

## Fidelity measurement in school-based mental health interventions: a systematic review\*

Douglas Garcia<sup>1</sup>

 <https://orcid.org/0000-0001-5245-6075>

Daniela Ribeiro Schneider<sup>2</sup>

 <https://orcid.org/0000-0002-2936-6503>

Roberto Moraes Cruz<sup>2</sup>

 <https://orcid.org/0000-0003-4671-3498>

**Aims:** characterize the measurement of fidelity of mental health interventions for children in primary schools. Data were collected at the ERIC, LILACS, APA, PubMed, Scopus, SciELO and Web of Science databases. We included 45 empirical articles, published between 2007-2016, which were analyzed in relation to categories defined previously. The results indicate variations in the definition, dimensionality and form of fidelity measurement, with few indicators of validity and accuracy of the instruments, which may bias the evaluation of the implementation process and the internal validity of the results of the interventions.

**Descriptors:** Guideline Adherence; Mental Health; Program Evaluation; Psychometrics.

\* Paper extracted from master's thesis "Implementation fidelity evaluation of a school-based preventive program in children's mental health", presented to Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001.

<sup>1</sup> Prefeitura Municipal de Palhoça, Centro de Referência de Assistência Social, Palhoça, RS, Brazil

<sup>2</sup> Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.

### How to cite this article

Garcia D, Schneider DR, Cruz RM. Fidelity measurement in school-based mental health interventions: a systematic review. SMAD, Rev Eletrônica Saúde Mental Álcool Drog. 2019;15(3):1-9. doi: <https://dx.doi.org/10.11606/issn.1806-6976.smad.2019.150238>

## **Mensuração da fidelidade de intervenções em saúde mental baseadas na escola: uma revisão sistemática**

Objetivo: caracterizar a mensuração da fidelidade de intervenções em saúde mental em crianças de escolas do ensino fundamental primário. Os dados foram levantados nas bases de dados ERIC, LILACS, APA, PubMed, Scopus, SciELO e Web of Science. Foram incluídos 45 artigos empíricos, publicados entre 2007 e 2016, os quais foram analisados em relação a categorias definidas previamente. Os resultados indicam variações na definição, dimensionalidade e forma de mensuração da fidelidade, havendo poucos indicadores de validade e precisão dos instrumentos, o que pode enviesar a avaliação do processo de implementação e a validade interna dos resultados das intervenções.

Descritores: Fidelidade a Diretrizes; Saúde Mental; Avaliação de Programas e Projetos de Saúde; Psicometria.

## **Medición de la fidelidad de intervenciones en salud mental basadas en la escuela: una revisión sistemática**

Objetivo: caracterizar la medición de la fidelidad de intervenciones en salud mental en niños de escuelas de enseñanza primaria primaria. Los datos fueron recogidos en las bases de datos ERIC, LILACS, APA, PubMed, Scopus, SciELO y Web of Science. Se incluyeron 45 artículos empíricos, publicados entre 2007 y 2016, los cuales fueron analizados en relación a categorías definidas previamente. Los resultados indican variaciones en la definición, dimensionalidad y forma de medición de la fidelidad, habiendo pocos indicadores de validez y precisión de los instrumentos, lo que puede enviar la evaluación del proceso de implementación y la validez interna de los resultados de las intervenciones.

Descriptorios: Adhesión a Directriz; Salud Mental; Evaluación de Programas y Proyectos de Salud; Psicometría.

## Introduction

Fidelity is a construct defined as the degree to which a planned intervention is implemented<sup>(1-2)</sup>. The construct is part of a intervention implementation process, which allows understanding *how* and *why* interventions work or stop working<sup>(1-3)</sup>. It is important to measure this construct in health care programs and interventions as its estimations permit assessing if there was any intended change in a specific outcome that can be credited to the implemented intervention<sup>(3-4)</sup>. Hence, measuring fidelity is significant in order to assess the intervention results validity<sup>(3-4)</sup>.

The problem emerges from the fidelity measurement operationalization, since it is possible to notice differences between the constituent definitions of construct, between the construct dimensions, and the scarcity of information regarding the psychometric properties in the tools used for the assessment.

From analyzing the term, we can observe the fidelity construct is also referred to in literature as fidelity of implementation<sup>(5)</sup>, integrity<sup>(1)</sup>, adherence<sup>(6)</sup>, among others. Despite the variety of terms, the definitions used to conceive the construct tend to resemble each other. We can observe this in the concepts of fidelity and integrity<sup>(7)</sup> and integrity<sup>(1)</sup>, which are, respectively, defined as "conformity to predicted elements and absence of unpredicted ones", and degree to which programs are implemented accordingly to what was planned". However, it is essential to take notice of possible dispersions in the definitions used for the constructs. The definition for adherence, for instance, albeit similar, is sometimes assessed as only one of the dimensions of fidelity<sup>(1)</sup>; whereas other times as an analogue construct<sup>(8)</sup>.

Fidelity dimensionality is another aspect of measurement that can present variations. According to Dane & Schneider<sup>(1)</sup> - often cited as reference in articles that measure the construct - fidelity is composed by five dimensions: adherence, degree to which the appliers follow the program method, and complement what is prescribed in the manual; application quality, the skill and comprehension degree of the appliers acting in the intervention; commitment or responsiveness, degree to which the participants are properly engaged in the tasks; and differentiation, degree in which the intervention meets distinct immediate and intermediate results.

The measurement of fidelity in research may, however, vary from the evaluation of only one dimension, e.g. adherence, or more than one<sup>(8-9)</sup>. We emphasize the multidimensional model with five dimensions is

founded in content validity, not having factorial validity. Therefore, the measured dimensionality in the studies is conditioned to the theoretical model of fidelity by the research author and focus<sup>(8-10)</sup>.

The psychometric features of the tools used to measure fidelity are not often shown in the research, which hinders the validity evaluation and the accuracy of the construct under assessment<sup>(11-12)</sup>. Furthermore, self-report procedures, regularly used to measure fidelity, tend to compromise the measurement accuracy<sup>(11)</sup>. Actually, the dimensions used are selected more for research methodological choice than for an underlying theoretical structure in the measurement process<sup>(11)</sup>.

Despite the measurement limitations, fidelity is a construct generally utilized by reason of permitting empirical validation inference, theoretical models subjacent to the intervention<sup>(13-14)</sup>. This construct also allows putting the intervention to test, as in effectiveness research. In such context, the fidelity level serves as an indicator whether or not an intervention worked - depending on the outcomes evaluated. Moreover, in intervention effectiveness research, the evaluation of fidelity permits pinpointing which the main intervention components are, i.e. in practical terms which ones are responsible for the change<sup>(3)</sup>.

In the scope of Brazilian children and teenager mental health, the evaluation of intervention implementation process became important mostly with the *Programa Elos* for drug use prevention in public schools<sup>(15-18)</sup>. This program is an adaptation of the Good Behavior Game, a classroom management program for elementary school children that has been effective in other countries, and is now being assessed for efficacy and effectiveness in Brazil<sup>(19-21)</sup>. In this sense, fidelity is a construct that is part of the implementation process, and also important to understand the result that programs such as *Programa Elos* can show.

Given the problem delineated around fidelity measurement, this systematic review has, as goal, to portray this construct measurement in interventions focused on mental health in elementary schools for the last 10 years. Due to this objective, we aimed at: characterizing methodological aspects in the evaluation process and fidelity construct measurement; and characterizing the fidelity construct in its conceptual matrix, dimensionality, and function in research scope.

## Method

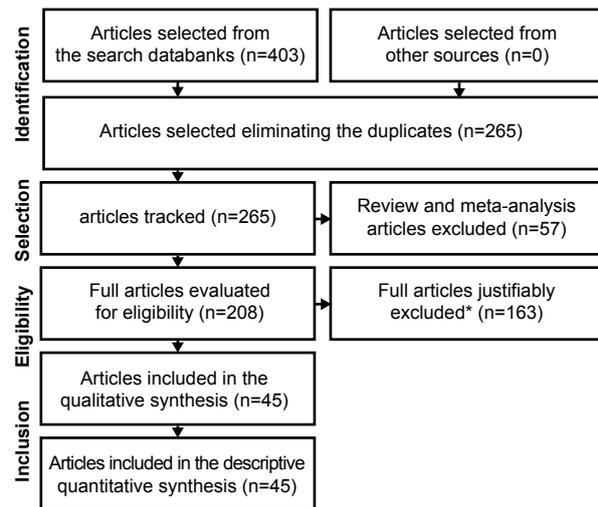
In order to systematically review the literature concerning fidelity measurement, we considered the

PRISMA protocol indicators<sup>(22)</sup>. As for data survey, we opted for the ERIC, LILACS, APA, PubMed, Scopus, SciELO, and Web of Science bases. These were selected for presenting relevant indexes for health care and education, with both national and international outreach. We also used the following equation in our data survey: *(Program OR intervention OR prevention) AND (school-based) AND (fidelity) AND ("program implementation" OR "program evaluation" OR "integrity" OR "adherence" OR "curriculum implementation")*. As to refine the results, we filtered to just find the results in the articles titles and abstracts, also using it in Portuguese and Spanish languages.

Among the article eligibility criteria, we defined: articles in Spanish, English or Portuguese; articles published between 2007 and 2016; and empirical articles on the evaluation of implementation fidelity interventions regarding mental health of only children between the ages of 3 and 12. We excluded review articles, meta-analyses, and duplicate articles. After defining these criteria, we read the abstracts and surveyed the studies that were suitable to the research objectives. The flowchart in Figure 1 demonstrates the selection process results, which led to a sample of 45 articles.

To conduct the data analysis, the articles were fully read for assessing the variables of interest, grouped in three categories: program characterization - program name, number of sessions, types of implementer, and target audience); research methodological outlining (type of outlining and sample characterization); and fidelity measurement characteristics (types and number of instruments, their psychometric parameters, measurement/scale characterization, constitutive definition characterization and construct operational definition, and fidelity measurement purpose in the research protocol).

The three categories we used for data analysis, and their respective constituent variables were presumptively defined to delimit the addressed mental health context and characterize the fidelity measurement object of study compared to the identified state of the art<sup>(2)</sup>. Thus, the variable and category delimitation resulted from the need to fulfill the review objective, and were also based on the data found in the articles. After this variable delimitation, the data were extracted from the analyzed articles and we conducted an occurrence distribution of each variable in each study; to find tendencies that would more suitably fulfill the objectives. For this analysis operationalization, we used spreadsheets.



\*The criteria that disqualified the 163 articles were: not being interventions fully based in schools; not measuring fidelity (n = 4); being only protocols for research to be applied in the future (n = 2); intervention aimed at adolescents and families (n = 80); interventions without an outcome regarding mental health (n = 69); and language (n = 5)

Figura 1 - Flowchart of the characterization of our article screening for our systematic review

## Results

### Characteristics of the programs and interventions in the publications

From the chosen 45 articles, 38 were identified as distinct programs, predominantly applied in the United States (38 studies) and Europe (6 studies). Table 1 presents a characterization of the methodological outlinings in the studies.

Table 1 - Characterization of the methodological outlinings in the 45 studies included in the review. Florianópolis, SC, Brazil, 2017

Description of the observed outlining aspect	Number of studies
Type of method used in the study	
Experimental and quasi-experimental with control group	20
Experimental and quasi-experimental without control group	5
Descriptive and correlational studies	20
More common program outcomes	
Bullying and violence	10
Autism Spectrum Disorder	7
Disruptive Behaviors	7
Number of sessions during interventions*	
4-20 sessions	6
11-20 sessions	14
21-169 sessions	7
Type of research participants involved in the studies†	
Children	33
Teachers	36
Other professionals	9
Parents/Custodians or Guardians	6

(to be continued...)

Table 1 – continuation

Description of the observed outlining aspect	Number of studies
Sample of children	
3 - 200 children	17
201 - 1000 children	10
1001 - 7413 children	6
Sample of program implementers	
3 – 100 implementers	23
101 - 200 implementers	6
201 - 2022 implementers	7
Target audience†	
Children from 3 to 4 years old	4
5 year old children	10
Children from 6 to 12 years old	42
<b>Total</b>	<b>45</b>

\*Some interventions have their duration represented in months or years, making it impossible to determine the number of sessions. † Some studies involved more than one category, hence the total sum being bigger than the total number of articles in the review

Teachers were the main responsible for program implementation. However, other school professionals, such as principals, psychologists, nurses, counselors, and doctors were identified as implementers in 9 studies.

#### **Fidelity construct measurement characteristics**

The constituent definition of the fidelity construct, which regards the phenomenon theoretical definition was identified in only 28 articles. By identifying this definition, we observed the references used by the article authors to define the fidelity construct, in order to characterize a conceptual matrix able to explore the predominant theoretical perspective. Figure 1 presents a synthesis of this analysis, resulting in the conceptual matrix, which displays the occurrence and co-occurrence of each frequency, alongside with the studies that showed the constituent definition for fidelity<sup>(1,3,5,23-29)</sup>.

According to the conceptual matrix, it is possible to note the the fidelity construct characterization presents a more frequent theoretical ground in three references, namely: Dusenbury, Brannigan, Falco and Hansen<sup>(3)</sup> in 10 articles; Durlak and Dupre<sup>(28)</sup> in 6 articles, and Dane and Schneider in 12 articles. Dane and Schneider, through a review of literature with studies published between 1980 and 1994 on preventive programs, investigated fidelity in 162 articles, and found that in only 24% showed fidelity-specific evaluation procedures; demonstrating that, despite the fragility in the definitions applied, a five-dimension view - adherence, exposure, implementation quality, responsiveness, and differentiation - was prevalent.

Dusenbury, Brannigan, Falco e Hansen<sup>(3)</sup>, who performed a systematized review, covering 25 years of drug use prevention programs, also admit the five-dimension model proposed by Dane and Schneider<sup>(1)</sup>,

fostering in demonstrating the main aspects that affect the fidelity implementation, such as: the program characteristics; the received training characteristics; the implementer characteristics; and organizational characteristics.

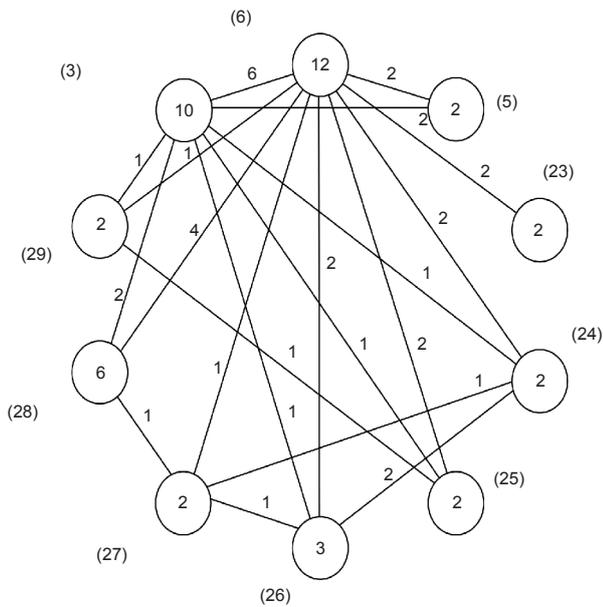
Durlak e Dupre<sup>(28)</sup> - in a more comprehensive review, aimed at identifying the implementation process impact on the evaluated outcome, as well as the intervening variables on this outcome - approach fidelity as a central construct, and define it using Dane and Schneider's theoretical basis<sup>(1)</sup>. The increase in evidence shown by Durlak and Dupre<sup>(28)</sup> is in determining the variables that affect implementation. Therefore, under this perspective it is observed that the conceptual matrix is superposed in terms of dimensions, which come from reviews that have focused on synthesizing the fidelity approach over the years.

From the 45 articles, only 13 measured fidelity in more than one dimension: three articles with two dimensions; six with three dimensions; two with four dimensions; one with seven dimensions; and another with eight dimensions. None of the remaining articles presented the utilized fidelity measurement dimensionality. The most prevalent dimensions associated to fidelity were: adherence (10 studies); exposure or dose (5 studies); implementation quality (4 studies); and responsiveness or engagement (4 studies).

Besides the most prevailing dimensions, other less common ones were also used in th studies, such as: defined expectations; instructed behavior expectations; reward system for the behavioral expectations; reply system to behavioral violations; tracking and assessment; and district level management and support (Figure 2).

The operational definitions, or the tools used to measure fidelity were fully shown in three studies; whereas in other 17, only some of the items were presented. The displayed items were in conformity with some of the fidelity construct dimensions, and in compliance with the conceptual dimensionality considered by the authors; as well as regarding the proposed research interests and outlining.

In 35 studies, the fidelity measurement occurred through only one tool; as for the remaining 10, two or more tools were applied. Questionnaires/inventories were predominant in 22 articles, and in 17, checklists prevailed. The measurement levels varied between dichotomous scales and rating scales, from 3 to 10 points. In 14 studies, the data was collected via implementer self-report, in 22, through external researcher report, and in 9, multiple report, associating measures gathered both by the implementer and by the researcher.



The numbers in parentheses are the references. The figures in each circle are the number of times that the reference was used in the constitutive definition of fidelity. The numbers next to the lines refer to the amount of co-occurrence of two authors connected by the line as references used in defining the construct

Figure 2 - Concept map of the references used to define the fidelity construct in the studies articles

The psychometric reliability and validity indicators were identified in studies that showed these parameters in one or more instruments. In the reliability indicators, we identified the Cronbach's alpha in 14 articles, inter-rater reliability in 11 articles, and kappa coefficient in two articles. As for the validity indicators, they were: construct validity via exploratory or confirmatory factor analysis, in 4 articles; criterion validity in two articles, and content validity, in one article. From an amount of 15 articles that referenced the fidelity measurement tool, five did not present the psychometric parameters - possibly due to such information being already in the tool reference.

Lastly, the fidelity construct measurement finality in the research varied in three approaches: fidelity as a result validation resource produced by the intervention on the outcome of interest (22 articles); fidelity as the main study object to investigate intervening variables in its variation (14 articles); and a mix that related fidelity as a validating outcome, as well as assessing other variables mediating or moderating relation over fidelity (9 articles).

Albeit the evidences we showed in this review, it is pertinent to highlight some limitations that might have influenced the results in the study, such as the article selection bias stated in the eligibility criteria delimitation, and the instrumental focus on the fidelity construct. As for what regards the study selection bias, the choice for articles within the mental health field restricted to school-based interventions with children was such

owing to the hypothesis of existing idiosyncrasies and similarities in the study objects in these contexts; which could favor the analyses and comparisons in a more homogeneous sample.

By reason of the aforementioned limitation and of the expansion of the presented discussions, we note that propositions from other similar research protocols that encompass other target audiences and interventions would likely bring pertinent contributions, in an enlarged mental health panorama. In what refers to the limitation regarding the focus of measuring fidelity in a mental health context, we regard that the contribution was predominantly concerning the critical evaluation of fidelity psychometric instrumentalization assessment. However, this condition implications gap still prevails, in an expanded mental health field - an environment in which other studies can be prompted.

## Discussion

The diversity of definitions on the fidelity construct, associated with the lack of consensus in the characterization of implementation variables - as dimensions, or as control variables in research outlining - present important development in measurement operationalization and, consequently, to program evaluation.

The operationalization in fidelity measurement among the studied articles is limited by some aspects, such as the use of constituent definitions based on strictly theoretical models, the dimensionality delimitation in conformity with the research objective, and the lack of validity and accuracy psychometric criteria in most tools.

The theoretical models of Dane and Schneider<sup>(1)</sup> and Dusenbury, Brannigan, Falco and Hansen<sup>(3)</sup>, the most frequent references to constitutively define fidelity, show dimensional overlapping. However, these were formed from literature or systematic reviews, meaning that this model validity is, strictly speaking, apparent or of content, since there was no empirical evidence of these dimensions through a construct validation in the articles.

Studies that did not present the constituent definitions corroborate the fidelity measurement fragilization, since even its content validity can be unknown; what compromises not only the measurement, but also the intervention result validity<sup>(11)</sup>. In this point, it is important to emphasize the existence of some efforts to standardize implementation-related constructs, e.g. fidelity, through the elaboration of glossaries for standard terms<sup>(11,30-32)</sup>. Nonetheless, even with terminology standardization, validity measurement still demands a theoretical definition of the phenomenon;

producing operational definitions that delimit at least what kind of behaviors comprise fidelity, for which intervention, and which subject<sup>(33)</sup>.

The insufficient presentation of tool items or operational behaviors that track the fidelity construct in the studies assessed in this review is a hindering aspect in internal measurement consistency, apart from hampering the comprehension on the extension of the evaluated phenomenon. For both the item operationalization in the measurement theory<sup>(34)</sup> and fidelity measurement recommendations<sup>(33)</sup>, the objective behavior description delineated in a group of items regarding the conceptual definition of the construct proposal is an important condition to the theoretical and operational measurement consistency. With respect to this, few studies show the construct items or its operational definitions, which interferes in evaluating the measurement theoretical coherence.

The purpose of measuring fidelity within some research outlinings showed the construct as a methodological resource to ensure that the variations in the outcome of interest are due to the program, thus legitimizing the intervention internal validity. Nevertheless, the very internal validity may be assessed in a biased way, since the fidelity measurement that delimits the intervention-result causal relation does not present validity parameters for most studies. After all, what is the validity of a program that has its efficacy measured and mediated by a non-validated measurement? In this sense, there is little advance in empirical validation of intervention underlying theoretical models, which give ground to the developments in establishing evidence-based practices in each intervention.

In a background in which the school context can be a key aspect in failure of children-directed mental health interventions<sup>(35)</sup>, consistent psychometric fidelity measurement in efficacy studies allows not only identifying evidence-based practices, but also make feasible cultural adaptation and program dissemination in various settings<sup>(35-36)</sup>.

## Conclusion

The search for consensual and detailed descriptions on health care interventions is important for the advance of knowledge and technological development; which can face editorial limitations and commercial interests in what regards interventions.

The fidelity construct measurement lacks psychometric estimators, having important implications to scientific advance on the knowledge of central intervention elements, and its relations with the

intended changes. Large-scale implementation challenges - especially for interventions deemed as potential public policies - are broad and require delineation that enables testing the intervention, and proposing reverse engineering when suggesting changes based in delineations that consider psychometrically consistent fidelity indicators correlated to outcome measures. To achieve that, fidelity and other implementation process-related constructs need technological research advances.

The restriction to mental health interventions in primary education - defined according to the delineation - is a methodological limitation in this research, and may generate some bias in the characterization of fidelity measurement. To include fidelity construct measurement in other kinds of intervention, and in contexts beyond health care, can show other results - which would extend the possibilities of phenomenon estimate in research contexts.

## References

1. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin Psychol Rev.* 1998; 18:23-45.
2. Hansen WB. Measuring fidelity. In: Sloboda Z, Petras H. *Defining prevention science, advances in prevention science.* Nova York: Springer Science and Business Media; 2014. p. 335-59.
3. Dusenbury L, Brannigan R, Falco M, Hansen WB. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res.* 2003; 18:237-56.
4. Melo MHS, Rodrigues DRSR, Conceição MIG. Avaliação de programas de prevenção e promoção em saúde mental. In: Murta SG, Leandro-França C, Santos KB, Polejack L. *Prevenção e promoção em saúde mental: fundamentos, planejamento e estratégias de intervenção.* Novo Hamburgo: Sinopsys; 2015. p. 212-29.
5. Century J, Rudnick M, Freeman C. A framework for measuring fidelity of implementation: a foundation for shared language and accumulation of knowledge. *Am J Eval.* 2010; 31:1-17.
6. Gony EA, Sutherland KS, Farrell AD, Sullivan TN, Doyle ST. Measuring teacher implementation in delivery of a bullying prevention program: the impact of instructional and procedural adherence and competence on student responsiveness. *Prev Sci.* 2015; 16(3):440-50.
7. McGrew JH, Bond GR, Dietzen L, Salyers M. Measuring the fidelity of implementation of a mental health program model. *J Consul Clin Psychol.* 1994; 62(4):670-8.

8. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci.* 2007;2:40. doi: 10.1186/1748-5908-2-40
9. Pérez D, Van der Stuyft P, Zabala MC, Castro M, Lefèvre P. A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. *Implement Sci.* 2016; 91:1-11.
10. Santos KB, Murta SG. A implementação de programas de promoção e prevenção no âmbito da saúde mental. In: Murta SG, Leandro-França C, Santos KB, Polejack L. *Prevenção e promoção em saúde mental: fundamentos, planejamento e estratégias de intervenção.* Novo Hamburgo: Sinopsys; 2015. p. 192-211.
11. Martinez RG, Lewis CC, Weiner BJ. Instrumentation issues in implementation science. *Implement Sci.* 2014;49:118. doi: <https://doi.org/10.1186/s13012-014-0118-8>
12. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. *Implement Science.* 2013; 8:22. doi:<https://doi.org/10.1186/1748-5908-8-22>
13. Basch CE, Slipevich EM, Gold RS, Duncan DF, Kolbe LJ. Avoiding type III errors in health education program evaluations: a case study. *Health Educ Q.* 1985; 12(4):315-31.
14. Linnan L, Steckler A. Process evaluation for public health interventions and research: An overview. In: Steckler A, Linnan L. editors. *Process evaluation for public health interventions and research.* San Francisco (CA): Jossey-Bass; 2002. p. 1-21.
15. Horr JF. Avaliação da satisfação do processo de implantação do programa preventivo Unplugged na perspectiva dos educandos [dissertação]. Florianópolis: Universidade Federal de Santa Catarina; 2015. 187 p.
16. Strelow M. Avaliação da implementação de programa preventivo em saúde mental através da aceitabilidade de crianças participantes [dissertação]. Florianópolis: Universidade Federal de Santa Catarina; 2016. 157 p.
17. Lopes J. Avaliação do processo de implementação de programa de prevenção escolar do uso de drogas na percepção dos professores participantes [tese]. Florianópolis: Universidade Federal de Santa Catarina; 2016. 258 p.
19. Embry DD. The Good Behavior Game: a best practice candidate as a universal behavioral vaccine. *Clin Child Fam Psychol Rev.* 2002; 5:273-97.
20. Donaldson JM, Vollmer TR, Krous T, Downs S, Berad KP. An evaluation of the Good Behavior Game in kindergarten classrooms. *J Appl Behav Anal.* 2011; 44:605-9.
21. Leflot G, van Lier PAC, Onghena P, Colpin H. The role of children's on-task behavior in the prevention of aggressive behavior development and peer rejection: a randomized controlled study of the Good Behavior Game in Belgian elementary classrooms. *J Sch Psychol.* 2013; 51:187-99.
22. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA Statement; 2009. Disponível em: [www.prisma-statement.org](http://www.prisma-statement.org)
23. Lee CS, August GJ, Realmuto GM, Horowitz JL, Bloomquist ML, Klimes-Dougan B. Fidelity at a distance: assessing implementation fidelity of the Early Risers Prevention Program in a going-to-scale intervention trial. *Prev Sci.* 2008; 9:215-29.
24. Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: issues and research recommendations. *Clin Psychol.* 2005; 12:365-83.
25. Domitrovich CE, Greenberg MT. The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *J Educ Psychol Consult.* 2000; 11:193-221.
26. Gresham FM, Gansle KA, Noell GH, Cohen S, Rosenblum S. Treatment integrity of school-based behavioral intervention studies: 1980-1990. *School Psych Rev.* 1993; 22:254-72.
27. Waltz J, Addis ME, Koerner K, Jacobson NS. Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *J Consult Clin Psychol.* 1993; 61:620-30.
28. Durlak JA, Dupre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol.* 2008; 41:327-50.
29. Yeaton WH, Sechrest L. Critical dimensions in the choice and maintenance of successful treatments: strength, integrity, and effectiveness. *J Consult Clin Psychol.* 1981; 49:156-67.
30. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health.* 2011; 38(2):65-76.
31. Rabin BA, Brownson RC, Haire-Joshu D, Kreuter MW, Weaver NL. A glossary for dissemination and implementation research in health. *J Public Health Manag Pract.* 2008; 14(2):117-23.
32. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Adm Policy Ment Health.* 2011; 38(1):4-23.

33. Ledford JR, Wolery, M. Procedural fidelity: an analysis of measurement and reporting practices. *J Early Interv.* 2013; 35(2):173-93.
34. Pasquali L. Instrumentação psicológica: fundamentos e práticas. Porto Alegre: Artmed; 2010. 559 p.
35. Bosworth K. Defining prevention science in school settings, complex relationships and process. Nova York: Springer Science and Business Media; 2015. 378 p.
36. Bishop DC, Pankratz MM, Hansen WB, Albritton J, Albritton L, Strack J. Measuring fidelity and adaptation: reliability of a instrument for school-based prevention programs. *Eval Health Prof.* 2014; 37(2):231-57.

Received: Nov 21<sup>th</sup> 2017

Accepted: Dec 7<sup>th</sup> 2018

Corresponding Author:  
Douglas Garcia  
E-mail: garciadouglas90@gmail.com  
 <https://orcid.org/0000-0001-5245-6075>

**Copyright © 2019 SMAD, Rev. Eletrônica Saúde Mental Álcool Drog.**  
This is an Open Access article distributed under the terms of the Creative Commons (CC BY-NC).  
This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.