



# Is Implementation Fidelity Important? A Systematic Review on School-Based Mental Health Programs

Rodrigo Rojas-Andrade<sup>1,2</sup>  · Loreto Leiva Bahamondes<sup>2</sup>

Published online: 2 February 2018

© California Association of School Psychologists 2018

## Abstract

Although implementation fidelity (IF) is an important factor in interpreting the outcomes of intervention programs, so far there is little knowledge about how it actually relates to them. The purpose of this systematic review was to identify which component of IF (adherence, quality of the intervention, exposure to the intervention and receptiveness) is the most important one for attaining the expected results in school-based mental health programs. A search in four electronic databases (APA PsycNET, PUBMED, EBSCO, and ISI-WEB Science) yielded 31 articles published between 2006 and 2016 that met all the established inclusion criteria. To determine the associations present, the proportion was calculated between the number of times that the components of IF were significantly linked to outcomes and the total number of times that the association was evaluated. It was observed that the various components of IF are linked to outcomes 40% of the time and that the strongest association is established with students' exposure and receptiveness to the intervention. Lastly, findings and their implications for future research are discussed.

**Keywords** Implementation fidelity · Effectiveness · School programs · Preventive interventions · School-based mental health programs

The aim of school mental health programs (SMHP) is to help the students to achieve their goals through the improvement and strengthening of their emotional well-being, their psychosocial skills, and their positive teaching-learning environments. Such programs should be based on empirically tested intervention models to ensure that the actions taken will produce the desired outcomes.

This realization and the interest in obtaining results has led to a growth of evidence-based interventions (EBI) in line with an approach that considers and summarizes empirical experience, professional expertise, and student characteristics in connection with the intervention (Weist and Lever 2014).

One of the principles of this approach is that replication will make it possible to reproduce the results found in experimental tests; however, this does not always occur, because when EBIs are implemented in complex and heterogeneous contexts, such as schools, they seem to lose strength and applicability (Proctor et al. 2009). In fact, multiple SMHP have been conducted (e.g., [www.casel.org](http://www.casel.org), [www.blueprintsprograms.com](http://www.blueprintsprograms.com)) whose effectiveness has been tested rigorously, but little is known about how to implement them properly in schools (Forman et al. 2013; Sarno et al. 2014).

In this context, the last decade has witnessed the development of implementation science, focused on understanding how to transfer the benefits of EBIs to the real world by studying the processes and components of their implementation in everyday intervention contexts (Bhattacharyya et al. 2009).

Studying implementation involves understanding the social contexts in which the actions are executed, and examining the technical resources and organizational conditions that support the proper execution of an intervention. In particular, it involves determining how the executed actions conform to a number of contexts while maintaining fidelity to the intervention model (Dupaul 2009; Perepletchikova 2011; Schulte et al. 2009).

Implementation fidelity (IF)—or treatment integrity—is one of the key aspects of implementation research and refers

---

✉ Rodrigo Rojas-Andrade  
rojasa01@docentes.academia.cl

Loreto Leiva Bahamondes  
loretoleivab@u.uchile.cl

<sup>1</sup> School of Psychology, Faculty of Social Sciences, University Academy of Christian Humanism, Condell 343, Providencia, Santiago, Chile

<sup>2</sup> Department of Psychology, Faculty of Social Sciences, University of Chile, Capitan Ignacio Carrera Pinto 1045, Ñuñoa, Santiago, Chile

to the degree to which an intervention is conducted in accordance with its intervention model (Perepletchikova 2011; Schulte et al. 2009). Knowing the threshold at which IF starts generating the desired (or undesired) results makes it possible to estimate the efforts required to implement an intervention adequately, this is fundamental for political-technical decision-making, because investing resources in intervention programs without applying them correctly is as pointless as investing in ineffective programs (Durlak 2015).

IF is a complex variable that is obtained by measuring several basic components of the operative model (Dane and Schneider 1998; Dupaul 2009; Hagermoser-Sanetti and Kratochwill 2009; Sarno et al. 2014). Its complexity arises from its multilevel and multidimensional nature, as it incorporates different dimensions of the various participants who are involved in nested relationships in an intervention. Even though several taxonomies are available (e.g., Hagermoser-Sanetti and Kratochwill 2009; Schulte et al. 2009), they all consider at least four components of IF in SMHP that are associated with the interventionist and with the students taking part in the intervention. At the interventionist's level, IF assesses the extent to which this practitioner's work conforms to the planned actions and the prescribed quality level. In particular, two aspects are considered: adherence, which concerns the degree of fulfillment of or fidelity to the practical components specified in the operative model (Schultes et al. 2015), and intervention quality, which refers to the degree of skill, enthusiasm, and commitment in the execution of the actions (Dane and Schneider 1998). At the level of student participants, IF assesses whether the intervention was adapted to fit them (i.e., if they actually were the target population) and whether they received the prescribed number of sessions in the right way for attaining the expected results. Specifically, the model also considers intervention exposure, which refers to the doses received according to what was planned (Coddling and Lane 2015), and receptiveness, which refers to the degree of relevance of the intervention or the commitment that the participants must display towards it (Low et al. 2014b).

## Previous Reviews

To date, a number of reviews have been published showing that IF has been insufficiently reported (Sutherland et al. 2013). As shown in Table 1, this trend has remained stable over the last 35 years of research, never exceeding 50% of the articles published in specialized journals.

One reason for this is that researchers assume IF to be present when using experimental designs. These studies are focused on control over the execution of actions; therefore, the authors are not always interested in checking whether execution is implemented with fidelity, because they make great efforts to ensure that this be so by training, supervising, and

monitoring the actions of the individuals in charge of the execution. However, since not all researchers are concerned with the degree of IF that their efforts produced, it cannot be assured that the results obtained are due to the actions executed, which means that their conclusions must be interpreted cautiously. For this reason, several professional organizations and funding agencies have made a commitment to IF, thus encouraging researchers to increase their efforts in this area (DiGennaro Reed and Coddling 2014).

Given the recent interest in IF, and despite the relative consensus regarding its measures (Dane and Schneider 1998; Schulte et al. 2009), the researchers who report it do so incorrectly: they only mention the aspects considered to promote IF, such as the use of manuals, training, supervision, monitoring, and the implementation context, but only a few give specific details about how much IF was attained and in which component. As a result, reviews of IF have focused on describing these elements, showing that better intervention outcomes are achieved when IF is fostered. It has been demonstrated that well-defined interventions lead to better outcomes than ambiguous and unclear ones. In a pioneering study, Tobler (1986) reviewed 143 drug prevention programs aimed at adolescent school populations. He found that the results obtained were linked to the operational definition of the intervention, because programs that included such a definition obtained larger effect sizes than those that did not. More recently, in a study of 55 assessments of mentorship programs for adolescents, DuBois et al. (2002) found that those programs that used evidence-based practices (with adequate operational definitions) yielded better results. Another similar study conducted by Sklad et al. (2012) examined 75 articles on the effectiveness of universal socioemotional learning programs and found that outcomes are linked to the reported use of a manual.

Other reviews demonstrate that monitoring of implementation is associated with intervention outcomes. Smith et al. (2004) examined 84 studies on the effectiveness of anti-bullying programs and observed that those that involve systematic monitoring tend to be more effective than those without monitoring. The same finding was reported by DuBois et al. (2002), who warn that low-monitoring mentorship programs for adolescents can be detrimental to their young participants, especially if they come from deprived social backgrounds. Another review along similar lines conducted by Wilson et al. (2003) studied 221 articles on programs for reducing aggressive behavior and found that implementation quality (understood as the resolution of implementation problems through monitoring) is strongly related to the results obtained.

In the implementation field, the review by Durlak and DuPre (2008) is essential reading. Using several analytical strategies, these researchers conclude that IF is important, and they find that well-implemented programs (i.e., those in

**Table 1** Fidelity implementation report in school mental health and related fields

Author	Years included in review	Number of analyzed papers	Percent of papers that reported IF
Griffith et al. (2009)	1977-2005	44	52%
McIntyre et al. (2007)	1991-2005	142	30%
Bruhn et al. (2015)	1993-2012	79	46%
Hagermoser-Sanetti et al. (2011)	1995-2008	210	50%
Hagermoser-Sanetti et al. (2011)	1999-2009	72	42%
Wheeler et al. (2014)	2000-2009	33	49%

which IF is promoted) have effect sizes two to three times larger than those of badly implemented ones and that, under ideal conditions, interventions with high IF can be up to 12 times more effective than poorly implemented ones. These results suggest that studies striving to ensure adequate conditions for promoting IF have better outcomes than those that fail to do so. However, they provide no practical information about, for instance, how a manual must be used or how much it is possible to deviate from it to achieve the expected results. These data are essential for replicating interventions in everyday school contexts, where the possibilities of being totally faithful to the intervention model are slim (Sarno et al. 2014).

On the other hand, although these results stress the importance of monitoring and supporting the implementation of operationally well-defined interventions, they do not shed light on the specific relationship between the implementation level of each component and the outcomes observed. Despite the importance of this issue, studies usually fail to connect IF with outcomes (Schoenwald and Garland 2013). This is why only two reviews were found that specifically refer to such relationship: in the first, Dane and Schneider (1998) reveal that the higher the dose, the better the results, while in the second, Durlak and DuPre (2008) note that in 79% of the interventions reviewed there is a significant link between dosage and adherence and at least half of the intervention results.

## Purpose of this Review

Understanding the complexity of IF in SMHP requires exploring how IF components associate with results and weighing the relative importance of their influence on them. In this regard, although some authors assume that adherence is the heart of IF because it hosts the specific strategies and techniques originating from the change model, it is necessary to delve deeper into the other components. This review sought to shed light on the link between IF components and the expected outcomes of SMHP. To accomplish this, a descriptive analysis of studies that addressed this connection was conducted by counting the number of times the components of IF were significantly linked to the results measured.

## Method

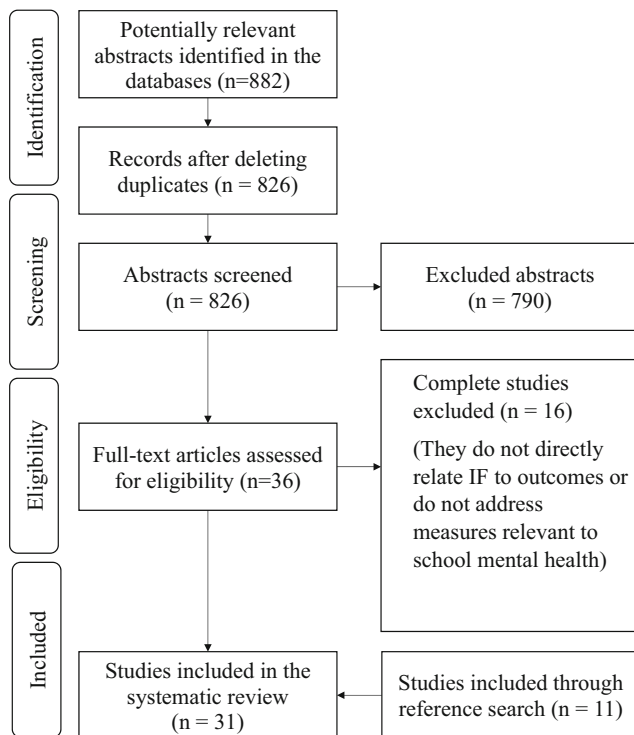
### Literature Search Strategy

Two search strategies were used to ensure an exhaustive search of the existing literature. The first was to identify primary studies in the online databases APA PsycNET ( $n = 315$ ), PUBMED ( $n = 216$ ), EBSCO ( $n = 95$ ), ISI-WEB Science ( $n = 89$ ), and Scopus Science Direct ( $n = 167$ ) using the following keywords: *fidelity, integrity, implementation, adherence, dosage, dose, exposure, quality, professional competence, engagement, receptiveness, outcome, effectiveness, efficacy\*, school\*, preschool\*, and class\**.

A total of 882 articles were reviewed by the author and her assistant by reading their titles and abstracts. Thirty-six articles were selected for potential inclusion in the review; after a detailed examination, this number went down to 20. The second strategy was to review the reference section of each selected article, which yielded 11 additional articles. Thirty-one articles in total were included in this review (see Fig. 1).

### Inclusion Criteria

Eligible studies for this review were those that: a) reported measures of the relationship between an IF component (adherence, intervention quality, dose, and/or receptiveness) and the results of the intervention; b) assessed school-based mental health programs, defined as interventions intended to promote, prevent, or treat problems associated with students' emotional well-being (externalizing and internalizing), psychosocial skills, and positive teaching-learning environments (such as school climate, school relationships, and bullying); c) were published between 2006 and 2016. This period was selected because, according to some authors, mentions of IF in articles began to increase in the mid-2000s (Hagermoser-Sanetti et al. 2011); d) were published in peer-reviewed academic journals only. Although this criterion introduces a publication bias, it was included given the importance of examining rigorous scientific information regarding the link between IF and outcomes.



**Fig. 1** Flow chart on the different phases of the systematic

## Exclusion Criteria

Studies excluded from this review were those that: a) reported qualitative or review-based results; b) failed to provide statistical information about the relationship between IF and outcomes; c) reported results of school interventions aimed only at improving certain aspects of learning processes (e.g., literacy) or focused on physical health and risk behaviors (e.g., nutrition programs or sexual risk behaviors).

## Special Treatment of Research Reports

Interventions conducted in different school levels (e.g., primary and secondary education) reported in a single study were coded and analyzed separately. The same procedure was followed when one study employed different measures to assess and determine the associations of a single IF component (e.g., adherence according to two informants or different temporal scales of exposure to the intervention).

## Coding

### Independent Variables: Components of Implementation Fidelity

Assuming that an intervention is operationalized through the measurement of IF components, these were regarded as independent variables in relation to the outcomes of SMHP. The underlying assumption is that measurement of these variables

constitutes a window into the intervention's black box, usually regarded as a dichotomous variable (treatment vs. control). Therefore, the four major components of IF were considered (Dane and Schneider 1998; Dupaul 2009; Schulte et al. 2009), and the presence or absence of their measurement in each study was assessed.

**Adherence** This element concerns the degree of fulfillment of or fidelity to the practical components specified in the operative model. It was deemed to be present when the researchers measured the use of specific techniques, the application of general principles, or the fulfillment of key phases of the intervention process.

**Intervention Quality** This element concerns the degree of competence with which the interventionist executes the actions. Intervention quality was deemed to be present when the researchers measured the interventionists' knowledge about the intervention, the interventionists' skill as demonstrated when performing the actions, and/or their attitudes towards the actions performed (e.g., enthusiasm and commitment).

**Exposure to the Intervention** This component concerns the number of intervention sessions in which a student participates. Exposure to the intervention was deemed to be present when the researchers measured the number, or the proportion of sessions received compared to the total number of sessions (e.g., annual, biannual, monthly, weekly, or daily).

**Receptiveness** This element concerns the degree to which students were committed to the intervention. Receptiveness was deemed to be present only when the researchers measured the participants' attitude towards the intervention (e.g., enthusiasm and commitment) through post-intervention surveys or questionnaires.

**Combined IF Indexes** This element concerns the construction or usage of IF indexes based on the combination of the components mentioned above.

### Characterization Variables: Type of Intervention and Measurement of IF

The types of intervention were characterized following elements of the taxonomy published by Humphrey et al. (2013) and other categories used in reviews of prevention programs in SMHP (Sklad et al. 2012; Weare and Nind 2011; Wilson et al. 2003).

**Scope of the Intervention** This element refers to the educational level in which the intervention is carried out. The label "universal" was assigned to the studies that report

interventions designed for school-wide application. The label “selective/targeted” was used for those reporting interventions aimed at a subgroup identified as at-risk of experiencing (or currently experiencing) social, emotional, and behavioral problems. “Mixed” was used when both approaches were present. In each case, the coding was based on the descriptions included in the sample and participants section.

**Structural Components** These are the types of activities performed during the intervention. The “skills teaching” code was used for studies that report interventions structured upon the basis of lessons and activities aimed at helping children develop and strengthen their social, emotional, and behavioral skills. The “strengthening of the school environment” code was used for studies that report interventions focused on improving school climate, culture, and norms. This category included interventions based on the Positive Behavior Intervention and Supports (PBIS) approach. The “mixed” code was used for studies reporting interventions based on both components. In each case, only the component addressed in the study was considered rather than the additional components in the full version of the program.

**Prescriptivity** This element concerns how prescriptive the specific actions of a given intervention are. The “top-down” label was assigned to studies reporting interventions based on planned actions and structured guidelines or manuals describing the implementation procedures, with the explicit obligation of performing them exactly as designed. The “bottom-up” code was assigned to studies reporting interventions that emphasize flexibility and local adaptation in the activities to be implemented. The “mixed” code was used when prescriptivity depends on the structural components of the intervention. In each case, the coding was based on the descriptions included in the introduction and procedures sections; in some cases, the websites of the programs assessed were also reviewed.

**Interventionist** This is the professional who executes the intervention. The “teacher” code was used when the teacher applied the components of the intervention to students; the “school team” code was selected when an internal school team implemented these actions; and the “another professional” code was assigned when the actions were conducted by social, medical, or other professionals.

**School Grade Level of the Participants** This refers to the grade level to which the student participants belong. Articles were coded as relating to preschool, primary education (grades 1–8), or secondary education (grades 9–12). In each case, the grade level in which the students received the intervention was considered, rather than the grade level in which the outcomes were assessed.

To characterize the measurement of IF, the following elements were considered: measurement instruments, measure validity, and measurement frequency.

**Instruments for Measuring IF** This component concerns the type of instrument used to assess IF. Studies were coded as: “permanent products” when assessment reports or cards were generated regarding the implementation process; “observations” when systematic observation checklists were used or when video recordings of the execution of the actions were coded; “self-reports” when questionnaires or surveys about the implementation process were used; “interviews” when conversational production techniques were employed; and “multiple instruments” when two or more approaches were used.

**Measure Validity** This refers to whether, or not validity indicators were reported for the IF measures used. Studies were coded as yes/no depending on the presence or absence of measures of validity.

**Frequency of IF Measurement** This component refers to the number of times that IF was measured. Studies were classified as: “session-by-session,” “weekly,” “monthly,” “bi-monthly,” “quarterly or four-monthly,” “five-monthly or bi-annually,” “annually,” or “once during the whole intervention”.

#### **Dependent Variables: Outcomes of the Interventions**

Five dependent variables were used in this section: internalizing mental health difficulties, externalizing mental health difficulties, socio-emotional skills, school relationships, and academic performance.

**Internalizing Mental Health Difficulties** This category included outcome variables associated with internalizing behaviors such as depressive and anxious symptoms, loneliness, suicide attempts, somatic problems, psychological well-being, or need of psychological care.

**Externalizing Mental Health Difficulties** This category included outcome variables associated with externalizing behaviors such as non-fulfillment of tasks, problems with classmates, aggressive, disruptive, hyperactive, or antisocial conduct, and substance use.

**Socio-Emotional Skills** This category included outcome variables associated with the five domains of socio-emotional learning, which include: self-awareness (e.g., self-esteem, recognition of emotions), self-control (e.g., stress management, impulse control), social awareness (e.g., empathy, respect for others), social skills (e.g., active listening, cooperation), and responsible decision-making (e.g., problem-solving, anticipation of consequences).

**School Performance** This category included outcome variables associated with academic performance (e.g., mathematics and language test scores) and dropout (absenteeism, dropout).

**School Relationships** This category included outcome variables associated with the perception of social environments of teaching-learning, such as classroom climate or bullying level at school or in the classroom.

### Coding Reliability

Coding was performed by the research team. In two training sessions, the team was instructed by the first author regarding the general topic under study and the search and coding procedures used. Each stage of the process was guided by a reference document that provided step-by-step instructions for executing the actions of the intervention. The search and coding processes were conducted in full by both team members, whose discrepancies were solved during weekly work sessions. When no agreement was reached, the third author was contacted to provide assistance as an expert referee.

### Proportion Ratio and Data Analysis

Given the lack of consensus regarding the measurement of IF components (Lewis et al. 2015; Schoenwald and Garland 2013) and the diversity of outcome variables in SMHP (Durlak et al. 2011; Sklad et al. 2012; Weare and Nind 2011; Wilson et al. 2003), the counting technique was used to analyze the relationship between IF and results (Cooper 2017). Based on this data, a proportion ratio (PR) was calculated between the number of times that IF components were significantly associated with outcomes ( $p < .05$ ) and the total number of outcomes measured. That is, if a study considered five possible outcomes and only one of them was found to be significantly associated with a component of IF, the PR was .2. It should be noted that the PR was used in this review as a measure to describe the ratio of the number of significant relationships to the total number of relationships evaluated. Reference to “higher” or “lower” PR values merely indicates that a higher or lower number of significant associations between IF and outcomes was observed, and is not meant to indicate statistical significance.

## Results

### Characteristics of the Studies Reviewed

The characteristics of the studies reviewed are summarized in Table 2. Half of the studies linking IF to SMHP outcomes were found to published between 2014 and 2016 ( $n = 16; 52%$ ),

which indicates an upward trend. With respect to geographical location, most studies were conducted in English-speaking countries, especially the United States ( $n = 17; 55%$ ).

Seventy-seven percent ( $n = 24$ ) of the studies only assess universal interventions, although if mixed interventions are added (i.e., those that include both universal and targeted activities), this value reaches 93% ( $n = 29$ ). The same is true of the structural components of the intervention, most of which are associated with skills teaching ( $n = 23; 74%$ ) if we take into account the programs that conduct such actions alongside school environment strengthening ones. The most common interventionists are teachers, both as part of a team ( $n = 8; 26%$ ) and as main implementers ( $n = 21; 68%$ ). With respect to school level, most interventions were conducted in primary education ( $n = 23; 74%$ ).

It was found that the 31 studies reviewed considered 171 outcome variables. Of these, 31% ( $n = 55$ ) concern mental health difficulties, 32% ( $n = 57$ ) refer to socio-emotional skills, 20% ( $n = 36$ ) involve academic performance, and 13% ( $n = 23$ ) reflect school relationships. With respect to the components of IF measured, the most prevalent indicators were adherence ( $n = 24; 77%$ ) and dosage ( $n = 18; 58%$ ). Nearly 40% (12) of the studies measure IF through self-report instruments, followed by those that employ more than one measure (9; 29%), including a combination of interviews, observations, and self-reports. 45% (14) provide data on the validity of their measures, although only consistency indicators tend to be included (Cronbach's alpha or inter-rater agreement). In addition, it was found that most researchers measure IF more than once during the intervention ( $n = 22; 71%$ ), commonly on a session-by-session basis ( $n = 11; 35%$ ).

### Associations between Outcomes and IF

Considering the special treatment of some studies included in the review, 41 interventions were analyzed. In total, the relationship between IF components and the outcome variables considered was measured 273 times, with a positive association being found in 40% of cases ( $n = 97$ ; See Table 3).

Examining the associations by IF component revealed that participant receptiveness is the element most frequently linked to outcomes (60%;  $n = 31$ ), in contrast with the aggregate IF index, which only reaches 10%. With respect to IF components and outcome variables, it was observed that the significant association rate of adherence did not surpass 30%. Intervention quality was associated with internalizing difficulties on all occasions ( $n = 3$ ), although its connection with school relationships was null. Dosage was associated up to 50% ( $n = 17$ ) with outcome variables, except for school relationships, with which there was no association. The combined fidelity index displayed no association with any of the mental health difficulties considered, and was linked to the rest of the variables in less than 30% of the cases. The component that

**Table 2** Characteristics of the 31 studies reviewed

Characteristics	n	%
Publication years		
2006-2009	3	10
2010-2013	12	39
2014-2016	16	52
Geographical location		
Australia	3	10
Finland	2	6
England	6	19
Ireland	1	3
Norway	2	6
United States	17	55
Scope of the intervention		
Universal	24	77
Selective/targeted	2	6
Mixed	5	16
Structural components		
Skills teaching	14	45
Strengthening of the school environment	8	26
Mixed	9	29
Prescriptivity		
Bottom-up	10	32
Top-down	16	52
Mixed	5	16
Interventionist		
Teacher	21	68
School team	8	26
Another professional	2	6
School level of the participant		
Preschool	4	13
Primary education	23	74
Secondary education	4	13
Number of outcome evaluated		
Internalizing difficulties	10	6
Externalizing difficulties	45	25
Social-emotional skills	57	32
School performance	36	20
School relationship	23	13
Components of implementation Fidelity		
Adherence	24	77
Intervention Quality	8	26
Exposure to the intervention	18	58
Receptiveness	6	19
Combine IF indexes	3	10
Instrument for measuring IF		
Self-report	12	39
Interviews	1	3
Observation	6	19
Permanent Product	3	10
Multiple	9	29

**Table 2** (continued)

Characteristics	n	%
Measure Validity		
Yes	14	45
No	17	55
Frequency of IF measurement		
Annually	9	29
Every five or six months	3	10
Every three or four months	3	10
Weekly or monthly	5	16
Session by session	11	35

displayed the highest number of significant associations with the outcomes assessed was receptiveness, operationalized as student commitment. Receptiveness was linked in 100% of cases with internalizing difficulties, in 60% of cases with externalizing difficulties, in 70% of cases with socio-emotional skills, in 40% of cases with school performance, and in 30% of cases with school relationships.

### Comparison of the Associations between Outcomes and IF According to Characterization Variables

The 273 assessments of the relationship between IF and outcomes were analyzed according to the selected characterization variables. In the studies that analyze universal interventions, the associations between IF and outcomes were found to be twice as high as those in studies reporting selective and/or mixed intervention data. PR values were 6% higher in top-down than in bottom-up interventions, although mixed interventions displayed 23% fewer significant associations. In general, PR values were found to be twice as high in primary education as they are in preschool and secondary education.

A look at the characteristics of IF measurements reveals that, when permanent products and self-report instruments were used, PR was approximately two times larger than when measured with other instruments. An unexpected result was that the studies reporting the validity of IF measures displayed lower PR values than those that did not provide this data. Finally, it was observed that PR values were 5% higher when researchers measured IF only once per year (See Table 4).

### Discussion

The purpose of this review was to shed light on the link between IF components and the outcomes of SMHP. Thirty-one articles were found that directly addressed this association. Of these, approximately half were published two years ago. This reveals a recent interest in empirically testing the hypothesis

**Table 3** Proportion ratio of the significant relationships between results and IF

Outcome	All Components		Adherence		Intervention Quality		Exposure to the intervention		Receptiveness		Combine IF indexes	
	PR	(SR/NR)	PR	(SR/NR)	PR	(SR/NR)	PR	(SR/NR)	PR	(SR/NR)	PR	(SR/NR)
All outcome	0.4	(97/273)	0.3	(29/104)	0.2	(5/22)	0.4	(30/77)	<b>0.6</b>	(31/53)	0.1	(2/17)
Internalizing difficulties	0.4	(7/16)	0.2	(1/5)	<b>1.0</b>	(2/2)	0.3	(1/4)	<b>1.0</b>	(3/3)	0.0	(0/2)
Externalizing difficulties	0.3	(25/76)	0.2	(4/18)	0.3	(1/4)	0.3	(9/29)	<b>0.6</b>	(11/19)	0.0	(0/6)
Social-emotional skills	0.4	(42/94)	0.3	(10/29)	0.2	(2/11)	<b>0.5</b>	(17/31)	<b>0.7</b>	(12/17)	0.2	(1/6)
School performance	0.3	(15/44)	0.3	(9/29)	–	–	0.4	(3/7)	0.4	(2/5)	0.3	(1/3)
School relationship	0.2	(8/43)	0.2	(5/23)	0.0	(0/5)	0.0	(0/6)	0.3	(3/9)	–	–

In bold. Values greater than or equal to .5

Note: *NR* number of relations evaluated, *SR* number of significant relationships lower than 5%, *PR* proportion ratio

**Table 4** Proportion ratio of the significant relationships between results and IF by Characteristics of the studies

Characteristics	PR	(SR/NR)
Scope of the intervention		
mixed	0.2	(11/52)
Selective/targeted	0.2	(4/19)
universal	0.4	(82/202)
Structural components		
Skill teaching	0.4	(66/156)
Strengthening of the school environment	0.3	(16/48)
mixed	0.2	(15/69)
Prescriptivity		
bottom-up	0.4	(14/38)
Mixed	0.2	(15/75)
top-down	0.4	(68/160)
School level of the participants		
Preschool	0.2	(15/63)
Primary education	0.4	(77/189)
Secondary education	0.2	(5/21)
Instrument for measuring IF		
Self-report	0.45	(56/124)
Interviews	0.00	(0/1)
Multiple	0.27	(21/79)
Observation	0.25	(13/52)
Permanent Product	0.41	(7/17)
Measure Validity		
No	0.42	(58/138)
Yes	0.29	(39/135)
Frequency of IF measurement		
More than once a year	0.35	(80/230)
Annually	0.40	(17/43)

In bold. Values greater than or equal to .5

Note: *NR* number of relations evaluated, *SR* number of significant relationships lower than 5%, *PR* proportion ratio

of the impact of IF on outcomes, especially in the United States, where most of these studies were conducted.

The interventions assessed are mostly universal, top-down, executed by teachers, aimed at primary school students, and focused on the improvement of socio-emotional skills, which is consistent with the international trend of evidence-based SMHP (Kutash et al. 2006). In this regard, it should be noted that despite extensive evidence in support of programs that target socio-emotional learning (Durlak et al. 2011), these are not free from criticism regarding the universality of their benefits outside their country of origin (Berry et al. 2015). Therefore, it is fundamental to explore the active ingredients of interventions and the processes involved in their implementation over the last years (Durlak 2016).

With regard to IF measures, the most frequent were adherence and dosage, a result that matches previous reviews (Dane and Schneider 1998; Durlak and DuPre 2008). In this respect, adherence is the heart of IF for some researchers (Gresham 2009), because if one wishes to measure the operationalization of the intervention model empirically, the differences between what interventionists actually do and what they should do cannot be ignored. Similarly, other researchers (Coddling and Lane 2015; McGinty et al. 2011) have stressed the importance of dosage for the measurement of IF, because it is not enough for something to be done according to the guidelines, it is also important to know how much is received by students.

Adherence and dosage represent the quantitative aspect of IF. This means that researchers have chosen not to consider qualitative aspects such as intervention quality and interventionist–participant interactions, despite the importance of these IF components in other areas. In psychotherapy, for instance, the therapist’s competence is an essential factor in measuring IF (Perepletchikova et al. 2007). On the other hand, it was found that in response to the complexity of IF, some researchers have developed indexes that consider two or more components, usually adherence, dosage, and intervention quality.



With respect to measurement instruments, most researchers use self-reports, either exclusively or alongside interviews, observations, or permanent products. It is interesting to observe that nearly half of the studies report the psychometric properties of their measures, although most of them were inter-rater agreement and internal consistency indexes. In this regard, as other authors have pointed out (Dupaul 2009), robust psychometric instruments need to be developed, and this in turn requires further consensus on the definitions of IF and the key components common to all program types. The work of authors such as Abry et al. (2015) and Rimm-Kaufman et al. (2014) provides useful observation and self-report instruments in this direction.

As to the central topic of this review, an especially relevant aspect concerns the use of PRs as indicators of the association between IF and results. Even though the ideal procedure would have been to conduct a meta-analysis and calculate effect sizes, the diversity of dependent and independent variables forced us to seek new ways to tackle this issue. In this context, the PR assumes that if the researchers measured certain given dependent variables, it is because these emerged from the model supporting the intervention; therefore, the multiple components of IF should be linked to them. From the findings of this review, it may be concluded that the relationship between IF and SMHP outcomes is partial, as IF (understood as the sum of its components) was found to be associated with the measured outcomes of SMHP only 40% of the time, i.e., a low rate considering the theoretical importance of IF.

Regarding outcome variables, IF was found to be weakly associated with school relationships, school performance, and externalizing difficulties. This is a worrying finding, because one of the fundamental principles of SMHP is that interventions contribute to educational goals (Suldo et al. 2014), but this does not seem to be the case in the studies reviewed. This urges us to reflect on the ingredients of SMHP that actively contribute to academic improvement, because the fact that no significant associations were found between IF and these outcomes does not mean that interventions do not have a positive impact on academic performance. Actually, the available evidence shows that they do help (Durlak et al. 2011); thus, it is necessary to explore in depth the associations between the actions performed and their results. It could be hypothesized that the lack of a link between IF and outcomes arises from the presence of variables not considered in the intervention models (and thus not measured as part of IF) that exert a significant influence during the execution process; alternatively, it may be due to methodological problems in the measurement of IF components.

On the other hand, when the four components of IF were considered in connection with the multiple outcome variables examined, the relative importance of each was assessed. One of the main findings of this review is that adherence, despite

being the most frequently reported component, is weakly associated with outcome variables. The same is true of intervention quality. This is an unexpected finding, given the importance that authors ascribe to these dimensions as reflecting the interventionist's practices. On the other hand, the components of IF located in the participant were observed to display stronger associations. A case in point is receptiveness, which is expressed in students' commitment to the intervention and is linked to 70% of socio-emotional outcomes.

Even though these results must be cautiously weighed, the PRs found may indicate that, in the case of SMHP, the most important aspect is for students to participate as many times as they can in the activities planned and with as much commitment as possible, regardless of the intervention model used. In this regard, it would be interesting to test the hypothesis that change takes place to a large extent due to personal factors such as involvement or attitudes towards the intervention activities, that is, depending on how strongly the participants believe that what they are doing can help them (Low et al. 2014a, b). These results highlight the importance of fostering students' commitment to and enthusiasm for participating in school-based mental health interventions. This can be achieved through change readiness strategies and school-based mental health literacy (Macklem 2014).

When examining PRs in light of the characterization variables of the interventions, stronger links between IF and outcomes were found in universal top-down skills teaching programs for primary education students. This can be partly explained because school-wide readiness for the implementation of SMHP can create a suitable environment for commitment in all the parties involved, especially students.

Another relevant result is that more than half of the studies examined measure IF monthly or less frequently. Even though this appears to be a good decision because it makes it possible to describe the variation in IF throughout the intervention, the data collected in this review show that when researchers measure IF only once a year, more instances of significant association with outcomes are observed. This seems to suggest that an overall assessment of the implementation process is better than a detailed one.

The limitations of this study concern three aspects that should be considered in future research. First, given the recent attention paid to IF in applied research, several different outcomes and interventions were included. However, even though this approach provides for generalization, it restricts the depth of analyses. Future reviews should focus on a single type of intervention and its related outcome to eliminate possible biases introduced by broad categories such as those included in this review. A second limitation was introduced by the requirement of research rigor, represented by publication in peer-reviewed journals. Although this choice provided for collecting high-quality scientific information, it produced a major by omitting evidence from other sources. Interventions

are carried out in many schools, but their results are not always reported in academic outlets and are instead filed away in the gray literature of foundations or government agencies. Therefore, it is necessary to explore these documents to shed more light on the associations examined in this study, especially in underdeveloped or developing countries.

The third limitation was the use of PRs for estimating the connection between IF and outcomes. This review set out to question indirect estimations of the importance of IF on outcomes. However, given the characteristics and designs of the primary studies analyzed, it was unable to make progress towards estimating IF effect size. In this regard, the PR is still an indirect estimation, but has the advantage of addressing direct associations between IF and outcomes, unlike previous reviews that had to create a posteriori implementation indicators. In this context, it is necessary to issue clear guidelines for researchers interested in assessing IF and to reach a consensus regarding the operational definitions of its multiple components.

**Funding** This study was funded by National Fund for Scientific and Technological Development [Grant number 1171634, 2017]; Postgraduate studies funded by CONICYT-PCHA / PhD National / 2015-21150612.

## Compliance with Ethical Standards

**Ethical Approval** This article does not contain any studies with human participants performed by any of the authors.

**Conflict of Interest** Author A has received research grants from Postgraduate studies funded by CONICYT-PCHA / PhD National / 2015-21,150,612. Author B has received research grants from National Fund for Scientific and Technological Development [Grant number 1171634, 2017].

## References

References marked with an asterisk indicate studies included in the systematic review

- \*Abyr, T., Hulleman, C.S., & Rimm-Kaufman, S.E. (2015). Using indices of fidelity to intervention core components to identify program active ingredients. *American Journal of Evaluation*, 36(3), 320–338. <https://doi.org/10.1177/1098214014557009>.
- \*Askill-williams, H., Dix, K.L., Lawson, M.J., & Slee, P.T. (2013). Quality of implementation of a school mental health initiative and changes over time in students' social and emotional competencies. *School Effectiveness and School Improvement*, 24(3), 357–381. <https://doi.org/10.1080/09243453.2012.692697>.
- \*Barlow, A., Humphrey, N., Lendrum, A., Wigelsworth, M., & Squires, G. (2015). Evaluation of the implementation and impact of an integrated prevention model on the academic progress of students with disabilities. *Research in Developmental Disabilities*, 36, 505–525. <https://doi.org/10.1016/j.ridd.2014.10.029>.
- \*Berry, V., Axford, N., Blower, S., Taylor, R.S., Edwards, R.T., Tobin, K., ... Bywater, T. (2015). The effectiveness and micro-costing analysis of a universal, school-based, social-emotional learning programme in the UK: A cluster-randomised controlled trial. *School Mental Health*, 8(2), 1–19. <https://doi.org/10.1007/s12310-015-9160-1>.
- Bhattacharyya, O., Reeves, S., & Zwarenstein, M. (2009). What is implementation research?: Rationale, concepts, and practices. *Research on Social Work Practice*, 19(5), 491–502. <https://doi.org/10.1177/1049731509335528>
- Bruhn, A. L., Hirsch, S. E., & Lloyd, J. W. (2015). Treatment Integrity in School-Wide Programs: A Review of the Literature (1993-2012). *Journal of Primary Prevention*, 36(5), 335–349. <https://doi.org/10.1007/s10935-015-0400-9>
- \*Burke, R.V., Oats, R.G., Ringle, J.L., Fichtner, L.O., & DelGaudio, M.B. (2011). Implementation of a classroom management program with urban elementary schools in low-income Neighborhoods: Does program Fidelity affect student behavior and academic outcomes? *Journal of Education for Students Placed at Risk*, 16(3), 201–218. <https://doi.org/10.1080/10824669.2011.585944>.
- \*Clarke, A.M., Bunting, B., & Barry, M.M. (2014). Evaluating the implementation of a school-based emotional well-being programme: A cluster randomized controlled trial of Zippy's friends for children in disadvantaged primary schools. *Health Education Research*, 29(5), 786–98. <https://doi.org/10.1093/her/cyu047>.
- Codding, R. S., & Lane, K. L. (2015). A spotlight on treatment intensity: An important and often overlooked component of intervention inquiry. *Journal of Behavioral Education*, 24(1), 1–10. <https://doi.org/10.1007/s10864-014-9210-z>
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23–45. [https://doi.org/10.1016/S0272-7358\(97\)00043-3](https://doi.org/10.1016/S0272-7358(97)00043-3)
- DiGennaro Reed, F. D., & Codding, R. S. (2014). Advancements in procedural Fidelity assessment and intervention: Introduction to the special issue. *Journal of Behavioral Education*, 23(1), 1–18. <https://doi.org/10.1007/s10864-013-9191-3>
- \*Dix, K.L., Slee, P.T., Lawson, M.J., & Keeves, J.P. (2012). Implementation quality of whole-school mental health promotion and students' academic performance. *Child and Adolescent Mental Health*, 17(1), 45–51. <https://doi.org/10.1111/j.1475-3588.2011.00608.x>.
- \*Domitrovich, C.E., Gest, S.D., Jones, D., Gill, S., & DeRousie, R.M.S. (2010). Implementation quality: Lessons learned in the context of the head start REDI trial. *Early Childhood Research Quarterly*, 25(3), 284–298. <https://doi.org/10.1016/j.ecresq.2010.04.001>.
- DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta analytic review. *American Journal of Community Psychology*, 30(2), 157–198. <https://doi.org/10.1023/A:1014628810714>
- Dupaul, G. J. (2009). Assessing integrity of intervention Implementation : Critical factors and future directions. *School Mental Health*, 1(3), 154–157. <https://doi.org/10.1007/s12310-009-9016-7>
- Durlak, J. A. (2015). Studying program implementation is not easy but it is essential. *Prevention Science*, 16(8), 1123–1127. <https://doi.org/10.1007/s11121-015-0606-3>
- Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education*, 46(3), 333–345. <https://doi.org/10.1080/0305764X.2016.1142504>
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350. <https://doi.org/10.1007/s10464-008-9165-0>
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal

- interventions. *Child Development*, 82(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- \*Evans, S.W., Schultz, B.K., & Demars, C.E. (2014). High school-based treatment for adolescents with attention-deficit / hyperactivity Disorder: Results from a pilot study examining outcomes and dosage. *School Psychology Review*, 43(2), 185–202.
- \*Ferrer-Wreder, L., Cadely, H.S.E., Domitrovich, C.E., Small, M.L., Caldwell, L.L., & Cleveland, M.J. (2010). Is more better? Outcome and dose of a universal drug prevention effectiveness trial. *Journal of Primary Prevention*, 31(5–6), 349–363. <https://doi.org/10.1007/s10935-010-0226-4>.
- \*Flannery, K.B., Fenning, P., Kato, M.M., & McIntosh, K. (2014). Effects of school-wide positive Behavioral interventions and supports and Fidelity of implementation on problem behavior in high schools. *School Psychology Quarterly*, 29(2), 111–124. <https://doi.org/10.1037/spq0000039>.
- Forman, S.G., Shapiro, E.S., Codding, R.S., Gonzales, J.E., Reddy, L.A., Rosenfield, S.A., ... & Stoiber, K.C. (2013). Implementation science and school psychology. *School Psychology Quarterly*, 28(2), 77–100. <https://doi.org/10.1037/spq0000019>.
- Gresham, F. M. (2009). Evolution of the treatment integrity concept: Current status and future directions. *School Psychology Review*, 38(4), 533–540.
- Griffith, A. K., Duppong Hurley, K., & Hagaman, J. L. (2009). Treatment Integrity of Literacy Interventions for Students With Emotional and/or Behavioral Disorders: A Review of Literature. *Remedial and Special Education*, 30(4), 245–255. <https://doi.org/10.1177/0741932508321013>
- \*Haataja, A., Voeten, M., Boulton, A.J., Ahtola, A., Poskiparta, E., & Salmivalli, C. (2014). The KiVa antibullying curriculum and outcome: Does fidelity matter? *Journal of School Psychology*, 52(5), 479–493. <https://doi.org/10.1016/j.jsp.2014.07.001>.
- Hagermoser-Sanetti, L., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, 38(4), 445–459.
- Hagermoser-Sanetti, L., Gritter, K., & Dobey, L. (2011). Treatment integrity of interventions with children in the journal of positive behavior interventions from 1999 to 2009. *Journal of Positive Behavior Interventions*, 14(1), 29–46. <https://doi.org/10.1177/1098300711405853>
- \*Hamre, B.K., Pianta, R.C., Mashburn, A.J., & Downer, J.T. (2012). Promoting young children's social competence through the preschool PATHS curriculum and MyTeachingPartner professional development resources. *Early Education & Development*, 23(6), 809–832. <https://doi.org/10.1080/10409289.2011.607360>.
- \*Hirschstein, M.K., Edstrom, L.V.S., Frey, K.S., Snell, J.L., & MacKenzie, E.P. (2007). Walking the talk in bullying prevention: Teacher implementation variables related to initial impact of the steps to respect program. *School Psychology Review*, 36(1), 3–21.
- \*Humphrey, N., Lendrum, A., Barlow, A., Wigelsworth, M., & Squires, G. (2013). Achievement for all: Improving psychosocial outcomes for students with special educational needs and disabilities. *Research in Developmental Disabilities*, 34(4), 1210–1225. <https://doi.org/10.1016/j.ridd.2012.12.008>.
- \*Kiviruusu, O., Bjorklund, K., Koskinen, H., Liski, A., Lindblom, J., Kuoppamaki, H., ... Santalahti, P. (2016). Short-term effects of the “together at school” intervention program on children's socio-emotional skills: A cluster randomized controlled trial. *BMC Psychology*, 4(27).
- Kutash, K., Duchnowski, A. J., & Lynn, N. (2006). *School-based mental health: An empirical guide for decision-makers*. Tampa: University of South Florida, The Louis De La Parte Florida Mental Health Institute, Department of Child and Family Studies, Research and Training Center for Children's Mental Health.
- Lewis, C.C., Fischer, S., Weiner, B.J., Stanick, C., Kim, M., & Martinez, R.G. (2015). Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria. *Implementation Science*, 10. <https://doi.org/10.1186/s13012-015-0342-x>.
- \*Low, S., Cook, C.R., Smolkowski, K., & Buntain-Ricklefs, J. (2014a). Promoting social-emotional competence: An evaluation of the elementary version of second step®. *Journal of School Psychology*. <https://doi.org/10.1016/j.jsp.2015.09.002>.
- \*Low, S., Van Ryzin, M.J., Brown, E.C., Smith, B.H., & Haggerty, K.P. (2014b). Engagement matters: Lessons from assessing classroom implementation of steps to respect: A bullying prevention program over a one-year period. *Prevention Science*, 15(2), 165–176. <https://doi.org/10.1007/s11121-012-0359-1>.
- \*Low, S., Smolkowski, K., & Cook, C. (2016). What constitutes high-quality implementation of SEL programs? A latent class analysis of second step® implementation. *Prevention Science*, 17(8), 981–991. <https://doi.org/10.1007/s11121-016-0670-3>.
- Macklem, G. (2014). *Preventive mental health at school. Evidence-based services for students*. New York: Springer.
- McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the journal of applied behavior analysis 1991-2005. *Journal of Applied Behavior Analysis*, 40(4), 659–72. <https://doi.org/10.1901/jaba.2007.659>
- McGinty, A. S., Breit-Smith, A., Fan, X., Justice, L. M., & Kaderavek, J. N. (2011). Does intensity matter? Preschoolers' print knowledge development within a classroom-based intervention. *Early Childhood Research Quarterly*, 26(3), 255–267. <https://doi.org/10.1016/j.ecresq.2011.02.002>
- \*Obsuth, I., Cope, A., Sutherland, A., Pilbeam, L., Murray, A.L., & Eisner, M. (2016). London education and inclusion project (LEIP): Exploring negative and null effects of a cluster-randomised school-intervention to reduce school exclusion-findings from protocol-based subgroup analyses. *PLoS One*, 11(4), 1–21. <https://doi.org/10.1371/journal.pone.0152423>.
- \*Pas, E.T., & Bradshaw, C.P. (2012). Examining the association between implementation and outcomes: State-wide scale-up of school-wide positive behavior intervention and supports. *Journal of Behavioral Health Services and Research*, 39(4), 417–433. <https://doi.org/10.1007/s11414-012-9290-2>.
- Perepletchikova, F. (2011). On the Topic of Treatment Integrity. *Clinical Psychology: Science and Practice*, 18(2), 148–153. <https://doi.org/10.1111/j.1468-2850.2011.01246.x>
- Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75(6), 829–841. <https://doi.org/10.1037/0022-006X.75.6.829>
- \*Pettigrew, J., Graham, J.W., Miller-Day, M., Hecht, M.L., Krieger, J.L., & Shin, Y.J. (2015). Adherence and delivery: Implementation quality and program outcomes for the seventh-grade keepin'it REAL program. *Prevention Science*, 16(1), 90–99. <https://doi.org/10.1007/s11121-014-0459-1>.
- Proctor, E. K., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., & Mittman, B. (2009). Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research*, 36(1), 24–34. <https://doi.org/10.1007/s10488-008-0197-4>
- \*Reyes, M.R., Brackett, M. A., Rivers, S. E., Elbertson, N. A, & Salovey, P. (2012). The interaction effects of program training, dosage, and implementation quality on targeted student outcomes for the RULER approach to social and emotional learning. *School Psychology Review*, 41(1), 82–99.
- \*Rimm-Kaufman, S.E., Larsen, R.A.A., Baroody, A.E., Curby, T.W., Ko, M., Thomas, J.B., ... DeCoster, J. (2014). Efficacy of the responsive classroom approach: Results from a 3-year, longitudinal randomized

- controlled trial. *American Educational Research Journal*, 51(3), 567–603. <https://doi.org/10.3102/0002831214523821>.
- \*Rosenblatt, J.L., & Elias, M.J. (2008). Dosage effects of a preventive social-emotional learning intervention on achievement loss associated with middle school transition. *Journal of Primary Prevention*, 29(6), 535–555. <https://doi.org/10.1007/s10935-008-0153-9>.
- Sarno, J., Lyon, A. R., Brandt, N. E., Warner, C. M., Nadeem, E., Spiel, C., & Wagner, M. (2014). Implementation science in school mental health: Key constructs in a developing research agenda. *School Mental Health*, 6(2), 99–111. <https://doi.org/10.1007/s12310-013-9115-3>
- Schoenwald, S. K., & Garland, A. F. (2013). A review of treatment adherence measurement methods. *Psychological Assessment*, 25(1), 146–156. <https://doi.org/10.1037/a0029715.A>
- Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research: Multidisciplinary perspectives on the conceptualization, measurement, and enhancement of treatment integrity. *School Psychology Review*, 38(4), 541–546.
- Schultes, M. T., Jöstl, G., Finsterwald, M., Schober, B., & Spiel, C. (2015). Measuring intervention fidelity from different perspectives with multiple methods: The reflect program as an example. *Studies in Educational Evaluation*, 47, 102–112. <https://doi.org/10.1016/j.stueduc.2015.10.001>
- Sklad, M., Diekstra, R., De Ritter, M., Ben, J., & Gravesteyn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, 49(9), 892–909. <https://doi.org/10.1002/pits.21641>
- Smith, J. D., Schneider, B. H., Smith, P. K., & Ananiadou, K. (2004). The effectiveness of whole-school antibullying programs: A synthesis of evaluation research. *School Psychology Review*, 33(4), 547–560.
- \*Smokowski, P.R., Guo, S., Wu, Q., Evans, C.B., Cotter, K.L., & Bacallao, M. (2016). Evaluating dosage effects for the positive action program: How implementation impacts internalizing symptoms, aggression, school hassles, and self-esteem. *American Journal of Orthopsychiatry*, 86(3), 310. <https://doi.org/10.1037/ort0000167>.
- \*Sørli, M.-A., & Ogden, T. (2007). Immediate impacts of PALS: A school-wide multi-level programme targeting behaviour problems in elementary school. *Scandinavian Journal of Educational Research*, 51(5), 471–492. <https://doi.org/10.1080/00313830701576581>.
- \*Sørli, M. A., & Ogden, T. (2015). School-wide positive behavior support—Norway: Impacts on problem behavior and classroom climate. *International Journal of School & Educational Psychology*, 3(3), 202–217. <https://doi.org/10.1080/21683603.2015.1060912>.
- Suldo, S. M., Gormley, M. J., DuPaul, G. J., & Anderson-Butcher, D. (2014). The impact of school mental health on student and school-level academic outcomes: Current status of the research and future directions. *School Mental Health*, 6(2), 84–98. <https://doi.org/10.1007/s12310-013-9116-2>
- Sutherland, K. S., McLeod, B. D., Conroy, M. a., & Cox, J. R. (2013). Measuring implementation of evidence-based programs targeting young children at risk for emotional/Behavioral disorders: Conceptual issues and recommendations. *Journal of Early Intervention*, 35, 129–149. <https://doi.org/10.1177/1053815113515025>
- Tobler, N. S. (1986). Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *Journal of Drug Issues*, 16(4), 537–567. <https://doi.org/10.1177/002204268601600405>
- Weare, K., & Nind, M. (2011). Mental health promotion and problem prevention in schools: What does the evidence say? *Health Promotion International*, 26(SUPL. 1). <https://doi.org/10.1093/heapro/dar075>.
- Weist, M., & Lever, N. (2014). Further advancing the field of school mental health. In M. Weist, N. Lever, C. P. Bradshaw, & J. S. Owens (Eds.), *Handbook of school Mental Health. Research, training, practice and Policy* (2nd ed., pp. 1–16). New York: Springer.
- Wheeler, J. J., Mayton, M. R., Ton, J., & Reese, J. E. (2014). Evaluating treatment integrity across interventions aimed at social and emotional skill development in learners with emotional and behaviour disorders. *Journal of Research in Special Educational Needs*, 14(3), 164–169. <https://doi.org/10.1111/j.1471-3802.2011.01229.x>
- \*Wigelsworth, M., Humphrey, N., & Lendrum, A. (2013a). Evaluation of a school-wide preventive intervention for adolescents: The secondary social and emotional aspects of learning (SEAL) programme. *School Mental Health*, 5(2), 96–109. <https://doi.org/10.1007/s12310-012-9085-x>.
- \*Wigelsworth, M., Lendrum, A., & Humphrey, N. (2013b). Assessing differential effects of implementation quality and risk status in a whole-school social and emotional learning programme: Secondary SEAL. *Mental Health and Prevention*, 1(1), 11–18. <https://doi.org/10.1016/j.mhp.2013.06.001>.
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 71(1), 136–149. <https://doi.org/10.1037/0022-006X.71.1.136>
- \*Zhai, F.H., Raver, C.C., & Jones, S.M. (2015). Social and emotional learning services and child outcomes in third grade: Evidence from a cohort of head start participants. *Children and Youth Services Review*, 56, 42–51.

**Rodrigo Rojas-Andrade** Phd (c) in Psychology. Master in Community Psychology. Psychologist. Associate Professor of Department of Psychology, University Academy of Christian Humanism. Chile.

**Loreto Leiva Bahamondes** Phd in Psychology. Master in Public Health. Psychologist. Assistant Professor of Department of Psychology, University of Chile. Chile