C. R. (2010). School-based behavioral assessment within problem-solving models: Current status and future directions. *School Psychology Review*, *39*(3), 343-349.
- Cone, J. D. (1978). The behavioral assessment grid (BAG): A conceptual framework and a taxonomy. *Behavior Therapy*, *9*(5), 882-888.
- Conselho Federal de Psicologia (2003). *Resolução nº 002. Uso, elaboração e comercialização de testes psicológicos*. Brasília, DF: Conselho Federal de Psicologia.
- Conselho Federal de Psicologia (2018). *Resolução n.009. Diretrizes da avaliação psicológica e regulamentação Sistema de Avaliação de Testes Psicológicos-SATEPSI*. Brasília, DF: Conselho Federal de Psicologia.
- Del Prette, Z. A. P. & Del Prette, A. (2001). *Inventário de Habilidades Sociais - IHS: Manual de aplicação, apuração e interpretação*. São Paulo: Casa do Psicólogo.
- Deno, S. L. & Mirkin, P. K. (1977). *Data-based problem modification: A manual*. Reston: Council for Exceptional Children.
- De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers D. E., Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858-900.
- Dowdy, E., Twyford, J., & Sharkey, J. D. (2013). Methods of assessing behavior: Observations and rating scales. In D. H. Saklofske, C. R. Reynolds, V. L. Schwann, D. H. Saklofske, C. R. Reynolds, & V. L. Schwann (Eds.), *The Oxford handbook of child psychological assessment* (pp.623-650). New York: Oxford University Press.

- Fonseca, T. D., Freitas, C. S., & Negreiros, F. (2018). Psicologia escolar e educação inclusiva: A atuação junto aos professores. *Revista Brasileira de Educação Especial*, 24(3), 427-440. <http://dx.doi.org/10.1590/s1413-65382418000300008>
- Gatti, B. A. (2014). Formação inicial de professores para a educação básica: pesquisas e políticas educacionais. *Estudos em Avaliação Educacional*, 25(57), 24-54. <http://dx.doi.org/10.18222/eaee255720142823>
- Glover, T. A. & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology*, 45(2), 117-135.
- Guzzo, R. S. L., Mezzalana, A. S. C., & Moreira, A. P. G. (2012). Psicólogo na rede pública de educação: embates dentro e fora da própria profissão. *Psicologia Escolar e Educacional*, 16(2), 329-338.
- Huebner, E. S. & Gilman, R. (2003). Toward a focus on positive psychology in school psychology. *School Psychology Quarterly*, 18(2), 99-102.
- Jimerson, S. R., Sharkey, J. D., Nyborg, V., & Furlong, M. J. (2004). Strength-Based Assessment and School Psychology: A Summary and Synthesis. *California School Psychologist*, 9(1), 9-19.
- Johnston J. M., Foxx R. M., Jacobson J. W., Green G., & Mulick J. A. (2006). Positive behavior support and applied behavior analysis. *The Behavior Analyst*, 29(1), 51-74.
- Kamphaus, R. W. (2012). Screening for behavioral and emotional risk: Constructs and practicalities. *School Psychology Forum*, 6(4), 89-97.
- Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., ... Rahman, A. (2011). Child and adolescent mental health worldwide: Evidence for action. *Lancet*, 378(9801), 1515-1525. [http://dx.doi.org/10.1016/S0140-6736\(11\)60827-1](http://dx.doi.org/10.1016/S0140-6736(11)60827-1).
- Kounin, J. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehard & Winston.
- Marcondes, M. I. (1999). Teacher education in Brazil. *Journal of Education for Teaching*, 25(3), 203-213. <http://dx.doi.org/10.1080/02607479919493>
- Marcondes, M. I., Leite, V. F., & Ramos, R. K. (2017). Theory, practice and research in initial teacher education in Brazil: Challenges and alternatives. *European Journal of Teacher Education*, 40(3), 326-341. <http://dx.doi.org/10.1080/02619768.2017.1320389>
- Matijasevich, A., Murray, E., Stein, A., Anselmi, L., Menezes, A. M., S., Santos A., ... Victora, C. G. (2014). Increase in child behavior problems among urban Brazilian 4-year olds: 1993 and 2004 Pelotas birth cohorts. *Journal of Child Psychology and Psychiatry*, 55(1), 1125-1134. <http://dx.doi.org/10.1111/jcpp.12236>
- Mendes, M. S. S. (2010). Qualidade de ensino na escola pública: desafios e (im)possibilidades. *Psicologia: Ensino e Formação*, 1(2), 61-71
- Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children and adolescents*. Mahwah: Lawrence Erlbaum Associates.
- Oakland, T. & Wechsler, S. (1990). School psychology in Brazil: An examination of its research infrastructure. *School Psychology International*, 11(4), 287-293.
- O'Connell, M. E., Boat, T., & Warner, K. E. (Eds.). (2009). *Prevention mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academies Press.
- Organisation for Economic Co-operation and Development (2015). *Education policy outlook: Brazil*. Slovak Republic: OECD.
- Reppold, C.T., & Noronha, A. P. (2018). Impacto dos 15 anos do SATEPSI na avaliação psicológica brasileira. *Psicologia, Ciência e Profissão*, 38,6-15. <http://dx.doi.org/10.1590/1982-3703000208638>
- Reynolds, C. R. & Kamphaus, R. W. (2015). *Behavioral assessment system for children* (3rd ed.). New York: Pearson.
- Reynolds, C. R. & Richmond, B. O. (2008) *Revised Children's Manifest Anxiety Scale* (2nd ed.). Los Angeles: Western Psychological Services.
- Sant'Ana, I. M., Euzébio Filho, A., Lacerda Júnior, F., & Guzzo, R. S. L. (2009). Psicólogo e escola: a compreensão de estudantes do ensino fundamental sobre esta relação. *Psicologia Escolar e Educacional*, 13(1), 29-36. Recuperado março 3, 2019, de [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1413-85572009000100004&lng=pt&tlng=](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1413-85572009000100004&lng=pt&tlng=)
- Salvia, J., Ysseldyke, J. E., & Witmer, S. (2009). *Assessment: In special and inclusive education* (11th ed.). Belmont: Wadsworth.
- Sistema de Avaliação de Testes Psicológicos. (2019). Retrieved April 4, 2019, from <http://satepsi.cfp.org.br>.
- Shapiro, E. S. & Browder, D. M. (1990). Behavioral assessment. In J. L. Matson (Ed.), *Handbook of behavior modification with the mentally retarded* (2nd ed., pp. 93-122). New York: Plenum Press.

- Shapiro, E. S. & Kratochwill, T. R. (Eds.) (2000). *Behavioral assessment in schools: Conceptual foundations and practical applications* (2nd ed.). New York: Guilford Press.
- Skinner, B. F. (1953). *Science and human behavior*. New York: The MacMillan Company.
- Smith, R. L., Eklund, K., & Kilgus, S. P. (2018). Concurrent validity and sensitivity to change of Direct Behavior Rating Single-Item Scales (DBR-SIS) within an elementary sample. *School Psychology Quarterly*, 33(1), 83-93.
- Sulzer-Azaroff, B. & Mayer, G. R. (1991). *Behavior analysis for lasting change*. Orlando: Harcourt Brace Jovanovich.
- Valle, L. E. L. R. (2003). Psicologia escolar: um duplo desafio. *Psicologia Ciência e Profissão*, 23(1), 22-29. <http://dx.doi.org/10.1590/S1414-98932003000100004>
- Walrath, C. M., Mandell, D. S., Holden, E. W., & Santiago, R. L. (2004). Assessing the strengths of children referred for community-based mental health services. *Mental Health Services Research*, 6(1), 1-8.
- Wechsler, S. M., Oakland, T., Leon, C., Vivas, E., Almeida, L., Franco, A., ... Contini, N. (2014). Test development and use in five Iberian Latin American Countries. *International Journal of Psychology*, 49(4), 233-239. <http://dx.doi.org/10.1002/ijop.12068>
- Witt, J. C., Heffer, R. W., & Pfeiffer, J. (1990). Structured rating scales: A review of self-report and informant rating processes, procedures, and issues. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Personality, behavior, and context* (pp.364-394). New York: Guilford Press.

Received: February 12, 2019

Approved: March 18, 2019