



# A Scoping Review of Constructs Measured Following Intervention for School Refusal: Are We Measuring Up?

David Heyne<sup>1\*</sup>, Johan Strömbeck<sup>2,3</sup>, Katarina Alanko<sup>3</sup>, Martin Bergström<sup>4</sup> and Robin Ulriksen<sup>5</sup>

<sup>1</sup> Institute of Psychology, Leiden University, Leiden, Netherlands, <sup>2</sup> Magelungen Utveckling AB, Stockholm, Sweden, <sup>3</sup> Faculty of Arts, Psychology and Theology, Åbo Akademi University, Turku, Finland, <sup>4</sup> School of Social Work, Lund University, Lund, Sweden, <sup>5</sup> BI Norwegian Business School, Oslo, Norway

## OPEN ACCESS

### Edited by:

Douglas F. Kauffman,  
Medical University of the Americas –  
Nevis, United States

### Reviewed by:

Hu Jian,  
Jiangxi University of Finance  
and Economics, China  
Candido J. Ingles,  
Miguel Hernández University of Elche,  
Spain  
Ricardo Sanmartín,  
University of Alicante, Spain

### \*Correspondence:

David Heyne  
heyne@fsw.leidenuniv.nl

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 24 December 2019

**Accepted:** 23 June 2020

**Published:** 20 August 2020

### Citation:

Heyne D, Strömbeck J, Alanko K,  
Bergström M and Ulriksen R (2020) A  
Scoping Review of Constructs  
Measured Following Intervention  
for School Refusal: Are We Measuring  
Up? *Front. Psychol.* 11:1744.  
doi: 10.3389/fpsyg.2020.01744

Reviews of the effectiveness of interventions for school refusal (SR) rely upon well-conducted primary studies. Currently there are no guidelines for those conducting primary studies about the measurement of outcome following intervention for SR. Most people would agree that it is important to measure school attendance as an outcome but there has been little discussion about other constructs that warrant measurement. To facilitate this discussion and support the development of guidelines, we conducted a scoping review of constructs measured in studies evaluating intervention for SR. We screened the title and abstract of 3,213 publications found in peer-reviewed journals between 1980 and 2019. After full text review of 271 publications, 50 publications describing 51 studies were included. Results address the frequency with which constructs were measured, along with instruments used, informants, and time-points for measurement. Based on the results, we offer guidelines for choosing constructs to measure following intervention for SR and considerations for how to measure the constructs. Guidelines can increase consistency across primary studies, with benefits for future meta-analyses and international comparisons. They also provide support for practitioners contemplating routine evaluation of their interventions for SR. Ultimately, a core outcome set for SR can be developed.

**Keywords:** school refusal, intervention, outcome, scoping review, guidelines, assessment, children, adolescents

## INTRODUCTION

When a young person<sup>1</sup> is reluctant or refuses to attend school because of emotional distress, this is referred to as school refusal (SR; Heyne et al., 2019). The emotional distress may take various forms (e.g., excessive fearfulness, depressive affect, temper tantrums, unexplained physical symptoms), and the reluctance or refusal may result in late arrival, occasionally missing whole days, or missing consecutive weeks, months, or years (Heyne et al., 2019). Because SR is often associated with absence from school, it can negatively impact academic achievement

<sup>1</sup>We use the terms young people and youth to refer collectively to children and adolescents.

(Gottfried, 2014; Gershenson et al., 2017) and socioemotional outcomes (Malcolm et al., 2003; Gottfried, 2014). School absenteeism predicts school drop-out (Schoeneberger, 2012) which is predictive of unemployment (Attwood and Croll, 2006). It is suggested that SR can greatly impact a youth's quality of life (Torrens Armstrong et al., 2011) and that families are affected when a young person has difficulty going to school (Bryce and Baird, 1986). School absenteeism also places extra burden on school staff (Thornton et al., 2013; Balu and Ehrlich, 2018).

SR occurs among 1–7% of youth in the general population and 5–16% of youth seen in clinical settings (Egger et al., 2003; Heyne and King, 2004; Steinhausen et al., 2008; Havik et al., 2015). SR is a complex problem (Ollendick and King, 1998) associated with a broad range of interacting risk factors (Ingul et al., 2019) and there is a long history of research on SR (Heyne et al., 2019). Interventions have been developed within different disciplines (Heyne, 2006) and evaluated in randomized controlled trials, non-randomized trials, multiple baseline case series, and case studies (see Heyne et al., 2002; Pina et al., 2009; Maynard et al., 2018).

There is a great need to build the knowledge base around interventions for absenteeism (Heyne, 2019) and for SR more specifically (Elliott and Place, 2019). For example, in the field of SR there are questions about the benefits of combining psychosocial and pharmacological interventions (Melvin and Gordon, 2019), the effectiveness of alternative educational programs (Brouwer-Borghuis et al., 2019), ways to improve outcomes for socially anxious youth not helped by current interventions (Heyne et al., 2015), and the long-term effects of intervention (Elliott and Place, 2019). Rigorous evaluation of interventions is needed to answer such questions (Tonge and Silverman, 2019).

Building a meaningful evidence base for SR interventions requires that those who evaluate interventions carefully consider the constructs of interest when measuring outcome. There are lists of assessment instruments and procedures for school attendance problems (Inglés et al., 2015; Kearney, 2016) and SR (Heyne and Rollings, 2002; Ingul et al., 2019) but these lists provide researchers and practitioners with minimal guidance about which constructs are most important when evaluating intervention. Narrative reviews and a systematic review of interventions for SR signal constructs of potential interest, but those reviews are limited in scope. For example, the Maynard et al. (2018) systematic review and meta-analysis of psychosocial interventions for SR reported on post-treatment school attendance and youth anxiety while other outcomes were not evaluated. Pina et al.'s (2009) narrative synthesis of the efficacy of psychosocial interventions for SR covered a broader range of outcomes, including depression and disruptive behavior. However, their review and the narrative reviews of others (Elliott and Place, 2019; Melvin and Gordon, 2019) aimed to synthesize data on the effectiveness of SR interventions and not to identify the range of constructs measured as outcomes.

There has been little discussion about which outcomes to include in the evaluation of intervention for SR, unlike in other fields (e.g., social-emotional learning; Ura et al., 2019). It is thus not surprising that there are no guidelines for the

evaluation of intervention for SR comparable to those in other fields (e.g., outcome measures recommended for people with depression and anxiety; Obbarius et al., 2017). Guidelines can enhance the evidence base for SR interventions by ensuring that important constructs are measured, in a consistent way, benefitting comparisons across studies, including future meta-analyses. This, in turn, enhances clinical decision-making. Guidelines also enhance the efficiency with which researchers and practitioners can choose constructs to measure.

The aim of the current study was to support the development of guidelines for measuring outcome following intervention for SR. The primary research question was: Which constructs have been reported in studies evaluating intervention for SR? A secondary question was: How have these constructs been measured? We conducted a review of literature across the last 40 years, undertaking a scoping review rather than a systematic review and meta-analysis. First, scoping reviews are used for reconnaissance (Peters et al., 2015), undertaking a broad review to clarify concepts in a research area, report on the nature of research activity and types of evidence being gathered, and identify gaps (Arksey and O'Malley, 2005; Peters et al., 2015). Second, we did not seek to answer clearly defined questions typically addressed via systematic review and meta-analysis such as the effectiveness of treatment based on the quantitative synthesis of empirical evidence (Pham et al., 2014; Peters et al., 2015). Nor did we examine the methodological quality of included studies, a procedure reserved for systematic reviews (Tricco et al., 2018). Scoping reviews also differ from integrative reviews inasmuch as the latter may combine data from theoretical and empirical literature (Whittemore and Knafl, 2005) which was beyond the scope of the current study.

## METHODS

### Inclusion Criteria

Studies were eligible for inclusion if they met the following criteria: (1) Language: published in Danish, Dutch, English, Finnish, Norwegian, or Swedish, the languages in which the authors are fluent; (2) Year: published between 1980 and February 2019 inclusive; (3) Type: published in a peer-reviewed journal, excluding conference abstracts and letters to the editor; (4) Accessibility: for full-text screening, studies needed to be accessible online or in libraries accessible to one of the authors; (5) Design: any study evaluating intervention for SR<sup>2</sup> (except reviews, study protocols, publications about intervention to prevent the onset of SR, and studies only addressing the prediction of outcome), such as randomized controlled trials, quasi-experimental designs, single case studies, and follow-up studies (even if the follow-up sample was included in an earlier study); and (6) Population: youth in primary or secondary school,

<sup>2</sup>No attempt was made to select or classify SR interventions according to the tiered system introduced by Kearney and Graczyk (2014) whereby emerging SR (Tier 2) is differentiated from severe or chronic SR (Tier 3) because most studies to date have provided little or no information about the severity or chronicity of SR among youth in the study samples.

between 5 and 18 years, who displayed at least the first three SR criteria presented by Heyne et al. (2019) even if other terms had been used to refer to SR (e.g., school phobia, school refusal behavior). Exclusion of studies occurred according to the order of the criteria presented (e.g., if a study fulfilled criteria 1, 2, 3, and 4, but not criterion 5, it was not screened according to criterion 6).

### Data Sources

A systematic search of 13 databases was conducted between January and February of 2019. Search terms were modified by author MB and a specialist librarian according to the database's thesaurus or subject terms (see **Appendix A**). The search yielded 6,437 publications: Academic Search Complete (1,188), Campbell Library (2), CENTRAL (272), Cinahl (672), Cochrane Database of Systematic Reviews (7), DARE (17), ERIC (319), HTA (3), PsycInfo (1,209), PubMed (1,245), Social Science Citation Index (1,238), Social Care Online (9), and SocIndex (256). This search of the 13 databases was supplemented by a search of the reference lists of published systematic reviews of intervention for SR, yielding another 17 publications. In line with the "snowball" technique (Pham et al., 2014) the authors identified 8 additional publications cited in publications already identified via the systematic search process. After removing duplicates, 3,213 publications were available for screening.

### Screening

A three-step screening process was used to establish the relevance of the 3,213 publications. Consensus meetings involving all authors were held at each step to discuss and resolve conflicts.

At Step 1, publication title and abstract were reviewed according to the eligibility criteria specified above. Each publication was independently reviewed by two researchers, with all five authors working in changing pairs. When pairs were unsure whether selection criteria were met, an inclusive approach was employed whereby the publication was included for full-text review at Step 2. This approach was considered appropriate because of the long history of confused terminology in the field of SR, whereby terms other than "school refusal" have been used to describe the phenomenon (see **Table 1** in Heyne et al., 2019). For example, publications with titles or abstracts that referred to "anxiety and school attendance/absence" or "somatization and school attendance/absence" were included for screening at Step 2. At the conclusion of Step 1, 271 publications had been selected for full-text review. Rayyan (Ouzzani et al., 2016) is a web-based program to facilitate systematic reviews, and this was used to manage selection at Step 1.

At Step 2, full-text review, two researchers independently reviewed each publication, also conducted in changing pairs. Conflicts occurred for 31 of the 271 publications reviewed (11%) and these conflicts were resolved by consensus discussion. During Step 2 it became apparent that some publications used data from a sample reported in an earlier publication. Twenty publications were reviewed a second time and eight were excluded after consensus discussions between the authors (see **Appendix B**). For example, we excluded a publication that described the longer-term functioning of adults who had refused to attend school during their youth. The authors did

not state or imply that the aim of their follow-up was to evaluate *intervention* for SR (Flakierska-Praquin et al., 1997) thus failing to fulfill inclusion criterion 5. It seemed that the aim of their study was to report on longer-term functioning in a naturalistic follow-up.

At Step 3, 33 case-related publications were re-reviewed, this time by two authors working collaboratively (DH, RU). Step 3 was included because it became apparent during Step 2 that publications about case-related material differed considerably in the extent to which outcome was reported. For example, while some publications presented empirical single case studies with a clear focus on outcome (Hagopian and Slifer, 1993) others described real or hypothetical cases simply to illustrate a particular issue such as case conceptualization, while not evaluating the intervention (Hadi et al., 2014). Case-related publications were retained if the title, abstract, or introduction stated or implied that the aim of the study was to evaluate an intervention for SR. Nine of the 33 case-related publications

**TABLE 1** | Characteristics of studies evaluating intervention for school refusal.

Characteristic	N (%)	Characteristic	N (%)
<b>Publication year</b>		<b>Sample Size</b>	
1980–1989	9 (18)	<10	27 (53)
1990–1999	15 <sup>ab</sup> (29)	10–49	15 <sup>ab</sup> (29)
2000–2009	14 <sup>b</sup> (27)	≥50	9 <sup>b</sup> (18)
2010–2019	13 (25)	<b>Mean Age<sup>d</sup></b>	
<b>Country<sup>c</sup></b>		6–9	6 <sup>b</sup> (15)
USA	19 <sup>ab</sup> (37)	10–14	29 <sup>ab</sup> (73)
Australia	10 <sup>b</sup> (20)	15–19	5 (13)
UK	7 (14)	<b>Gender (% Male)</b>	
Japan	5 (10)	≤33	10 (20)
Netherlands	3 (6)	34–66	23 <sup>ab</sup> (45)
India	2 (4)	≥67	15 (29)
China	1 (2)	Not specified	3 (6)
Finland	1 (2)	<b>Intervention</b>	
Singapore	1 (2)	Psychosocial (other than CBT)	13 (25)
Spain	1 (2)	CBT	12 <sup>b</sup> (24)
Sweden	1 (2)	Behavioral	10 (20)
<b>Language</b>		Not specified	4 (8)
English	51 (100)	Medication + other	4 (8)
Danish	0 (0)	CBT + psychosocial <sup>e</sup>	3 (6)
Dutch	0 (0)	Medication + CBT	2 <sup>b</sup> (4)
Finnish	0 (0)	Medication alone	2 <sup>a</sup> (4)
Norwegian	0 (0)	Virtual reality	1 (2)
Swedish	0 (0)	<b>Type of study</b>	
Case study	24 (47)	Case study	24 (47)
Group	19 (37)	Group	19 (37)
Follow-up only	8 (16)	Follow-up only	8 (16)

<sup>a</sup>Two studies were reported in one publication. <sup>b</sup>The sample was reported in more than one publication. <sup>c</sup>Based on the location of the first author. <sup>d</sup>Follow-up studies were excluded. <sup>e</sup>Chhabra and Puar (2016) employed psychosocial interventions alongside CBT, namely narrative therapy plus counseling with family, teachers, and peers. Last et al. (1998) compared CBT with an educational-support therapy. Tolin et al. (2009) employed CBT and other interventions as needed, such as motivational interviewing.

were excluded (see **Appendix C**). At the conclusion of Step 3, 50 publications were included for data extraction. At all steps, further duplicates were removed once identified.

## Data Extraction

An Excel spreadsheet was developed (JS) to record data extracted from the selected publications. During a consensus meeting with all authors the spreadsheet was reviewed, and minor modifications were made. Data to be included in the spreadsheet were study characteristics (e.g., year of publication, country, type of study, type of intervention evaluated), instruments used to measure outcome, and methodological characteristics (e.g., informants, measurement time-points). All instruments used at post-treatment or follow-up were assumed to be measures of outcome, because studies did not consistently state the purpose of each instrument. Authors JS, KA, and DH independently extracted data for inclusion in the spreadsheet. Consensus meetings were used to resolve uncertainties.

## Data Synthesis

The constructs that were measured following intervention were deduced from the titles of the instruments reported in the studies (authors DH, JS, RU). When unsure, authors conducted searches in Web of Science and Google to clarify the nature of the construct(s) measured via a particular instrument. Descriptive statistics (frequencies, percentages, means, standard deviations) were used to summarize the data.

## RESULTS

Fifty publications met the inclusion criteria, one of which comprised two separate studies, yielding a total of 51 studies (see **Figure 1**).

### Study Characteristics

**Table 1** summarizes the characteristics of the 51 studies. Studies were conducted in the USA (37%), Europe (25%), Australia (20%), and Asia (18%), and almost half (47%) were case studies. There were 19 group-based studies (37%), comprising randomized controlled trials (18%), non-randomized controlled trials (4%), and single arm studies with pre-test and post-test but no control group (16%)<sup>3</sup>. Eight studies (16%) only reported follow-up. The mean number of youth per study type was 36 for group-based studies, 48 for follow-up studies, and 2 for case studies. Excluding follow-up studies, the age of the youth ranged between 6 and 18 years ( $M = 12.7$ ;  $SD = 2.4$ ). Across all studies, 55% of youth were males.

Collectively, cognitive behavioral therapy (CBT; 24% of studies) and behavioral intervention (20%) were the most common types of intervention. Other psychosocial interventions were evaluated in 25% of the studies, including multimodal treatment, Morita therapy, parent counseling, collage therapy, and hypnosis. Medication was evaluated as a stand-alone intervention (4% of studies), in conjunction with CBT (4%),

or in conjunction with other interventions (8%) such as individual psychotherapy for the child and casework with parents (Berney et al., 1981).

## Constructs Measured as Outcome of Intervention

The constructs measured as outcome are found in **Table 2**. Among the 29 constructs, school attendance was the most common (73% of studies). Other relatively common constructs were anxiety (39%), depression (37%)<sup>4</sup>, emotional and behavioral symptoms (37%), global functioning (29%), fear (20%) and/or fear of school (16%), and self-efficacy (16%). Constructs measured infrequently included the function of the refusal to attend school (4%), self-esteem (4%), self-concept (2%), psychological well-being and stress (2%), quality of life (2%), and social adjustment (2%). All but two constructs pertained to characteristics of the young person, namely parent self-efficacy for managing school attendance problems, and parent desire for their child to return to school, measured via a single item.

## Methods Employed to Measure the Constructs

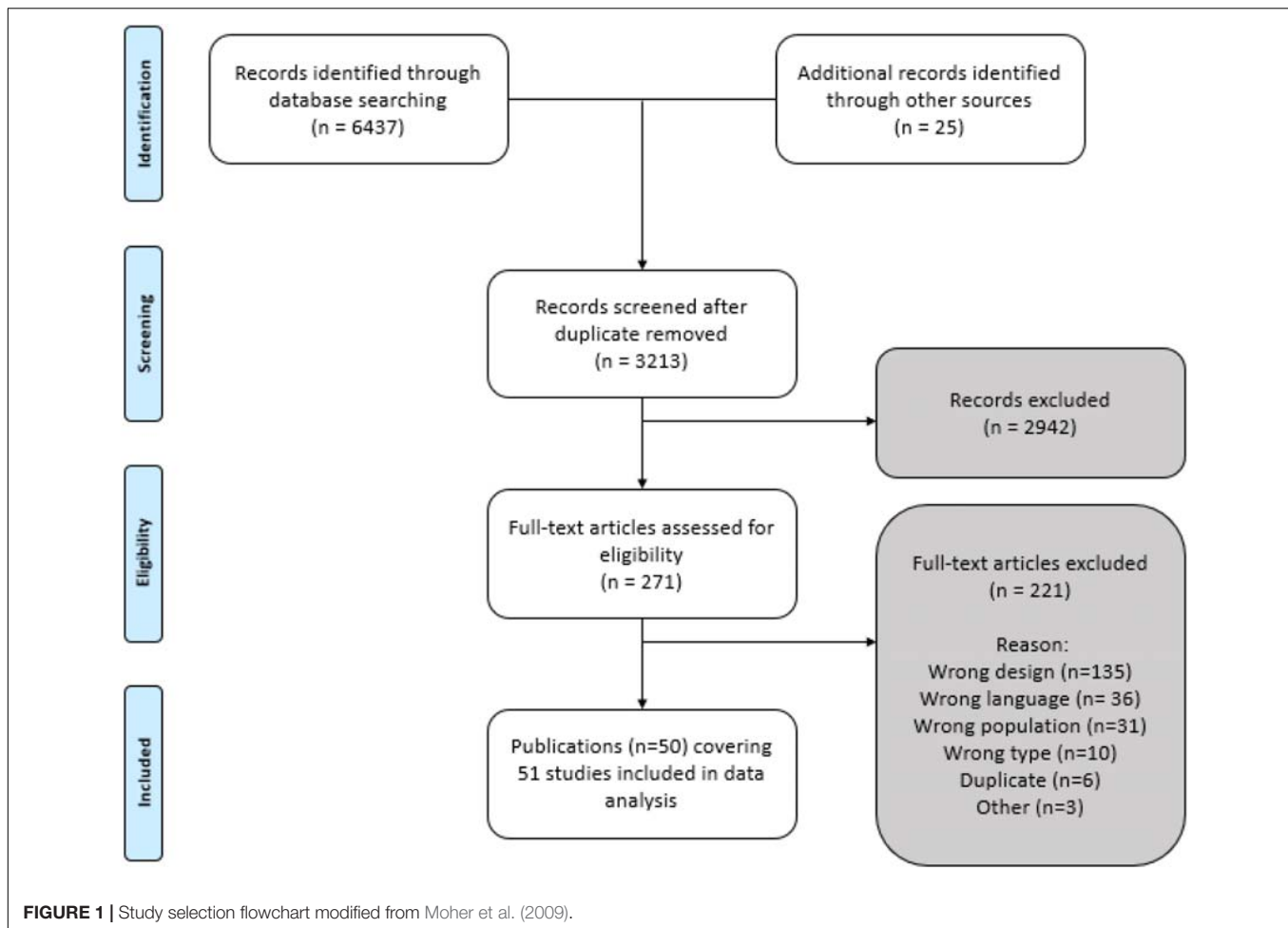
**Table 2** also presents the instruments used to measure constructs. It should be noted that school attendance was not measured via a specific instrument *per se*. Moreover, while 37 studies (73%) included school attendance as an outcome measure, information about the process for gathering attendance data was provided in just 16 of the 51 studies (31%). Attendance was reported in a variety of ways, including the number of days or weeks absent, the percentage of time absent, or via qualitative descriptions (e.g., “at 6 months Rob was spending full days in school”; Phillips and Wolpe, 1981).

The other 28 constructs were measured in many different ways. Across the 51 studies we identified 57 instruments used to measure outcome, the most common being the Child Behavior Checklist (Achenbach, 1991a) in 29% of studies, Children’s Depression Inventory (Kovacs, 1992) in 27%, Children’s Manifest Anxiety Scale or its revision (CMAS, RCMAS; Reynolds and Richmond, 1985; Reynolds and Richmond, 2008) in 24%, Fear Survey Schedule for Children (FSSC; Ollendick, 1983) in 18%, Self-Efficacy Questionnaire for School Situations (Heyne et al., 1998) in 16%, School Fear Thermometer (Heyne and Rollings, 2002) in 14%, and Global Assessment of Functioning Scale (GAF; American Psychiatric Association, 1994) in 14%.

**Table 3** summarizes the methods, informants, and time-points employed across the 51 studies. Excluding the gathering of attendance data, the most common data gathering methods were questionnaires (59%), clinician rating scales (35%), various types of interviews which were not described as diagnostic interviews (33%), other unspecified rating scales (24%), and interviews described as having a diagnostic purpose (16%). These diagnostic interviews were usually structured [e.g., Anxiety Disorders Interview Schedule for Children (ADIS-C/P); Silverman and Albano, 1996] but in one case an unstructured diagnostic

<sup>3</sup>This totals to 38% due to rounding.

<sup>4</sup>Three other studies (6%) measured both anxiety and depression with a single instrument.



interview was used (Hagopian and Slifer, 1993). Just one study, a case study (Conoley, 1987), employed an observational method to collect data, whereby a teacher observed the child in the classroom. In another case study the young person and parent were instructed to keep a diary (Chorpita et al., 1996).

With respect to informants, some studies used a multi-informant approach by eliciting information about outcome from two informants (20% of studies), three informants (25%), or four informants (6%). Respondents were youth in less than two-thirds of studies (61%), parents in less than one-half of studies (43%), and clinicians in just over one-third of studies (35%). In only nine studies (18%) did teachers or other school personnel report on outcome.

There was considerable variability in the time-points for measuring outcome. Of the 43 group-based studies and case studies (i.e., excluding the 8 studies that were solely follow-up studies), 17 (40%; 8 group-based studies and 9 case studies) involved at least three time-points for gathering data related to outcome: pre-intervention, post-intervention, and follow-up. Eleven of the 43 studies (26%) measured outcome at two time-points (pre-intervention and either post-intervention or another time-point), seven studies (16%; all case studies) included daily or weekly data gathering, and seven studies (16%; all case studies)

did not specify the time-points for measuring outcome. The eight follow-up studies, by their very nature, included measurement at some point after intervention ended, ranging from 1 to 20 years. Length of follow-up was 1–5 years in three studies (38% of follow-up studies), 6–15 years in four studies (50%), and 15–20 years in one study (12%).

## DISCUSSION

This is the first review of constructs measured following intervention for SR. Fifty-one studies met inclusion criteria: 9 studies published in the 1980s and 13–15 studies per decade across the last three decades. We discuss the constructs measured in the 51 studies, the way in which they were measured, and the strengths and limitations of the current study. Thereafter we offer guidelines for evaluating outcome following intervention for SR.

### Constructs Measured as Outcomes of Intervention

We identified 29 constructs measured as outcomes. Unsurprisingly, the construct measured most often was school attendance. Other constructs measured with moderate

**TABLE 2 |** Constructs measured after intervention for school refusal and instruments used to measure the constructs.

	<b>Group N = 19</b>	<b>FU N = 8</b>	<b>CS N = 24</b>	<b>Total N = 51</b>
<b>School attendance</b>	<b>16</b>	<b>3</b>	<b>18</b>	<b>37</b>
<b>Anxiety</b>	<b>12</b>	<b>2</b>	<b>6</b>	<b>20</b>
Children’s Manifest Anxiety Scale (CMAS/RCMAS)	7	0	5	12
State-Trait Anxiety Inventory (STAI/STAI-C)	3	1	2	6
Multidimensional Anxiety Scale for Children (MASC)	2	0	2	4
Anxiety Rating for Children (ARC/ARC-R)	2	1	0	3
Social Anxiety Scale for Children (SASC/SASC-R)	1	0	1	2
Self-Rating Anxiety Scale (SAS)	1	0	0	1
<b>Depression</b>	<b>10</b>	<b>2</b>	<b>7</b>	<b>19</b>
Children’s Depression Inventory (CDI)	8	0	6	14
Children’s Depression Rating Scale (CDRS/CDRS-R)	2	1	0	3
Beck Depression Inventory (BDI)	1	0	1	2
Self-rating Depression Scale (SDS)	1	0	1	2
Zung Depression scale	0	1	0	1
<b>Emotional and behavioral symptoms</b>	<b>9</b>	<b>2</b>	<b>8</b>	<b>19</b>
Child Behavior Checklist (CBCL)	7	1	7	15
Teacher’s Report Form (TRF)	3	0	3	6
Youth Self-Report (YSR)	1	0	1	2
Achenbach Young Adult Self-Report (YASR)	0	1	0	1
Devereux Behavior Rating Scales–School Form	0	0	1	1
Rutter Behavior Rating Scales (RBRS)	1	0	0	1
Strengths and Difficulties Questionnaire (SDQ)	1	0	0	1
Young Adult Behavior Checklist (YABCL)	0	1	0	1
<b>Global functioning</b>	<b>10</b>	<b>1</b>	<b>4</b>	<b>15</b>
Global Assessment of Functioning Scale (GAF)	6	0	1	7
Children’s Global Assessment Scale (CGAS)	4	0	0	4
Clinical Global Impression (CGI)	2	0	2	4
Subjective Units of Distress (SUDS)	1	0	1	2
Comprehensive Psychopathology Rating Scale (CPRS)	0	1	0	1
<b>Fear</b>	<b>7</b>	<b>1</b>	<b>2</b>	<b>10</b>
Fear Survey Schedule for Children (FSSC-2/FSSC-R)	7	0	2	9
Fear Questionnaire	0	1	0	1
<b>Fear of going to school / school-related fear</b>	<b>5</b>	<b>0</b>	<b>3</b>	<b>8</b>
School Fear Thermometer (SFT)	4	0	3	7
School-Related Fears Inventory <sup>a</sup> (IME)	1	0	0	1
<b>Self-efficacy for school-related situations</b>	<b>5</b>	<b>0</b>	<b>3</b>	<b>8</b>
Self-efficacy Questionnaire for School Situations (SEQ-SS)	5	0	3	8
<b>Diagnosis</b>	<b>5</b>	<b>1<sup>b</sup></b>	<b>2</b>	<b>8<sup>b</sup></b>
Anxiety Disorders Interview Schedule for Children (ADIS-C/P)	3	0	1	4
Diagnostic Interview for Children and Adolescents – Revised (DICA-R)	0	1	0	1
Diagnostic Interview Schedule for Children (NIMH DISC 2.3)	0	1	0	1
Missouri Assessment of Genetics Interview for Children (MAGIC)	1	0	0	1
Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS-P)	1	0	0	1
Unspecified diagnostic interview	0	0	1	1
<b>Anxiety and depression<sup>c</sup></b>	<b>0</b>	<b>2</b>	<b>1</b>	<b>3</b>
Leeds Anxiety and Depression Scale	0	2	0	2
Revised Child Anxiety and Depression Scale (RCIDAS)	0	0	1	1
<b>Adverse effects of medication</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>2</b>
New York state psychiatric institute side effect form	1	0	0	1
UKU-scales (side-effects)	0	0	1	1
<b>Function of refusal to go to school</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>2</b>
School Refusal Assessment Scale (SRAS)	1	0	1	2

(Continued)

TABLE 2 | Continued

	Group N = 19	FU N = 8	CS N = 24	Total N = 51
<b>Outcome of services – general health, social functioning</b>	<b>2</b>	<b>0</b>	<b>0</b>	<b>2</b>
Health of the Nation Outcomes Scales Child and Adolescent (HoNOSCA)	2	0	0	2
<b>Self-esteem</b>	<b>1</b>	<b>1</b>	<b>0</b>	<b>2</b>
Rosenberg Self Esteem Scale	0	1	0	1
Self-esteem Inventory	1	0	0	1
<b>Severity of diagnosis</b>	<b>0</b>	<b>0</b>	<b>2</b>	<b>2</b>
Clinical Severity Rating (part of ADIS)	0	0	2	2
<b>Cognitive and behavioral dimensions in motivation and engagement</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
Motivation and Engagement Scale – High School version (MES-HS)	1	0	0	1
<b>Consumer satisfaction + parent/adolescent desire for school return</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
School Refusal Program Consumer Satisfaction Questionnaire (SRP-CSQ)	1	0	0	1
<b>Daily hassles</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>1</b>
Daily Life Stressors Scale (DLSS)	0	0	1	1
<b>Dimensions of personality</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
Junior Eysenck Personality Questionnaire (JEPQ)	1	0	0	1
<b>Mental health</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>
General Health Questionnaire (GHQ 30)	0	1	0	1
<b>Overall improvement since start of intervention</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
Global Improvement Scale	1	0	0	1
<b>Parent self-efficacy for managing school attendance problems</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
Self-efficacy Questionnaire for Responding to School Attendance Problems (SEQ-RSAP)	1	0	0	1
<b>Personal functioning</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
Personal Performance Scale (PPS)	1	0	0	1
<b>Psychological well-being and stress</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
General Well-being Scale	1	0	0	1
<b>Psychopathology (dimensional)</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>
Maudsley Symptom Checklist	0	1	0	1
<b>Quality of life</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
KIDSCREEN-27	1	0	0	1
<b>Reading ability</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
Burt Reading Test (BRT)	1	0	0	1
<b>Self-concept (intrapersonal competence)</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>1</b>
Piers-Harris Self-concept Scale (P-H)	0	0	1	1
<b>Social adjustment</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>
Social Adjustment Scale	0	1	0	1

FU, Follow up; CS, Case study. <sup>a</sup>The Spanish title, reported in an English-language publication, is *Inventario de Miedos Escolares (IME)*. <sup>b</sup>One follow-up study used two different diagnostic interviews. <sup>c</sup>The constructs “anxiety” and “depression” were measured via a single instrument.

frequency were emotional and behavioral symptoms (including anxiety, fear, school-related fear, and depression), self-efficacy, and global functioning.

### School Attendance

School attendance is a pivotal measure of outcome following intervention for SR (King et al., 1998) described as a “gold standard” because it provides a real-world referent in ways that psychological rating scales do not (Tonge and Silverman, 2019). This seems to explain why school attendance was the most commonly measured construct.

It is logical that five of the eight follow-up studies did not measure attendance because many participants would have been older than school-age at the time data was gathered.

However, one quarter of the case studies did not include school attendance as an outcome even though youth in these studies were aged 6–18 years, so most were probably of school-going age when intervention was completed. Case studies may sometimes be authored by practitioners who find it more difficult to retrieve school attendance data than do research teams that have assistants to contact or visit schools to collect data.

In Maeda’s (2017) case study, attendance data was supplemented with peri-attendance information about the need for parents to escort the young person to school, probably because the therapy evaluated was Morita therapy which involves parents escorting their child to school. Maeda’s study signals the importance of not only determining whether a young person is at school but also how much or little effort

**TABLE 3** | Methodological characteristics of studies evaluating outcome of intervention for school refusal.

	Group N = 19	FU N = 8	CS N = 24	Total N = 51
<b>Measurement method</b>				
School attendance data	16	3	18	37
Questionnaire	16	5	9	30
Rating scale – clinician	12	2	4	18
Interview – other <sup>a</sup>	7	6	4	17
Rating scale – other <sup>b</sup>	6	0	6	12
Interview – diagnostic	5	1	2	8
Review of medical record	2	3	0	5
Other	3	1	1	4
Diary	0	0	1	1
Observation	0	0	1	1
Test	1	0	0	1
<b>Informants</b>				
Youth	17	5	9	31
Parent	11	2	9	22
Clinician	13	2	3	18
School personnel <sup>c</sup>	4	2	3	9
<b>Time-points</b>				
Pre-post and follow up	8	0	9	17
Only follow up	1	8	0	9
Daily/weekly	0	0	7	7
Not specified	0	0	7	7
Pre-post	5	0	1	6
Pre + after certain time <sup>d</sup>	5	0	0	5

<sup>a</sup>These interviews were not clearly for diagnostic purposes and were simply described as a structured or unstructured interview, clinical interview, telephone interview, etcetera. <sup>b</sup>Rater by informant other than clinician (e.g., youth). Teachers or school counselors. <sup>c</sup>School personnel. <sup>d</sup>Measurement at pre-intervention and then another time-point (e.g., after 4 and 8 weeks) which is not post-intervention.

is required – by the young person, parents<sup>5</sup>, and school staff – to ensure the young person is at school. Berney et al. (1981) also assessed ability to attend school, taking account of the need for parent escorting, although they noted that information was limited to arrival at school “and does not take into consideration what happens subsequent to that” (p. 112). Mansdorf and Lukens (1987) reported the percent of time two youths displaying SR could remain in school with the parent and alone, during intervention and at 3-month follow-up.

### Emotional Symptoms

Across the four decades pertinent to this study, definitions of SR have consistently specified the presence of emotional distress (Atkinson et al., 1985; Last et al., 1998; Bernstein et al., 2000; Heyne et al., 2011). It is thus not surprising that, apart from school attendance, the constructs most commonly measured relate to emotional distress (i.e., anxiety, fear, school-related fear, depression, anxiety and depression, and emotional and behavioral symptoms). The two instruments most commonly used to measure anxiety were the CMAS/RCMAS and the State-Trait Anxiety Inventory and its child version (Spielberger et al., 1973). The more recent Multidimensional Anxiety Scale for

Children (March et al., 1997) was used in 4 studies and no studies used the Screen for Child Anxiety Related Emotional Disorders (SCARED, Birmaher et al., 1999). Fear was almost always measured with the FSSC or its revision, and fear of school almost always with the SFT. Almost three-quarters of studies measuring depression used the CDI. Emotional and behavioral symptoms were most commonly measured via the CBCL for parent report and the Teacher’s Report Form (Achenbach, 1991b) for teacher report. The Youth Self-Report (Achenbach, 1991c) was only used in two studies, perhaps because of its length.

### Behavioral Symptoms

Because definitions of SR emphasize emotional distress and specify the absence of severe antisocial behavior, it may seem unsurprising that a little more than one-third of studies measured both emotional *and* behavioral symptoms. However, disorder-level oppositional behavior is reported among 21–44% of youth referred for SR (Heyne et al., 2015) suggesting the importance of measuring oppositional behavior.

### Self-Efficacy

Youth self-efficacy was measured in 16% of studies, all of which were evaluations of CBT for SR. In each study it was measured as a situation-specific construct via the SEQ-SS, not as a general self-efficacy construct. The SEQ-SS measures a young person’s

<sup>5</sup>We use the term “parents” to refer to parents and/or other caregivers.



perception of their ability to cope with school-related situations such as doing school-work, being away from parents, and answering questions about absence. It has been suggested that low self-efficacy for responding to school situations poses a risk for SR and, conversely, high self-efficacy may help explain school attendance even when a young person faces difficult situations at school (Ingul et al., 2019).

One study measured parent self-efficacy using the relatively recent Self-Efficacy Questionnaire for Responding to School Attendance Problems (SEQ-RSAP; Heyne et al., 2007). The study revealed a significant increase in parent self-efficacy following a CBT intervention which included parents (Heyne et al., 2011) providing initial support for measuring the construct of parent self-efficacy when intervention is conducted with parents.

### Global Functioning, Mental Health, and Diagnosis

Less than one-third of studies measured youths' global functioning as an outcome, most commonly via the GAF. Mental health is a similarly broad construct, and it was measured in two studies. In four studies the diagnostic status of emotional distress was assessed via the ADIS-C/P, and only four other studies used a diagnostic interview schedule to measure outcome. We observed that a focus on overall adjustment was less typical of case studies than group-based studies. Those conducting group-based studies are typically from research settings where global functioning and diagnosis are commonly assessed.

### Combinations of Constructs

One study combined outcomes related to attendance and diagnosis to determine the proportion of youth no longer fulfilling operational criteria for SR. Specifically, Heyne et al. (2011) reported that, subsequent to intervention, 55% of youth attended school more than 80% of the time *and* no longer had a diagnosable level of anxiety. This seems to be the first effort to conceptualize outcome as the absence of multiple SR criteria.

### Constructs Seldom Studied

We could expect that authors measure particular constructs following intervention because those constructs reflect the goals of the intervention. Similarly, constructs measured during intervention and at the end of intervention could be expected to reflect an author's theory of change. However, during full-text review of studies we noticed that few authors presented a rationale for measuring the constructs embodied in the instruments they used. More often, authors presented a rationale for choosing a specific instrument (e.g., its psychometric properties; suitability for a specific age group). It is thus unclear from this review whether the lack of attention to specific constructs should be interpreted as intentional (e.g., the author believed that the construct was not important enough to be included in a lengthy assessment battery) or unintentional (e.g., the author overlooked the importance of a construct).

We identified numerous constructs that were seldom measured. In some instances, the lack of attention to a construct is understandable. For example, only two studies used the School

Refusal Assessment Scale (Kearney and Silverman, 1993) or its revision (Kearney, 2002b) to measure outcome. This instrument was designed to facilitate intervention planning by indicating the function of the refusal to attend school; it was not designed as a measure of outcome. Only two studies measured self-esteem and one study measured self-concept. This might be explained in part by the fact that a large number of studies included in our review evaluated CBT or behavioral intervention. Traditionally, interventions using CBT-based theory and techniques were not focused on raising levels of self-esteem, and if measured, self-esteem was a secondary outcome rather than primary outcome (Kolubinski et al., 2018).

The lack of attention to other constructs is more surprising. Only one study used an instrument focused on social adjustment despite the fact that this construct is linked to SR historically (Buitelaar et al., 1994; Place et al., 2002; Egger et al., 2003) and recently (Ingul and Nordahl, 2013; Blöte et al., 2015; Heyne et al., 2015). We found indirect measurement of social adjustment via global ratings of functioning (e.g., the GAF), a single broad item rated by youth, parents, or clinicians [e.g., "peer relationships" in the HoNOSCA (Gowers et al., 1999)] and questionnaires which simultaneously measured a broad range of constructs (e.g., the CBCL includes a subscale for social problems). However, data from instruments such as the CBCL were usually reported at the broad-band level (i.e., internalizing behavior and externalizing behavior).

No studies measured family functioning as an outcome despite the fact that one-half to two-thirds of families of youth who display SR exhibit maladaptive family functioning, and CBT manuals for SR commonly include family-related work on communication and problem-solving (Heyne et al., 2015). In Bernstein et al. (2001) 1-year naturalistic follow-up study, the Family Adaptability and Cohesion Evaluation Scale II (FACES II; Olson et al., 1982) was used to measure cohesion, adaptability, and family type (balanced to extreme), but only as a predictor of outcome at follow-up.

Only two parent-related constructs were measured as outcomes: self-efficacy in one study, and desire for the child to return to school in another study. Measures of parenting styles and dimensions were not included in any studies, despite the potential impact of parenting on the outcomes of intervention for SR (Heyne et al., 2015). Prabhuswamy et al. (2007) presented pre-intervention data on psychosocial factors such as parental overindulgence and overprotection, but these constructs were not measured post-intervention. Furthermore, no studies reported on parent psychopathology despite the fact that it is often observed in the parents of youth displaying SR (Heyne et al., 2015).

In two recent studies of alternative educational settings for youth displaying SR, authors included positively-oriented constructs. Preece and Howley (2018) noted that staff in an intervention facility for youth unable to attend mainstream school regarded youths' well-being as essential for re-engagement with formal education, and the General Well-Being Scale (Heubeck and Neill, 2000) was used to monitor youth-reported progress in well-being. McKay-Brown et al. (2019)

argued that professionals' attention to youths' quality of life, in the context of a wraparound model of care, could enhance SR interventions. Quality of life was measured as an outcome via youth and parent reports on the KIDSCREEN-27 (Kidscreen Group Europe, 2006) which assesses youths' health and well-being.

## How the Constructs Were Measured

The gathering and reporting of attendance data varied considerably. For example, data was derived via parents' weekly reports of the number of hours in school (Bernstein et al., 2000) and daily records of attendance kept by teachers (King et al., 1998). Many authors did not specify the source of attendance data so readers cannot assess the reliability of outcomes based on that data. There was also variability in the way attendance data was reported, such as the number of days or weeks youth were absent; the percentage of absence in a period of time (e.g., across 2 weeks; across 4 weeks); achievement of a specific level of school attendance (e.g., at least 80 or 90%); and descriptions such as "at 6 months Rob was spending full days in school" (Phillips and Wolpe, 1981).

Across all constructs, respondents were youth in less than two-thirds of studies, parents in less than one-half of studies, clinicians in just over one-third of studies, and teachers in under one-fifth of studies. Differences emerge when analyzing the data according to study type. For example, while 17 of the 19 group-based studies (89%) elicited youth report, only 9 of the 24 case studies (38%) did so. It is possible that youth had been consulted during outcome evaluation in case studies, but a failure to specify the instrument(s) used to measure outcome led to an underestimation in the current study of the extent to which youth were the informants on outcomes. For example, Hargett and Webster (1996) reported that a young person's adjustment was monitored over 7 months and that there were no signs of SR, but there was no specification of the data source.

School attendance data aside, the most common methods for data gathering were questionnaires, rating scales completed by clinicians or others, and some form of interview whether for diagnostic or other purposes. We assume questionnaires were regularly used because they are easy to administer, score, and interpret. Even though interviews provide rich qualitative information about the lived experiences of participants, which is important for the development of evidence-based practice (American Psychological Association, 2006), quantitative information grants the most efficient method for comparing results across studies. Just one study used an observational method to gather data following intervention, and just one used a daily diary, despite the recommendation that observations and diaries be used in the assessment of SR (Ollendick and King, 1998).

The time-points for measuring outcome were varied, especially among the group-based studies. The impression we gained during full-text review is that authors rarely if ever justified the time-points they used, so it is difficult to explain the variability. Hargett and Webster (1996) used an ongoing approach to measure outcome, collecting bi-weekly data for 7 months until the end of the school year, and then for the

first 2 months of the next school year. This ongoing approach to measurement – albeit easier to conduct in a single case study relative to a group-based study – enabled the authors to demonstrate ongoing efficacy of the intervention, including no relapse back to refusal to attend school or an inability to stay at school after arrival.

## Strengths and Limitations

The scoping review method employed in the current study helped clarify which constructs have been measured as outcomes following SR intervention. We conducted a broad search across four decades, five languages, and various study types (group, case, and follow-up studies). The inclusion of case studies strengthens the relevance of this review for practice-based settings, beyond its relevance for research-based settings.

Unpublished studies were not included, which may have limited the range of constructs identified. At the same time, by restricting our search to peer-reviewed publications we incorporated a crude quality assessment check on the included studies. There was no further assessment of study quality because the aim was to review existing literature according to constructs and methods for measuring constructs, and not to synthesize evidence about intervention effectiveness (Pham et al., 2014).

As is typical of reviews, judgment was used to determine whether studies should be included or excluded. In our review, this included judgment about whether or not case studies met the additional inclusion criterion (i.e., "stated or implied intention to evaluate outcome"), and judgment about which constructs were being measured by the instruments used to evaluate outcome. The latter was necessary because the authors of studies rarely specified which constructs they intended to measure. Judgments were made in pairs, and if there was doubt the research team met for a consensus discussion.

We observed that the SR criteria reported in some studies were unclear. We thus excluded some studies that may well have been evaluations of intervention for SR. To assist future reviews, authors should specify which criteria for SR were (not) applicable to the youth in their study.

It was beyond the scope of this study to review the psychometric properties of the instruments identified across the 51 studies. A review of this kind will benefit decision-making about how best to measure constructs of interest. Existing reviews that contain reliability and validity information about instruments used to measure constructs such as "school engagement" (Fredricks and McColskey, 2012) could be used to guide a review of the psychometric properties of instruments used to measure the constructs included in the guidelines that follow.

## Recommendations

Following, we offer guidelines for measuring outcome following interventions for SR. Greater standardization of outcome measurement – which constructs are measured and how they are measured – facilitates comparison of outcomes across studies and the synthesis of data via meta-analysis. Ultimately, consumers of SR interventions benefit from greater standardization. That is, standardization enhances the accumulation of evidence about

the relative benefits of different options for intervention, and practitioners and researchers are more likely to measure and report outcomes that are important to the users of their research (Chu et al., 2015; Kirkham et al., 2016). Guidelines also aid efficiency in decision-making about outcome evaluation. For busy practitioners, this may increase the likelihood that they routinely evaluate progress so as to determine when and how the “scaffolding of support” to youth, families, and schools can be reduced.

### Guidelines for Choosing Constructs

According to Ollendick and King (1998) professionals often gathered information in certain ways simply because those ways were convenient. We contend that the choice of outcome constructs should be based on the relevance of the constructs to the goals of intervention and not influenced by convenience or habit (e.g., using an instrument because it is familiar).

If a researcher’s or practitioner’s goal is to “simply” help youth return to school (which is seldom a simple process), it might seem logical to limit outcome constructs to school attendance. However, SR is heterogeneous in its etiology and presentation (Heyne and Sauter, 2013; Gallé-Tessonneau and Heyne, 2020) as well as its impact, so intervention should focus on improving broader outcomes for youth, necessitating a wider palette of outcome constructs. In other words, the goals of intervention are likely to include general goals (e.g., increased school attendance) and specific goals informed by case formulation (e.g., increased social involvement).

Assuming researchers and practitioners choose to measure multiple outcomes, which should they be? In a review of SR intervention, Elliott and Place (2019) noted: “Researchers, therefore, need to be explicit about whether the primary outcome sought in their intervention studies is reduction in anxiety or increased school attendance.” Based on the findings in our review and our own reflections on the goals of intervention, we propose that evaluation of outcome includes – but also extends beyond – the constructs of school attendance and anxiety. If we assume that the constructs measured in the 51 studies reviewed here were chosen because they reflected the goals of the interventions offered and not because of convenience or habit, then researchers and practitioners are well advised to measure the more common constructs identified in the current study: school attendance; emotional functioning including anxiety, fear/fear of school, and depression; behavioral symptoms; global functioning; and self-efficacy.

School attendance is an important foundational competency for youth (Kearney et al., 2019a) and a gold-standard, real-world referent for evaluating interventions for school attendance problems (Tonge and Silverman, 2019). It is self-evident that it would be included in evaluations of SR. As noted in section “School Attendance,” peri-attendance variables are also relevant. A smartphone application could be developed to facilitate ecological momentary assessment of variables beyond school attendance and absence, such as the young person’s whereabouts (e.g., in class or often in the school nurse’s office), how much time parents spend at school with a separation anxious youth

participating in intervention for SR, and the youth’s emotional distress during the school day.

Emotional functioning includes the youth’s levels of fear, anxiety, and depression. Alongside general levels of emotional distress, researchers and practitioners may measure distress experienced within the school setting. Three relatively recent instruments not yet incorporated in outcome studies have face validity for SR intervention: the School Anxiety Scale–Teacher Report (Lyneham et al., 2008), the School Anxiety Inventory<sup>6</sup> (SAI; García-Fernández et al., 2011) and its short version (García-Fernández et al., 2014), and the School REfusal EvaluationN Scale (SCREEN; Gallé-Tessonneau and Gana, 2019). These provide more detailed information about youths’ emotional distress in the school context relative to instruments such as the MASC and its revision (MASC-2; March, 2013), the Revised Children’s Manifest Anxiety Scale (Reynolds and Richmond, 1985) and its revision (RCMAS-2; Reynolds and Richmond, 2008), and the Spence Children’s Anxiety Scale (Spence, 1998). The SCARED has a 4-item school phobia subscale but the subscale does not always emerge in analyses of the instrument’s factor structure (Inglés et al., 2015). Because psychosomatic symptoms are also prominent in cases of SR (Heyne et al., 2015) they should be measured alongside the other constructs associated with emotional distress.

Measures of behavioral symptoms can provide an indication of the frequency and severity of a young person’s resistance to school attendance. Parents can be asked to complete a daily logbook that includes ratings of noncompliance and disruption (Kearney and Albano, 2000) and more specifically the child’s resistance to efforts to get them to go to school (Kurita, 1991). Anecdotal evidence suggests that the decrease in a young person’s resistance to school attendance is a very important outcome for parents who are often emotionally and physically exhausted due to resistive behaviors often directed at them.

Global functioning provides a measure of overall outcome following SR intervention. It was infrequently incorporated in case studies included in this review. To benchmark global functioning of youth included in case studies against those included in group-based studies, authors preparing case studies will need to incorporate a measure of global functioning. The most common measure of global functioning was the clinician-rated GAF, which includes assessment of the impact of symptoms on daily life. It is important to also obtain youth, parent, and teacher perspectives on global functioning, perhaps via the impact supplement of the Strengths and Difficulties Questionnaire (Goodman, 1999).

Youth self-efficacy and parent self-efficacy are cognitive constructs relevant to outcome because they are a key target for change during CBT for SR and they have been found to increase following intervention (Heyne et al., 2015). Non-CBT interventions may also have a positive effect on school attendance due to youths’ increased self-efficacy for attending school and/or parents’ increased self-efficacy for responding to a child’s refusal

<sup>6</sup>This is an updated version of the Inventario de Miedos Escolares (IME) presented in Table 2. The IME only assesses anxiety-generating situations whereas the SAI also assesses the three anxiety response systems (cognitive, physiological, and motor or behavioral).

to attend. As such, it is valuable to include youth and parent measures of self-efficacy in the evaluation of CBT and non-CBT interventions. Furthermore, higher levels of self-efficacy at post-intervention may help prevent relapse during a follow-up period because youth and/or parents are more likely to engage in adaptive behaviors during those times when there are small setbacks.

Other constructs found to be less common or not measured at all also warrant attention. Following, we draw attention to constructs relevant to the young person's adjustment; motivation for change; family functioning; and side effects of intervention. We conclude this section with attention to individualized goals as constructs of interest, and variation in the constructs of interest according to the length of follow-up conducted.

An important facet of the young person's adjustment is their social adjustment (see section "Constructs Seldom Studied"). An instrument measuring this construct should be used, or at the very least authors should report outcome according to subscales that measure social functioning (e.g., "social problems" in the CBCL). We also encourage researchers and practitioners to measure well-being as a broader youth-focused construct, reflecting a holistic perspective on the success of intervention. School engagement is an important construct because of the potential for relapse if there is little school engagement following return to school. McKay-Brown et al. (2019) used the Motivation and Engagement Scale (Martin, 2014) to measure "educational functioning," noting that the instrument measures adaptive and maladaptive factors related to learning behaviors "that are linked to school engagement" (p. 96). According to Amai (2020) a related construct of "sense of school adaptation" has been widely measured in Japan using Furuichi and Tamaki's (1994) School Adaptation Scale (e.g., I look forward to going to school; I want to go to school even if I feel a little bad). Other youth-focused constructs that appear to be associated with SR but have not been measured as outcomes include emotion regulation (Hughes et al., 2010) and negative automatic thoughts and thinking styles (Maric et al., 2012). Academic functioning is an important construct when evaluating interventions for absenteeism (Tonge and Silverman, 2019), including SR.

Parent and youth motivation for achieving school return was measured in just one study via a single item about desire for return to school (Melvin et al., 2017). Readiness for change is found to be related to outcome in studies of psychotherapy with adults and adolescents (Krebs et al., 2018) and seems important for understanding treatment progress among depressed adolescents (Rodriguez-Quintana and Lewis, 2019). It should receive more attention in studies of intervention for SR, with measurements at pre-intervention, mid-intervention, and post-intervention, as recommended by Rodriguez-Quintana and Lewis (2019) in relation to adolescent depression.

A systemic perspective on SR and interventions for SR calls for measurement of parenting and family functioning (see section "Constructs Seldom Studied"). Measuring these constructs during intervention, at post-intervention, and at follow-ups will allow us to establish the extent of change in functioning as well as the extent to which change in parent and family functioning is associated with change in school

attendance and other outcomes for youth. As noted, no studies included in the current review addressed these constructs. If the word-limit restrictions of journals lead authors to exclude data about parent and family functioning, such information should be included in supplementary online materials or adjunct publications so that the evidence base for changes in these constructs grows.

A few studies of pharmacological intervention measured adverse effects, and some studies of psychosocial intervention measured the experience of intervention. In Head and Jamieson's (2006) study, which was excluded from our review<sup>7</sup>, the focus of enquiry was broadened from "technical success" in terms of school attendance to whether and why youth, parents, and teachers regarded intervention as successful. Maeda (2012) provided a qualitative account of the impact of intervention (forced school attendance) on the parents: "Thus, intensive exposure therapy for school return could be a burden to children, parents, and school officials in spite of being effective for school return" (p. 309). Information about the experience of intervention may impact consumer uptake and persistence with intervention. Just one study included in our review reported on consumer satisfaction as an outcome. Consumer satisfaction warrants inclusion in all reports on outcome because it provides information that can help shape interventions in ways that enhance uptake, persistence, and outcome.

Individualized goals for intervention constitute important constructs to be measured. In Meyer et al. (1999) case study, "progress toward goals" was measured according to youth and parent behaviors. This approach to conceptualizing change may have been used because the young person had an intellectual disability. Nonetheless, it presents a model for all practitioners evaluating interventions, and where possible, for the evaluation of outcome in group-based studies. Another example of attention to specific goals is found in the work of Kearney and Silverman (1990). They used a standard set of outcome measures across seven cases but they expected differences on the measures per case, depending on the function served by each youth's refusal to attend school. It was suggested that "perhaps a more appropriate way to examine the data is to focus primarily on those measures that are pertinent to each functional category" (p. 354). Indeed, goal-based outcomes derived from youth and parent goals for intervention should be considered for inclusion in a battery of outcome measures (Law and Jacob, 2015).

The constructs of interest will also vary according to the length of follow-up being conducted. For example, in Flakierska et al. (1988) 15- to 20-year follow-up study of adults who had refused to attend school as youth, constructs included the number of visits to adult outpatient psychiatric care and the number of children they had. Clearly, longer follow-ups call for broader conceptualization of the constructs of interest.

Whichever constructs are measured, authors need to provide a clear rationale for choosing those constructs. This is in contrast to what we observed during the current review. For example,

<sup>7</sup>The study did not meet inclusion criterion 5; it was not an evaluation of intervention but a report on the experience of intervention.

interviews were used in many studies, presumably to measure diagnosis or global functioning, but a clear explanation for why an interview was used was not always included. Our observation that authors often neglected to specify the constructs of interest reflects an unfortunate long-standing phenomenon. Even in the 1980s Valles and Oddy (1984) noted that “measures of outcome are ill-defined” (p. 36). When the specification of constructs of interest becomes standard practice, authors are prompted to reflect on the goals they have for their intervention. Specification of constructs will also benefit future reviews of the effectiveness of interventions for SR. First, it will expedite the selection of primary studies for systematic review because the outcomes of interest are clearly specified in the primary studies. Second, it can facilitate the interpretation of outcomes from meta-analyses. For example, effect sizes might differ between primary studies because of differences in the conceptualization and operationalization of outcome, whereby some measured constructs may be more sensitive to change than others.

### Guidelines for Measuring Constructs

To advance the evidence base for SR interventions, careful consideration needs to be given to the methods for measuring chosen constructs. First, consistency in the choice of instrument or procedure to measure a construct such as school attendance will facilitate the comparison of outcomes across schools, school districts, and states (Hobbs et al., 2018), as well as across countries. For example, Fredricks and McColskey (2012) noted large variation in how constructs were measured in the field of school engagement, making it difficult to compare findings across studies. Second, when other methodological issues are consistent across studies, such as choice of informants and measurement time-points, then the interpretation of comparative results will be simplified. For example, Melvin and Gordon (2019) conducted a review of antidepressant medication for SR and suggested that study differences in the source of information (e.g., school versus parent) and the timing of assessment created measurement error variance which could explain the apparent lack of benefit when combining medication with CBT for SR. Third, greater consistency in how outcome data is reported will enhance data synthesis such as meta-analysis. In a meta-analysis of the relationship between anxiety and school attendance problems, Finning et al. (2019) concluded that methodological differences across studies limited scope for combining studies. Their study was not a meta-analysis of outcome, but the challenges they experienced in synthesizing data across studies would apply equally to meta-analyses of outcome following intervention for SR.

Following, we discuss six topics relevant to promoting consistency in the evaluation of outcome: accessing valid data on school attendance; using psychometrically sound instruments to measure other constructs; establishing uniformity in the choice of psychometrically sound instruments; establishing uniformity in time-points for measurement; specifying criteria for determining when desired outcome has been achieved; and incorporating various sources of data.

First, those who evaluate outcome need access to school attendance data that accurately describes a young person’s

attendance and non-attendance. Because schools are increasingly held to account for the registration of attendance and non-attendance (Hutt and Gottfried, 2019; OFSTED, 2019) it is reasonable to expect that researchers and practitioners could acquire school-based data per half-day, and ideally per lesson. This can be converted to a percentage of school attendance for a specified time-frame [e.g., number of lessons (or hours) attended in a 4 week period divided by the total number of lessons (or hours) scheduled in that time-frame]. When possible, information should also be gathered about peri-attendance variables such as late arrival to school and absence from a lesson whilst still at school (e.g., spending time in the school counselor’s office). Some schools may not (yet) collect such detailed data. To fill this gap, parents and youth can be asked to record attendance as well as peri-attendance variables. Presumably youth can provide a more accurate account of absence from class during attendance at school. To increase parent and youth compliance with the request for data, researchers and practitioners might use automated reminders (e.g., smartphone applications) or diaries managed via email contact. At the end of each week during the post-intervention and follow-up time-frames, youth and parents could be asked to list how many classes were held and which classes were attended. If possible, this would occur at the end of each day to reduce problems with recall at the end of the week. Because there are discrepancies in school- and parent-reported absences (Lomholt et al., 2020) and school- and youth-reported absences (Keppens et al., 2019) we recommend gathering data from all three sources and reporting outcomes for each group separately. Similarities and differences in outcomes based on youth-, parent-, and school-reported absences need to be taken into account when authors discuss the effectiveness of the intervention.

Second, constructs need to be measured via instruments with strong psychometric properties. Simply put, does the instrument measure the construct it is supposed to measure and in a reliable way? Important psychometric properties include construct validity, internal consistency reliability, test-retest reliability, and sensitivity to treatment effects (Spence, 2018). Psychometrically strong instruments are needed for group-wise significance tests when evaluating intervention effects in group-based studies. For case studies, two methods which practitioners can use to analyze clinically meaningful change are “crossing clinical thresholds” and “reliable change” (Wolpert et al., 2015). Such analyses also rely on measurement via psychometrically sound instruments. Other characteristics that render instruments more suitable for measuring outcome following intervention include applicability to a wide age range (e.g., to compare intervention effects for younger versus older students); availability for different respondents (e.g., youth, parents, teachers); availability of normative data and clinical cut-off scores (e.g., to determine whether change is clinically meaningful); and the time taken to administer the instrument. A more fundamental issue is the need to specify the instruments used to gather outcome data; this was not always the case in the studies we reviewed.

Third, when a majority of researchers and practitioners use the same instrument to measure a specific construct, the

comparison of outcomes across studies will be more robust. It will also be easier to combine data for meta-analyses, avoiding the methodological complications that occur when different measures of the same construct are used to create a common outcome measure (Bennett et al., 2013). As the field achieves greater consensus on which instruments are most valuable, efforts should be made to ensure instruments are available in many languages via thorough translation and adaptation processes that include forward and backward translation (Van Widenfelt et al., 2005).

Fourth, there should be more uniformity in the time-points for measuring constructs. To illustrate, one study assessed school attendance across the 4 weeks following intervention (Melvin et al., 2017) while earlier studies did so across 2 weeks (Heyne et al., 2011). We contend that the 4-week period provides a better test of the reliability of intervention effects with respect to school attendance, even though a 2-week time-frame is commonly employed in criteria for deciding whether a school attendance problem exists (Kearney, 2008). It is logical to apply the 4-week time-frame to the measurement of all chosen constructs. For example, 4 weeks after delivery of the final component of intervention, researchers or practitioners administer instruments to measure constructs such as anxiety and depression. If the CDI is used to measure depression, then 4 weeks after intervention finished the young person is invited to report symptoms of depression across the prior 2 weeks (the standard time-frame prescribed for administration of the CDI). Longer-term outcome is ideally based on follow-ups at 6 and 12 months after intervention finishes. This permits more robust evaluation of the young person's adjustment to the ongoing academic and social-emotional challenges of life at school. When interventions are prolonged (e.g., placement within an alternative educational setting), 6- and 12-months evaluations can take place during the course of the intervention, instead of waiting until a 12-month program is complete before evaluating its effects. Finally, alongside standard time-frames for measurement, researchers and practitioners may choose to conduct additional measurement at other time-points.

Fifth, there needs to be consistency in criteria for determining whether desired outcome is achieved, and clear specification by authors of the criteria they used. This will address Kearney and Silverman's (1990) observation 30 years ago that many studies of intervention for absenteeism used inconsistent or inadequate criteria for positive outcome. The constructs used to measure desired outcome will depend on the stated aims of the intervention. We found that many case studies included qualitative descriptions of outcome (e.g., "attendance and anxiety remained at acceptable levels"; Kearney, 2002a) whereas the group-based studies analyzed outcomes quantitatively (e.g., change in mean level of anxiety), although we also found qualitative descriptions in some group-based studies (e.g., "in general, youth no longer exhibited upset on arrival at school"; King et al., 2001). It is incumbent upon authors to report the proportion of cases fulfilling a specified outcome rather than relying on non-specific terms such as "in general" and "typically." With respect to school attendance, authors can report on the proportion of youth reaching a specific level of attendance,

alongside their reports of the average amount of attendance in the 4-week period since intervention finished. Chronic or persistent absenteeism is increasingly specified as 10% absence or more in a given time-frame (Heyne et al., in press) so a standard criterion for desired outcome would be attendance above 90%. To be able to compare outcomes across case studies and group-based studies, case studies could include a minimum level of quantitative data (e.g., whether or not the young person achieved more than 90% attendance at post-intervention). Authors can also consider combining constructs to determine the proportion of youth who simultaneously fulfill two or more criteria for desired outcome (see section "Combinations of Constructs").

Sixth, all stakeholders in interventions for SR – youth, parents, education professionals, and helping professionals – should be invited to report on the outcome of intervention to ensure a breadth of perspectives on outcome. In a meta-analysis of five decades of research on psychological interventions for youth, Weisz et al. (2017) argued that "it matters a lot" who reports on outcome, based on the observation that effect sizes differed across informants (p. 94). They emphasized the need for researchers and practitioners to obtain and integrate information from multiple informants and to be explicit about the source of outcome data. This emphasis on multi-source and multi-method assessment is not new to the field of school attendance and absenteeism (Ollendick and King, 1998; Kearney, 2002a). All stakeholders should also be consulted about the constructs that ought to be measured as outcomes following intervention for SR.

## CONCLUSION

Are we measuring up? In other words, are we as researchers and practitioners measuring outcome in a way that helps to build a meaningful evidence base for SR intervention?

With respect to constructs measured, there has been some consistency across studies but also considerable variability. School attendance is the only construct that was measured in more than two-thirds of the studies. The fact that other "common" constructs (i.e., emotional and behavioral symptoms, anxiety, fear and/or fear of school, depression, self-efficacy, and global functioning) were measured in a third of studies or less might reflect variability in what was regarded as important to measure, but it might also reflect a failure to consider the benefit of measuring such constructs. It is unlikely that the low rate at which these "common" constructs were measured was due to the unavailability of instruments because there have been instruments to measure the majority of these constructs since the 1980s. Unfortunately, authors infrequently provided justification for the choice of constructs measured. With respect to the way in which constructs were measured, there was also substantial variability. It is important that authors of future studies clearly specify the rationale for focusing on specific constructs and for using specific instruments to measure those constructs.

Despite current shortcomings in the evaluation of outcome following SR intervention, the current review yields initial guidelines for researchers and practitioners planning to evaluate

outcome. Measurement of the more common constructs identified in this review (i.e., attendance; emotional, behavioral, and global functioning; self-efficacy) can be supplemented with measurement of the young person's social adjustment and the well-being of the young person, parents, and family. These guidelines may yield greater uniformity in the evaluation of interventions, benefitting science and practice, and thus the youth, parents, and schools impacted by SR.

The current review can also serve as a platform for further work on the development of a core outcome set for SR, possibly via the international consensus-based process presented in COS-STAP (Core Outcome Set-STANDARDISED Protocol Items; Kirkham et al., 2019). The fact that SR is not included as a disorder in classification systems such as the DSM does not negate the need for a core outcome set. Work on a core outcome set should foster broad stakeholder input, broader than the perspective of the authors of this paper. Attention should be paid to changes in education and technology, such as competency-based education and virtual learning (Kearney et al., 2019b) necessitating a reconceptualization of traditional outcomes such as percentage of time spent at school.

Attention should also be paid to the development of a core outcome set for interventions focused on truancy, the school attendance problem characterized by “skipping” or absconding from school (Heyne et al., 2019). In a systematic review and meta-analysis of the effects of interventions for truancy, Maynard et al. (2013) found considerable variability in how study authors operationalized and reported outcomes related to school attendance, and there was sometimes a lack of clarity about what had been measured (e.g., excused and/or unexcused absences). Maynard and colleagues called for greater consistency in measuring and reporting school attendance when evaluating truancy interventions. Broadening the perspective, Keppens and Spruyt (2020) concluded an integrative review of interventions

to prevent truancy with a call to evaluate outcomes other than truancy-related absence, such as graduation in the longer-term. Their review also signals the need to think more broadly about truancy, not simply as a behavior to be changed but also as a symptom of the need for change in school bonding.

In conclusion, measuring up to the task of advancing the science and practice of SR intervention requires greater consensus on the evaluation of outcome. This review contributes to the discussion about guidelines for evaluating outcome. Before a core outcome set becomes available we encourage researchers and practitioners to carefully consider and justify their choice of constructs and measurement methods. A collective effort is also needed to increase consistency in the choice of psychometrically sound instruments for measuring important constructs.

## AUTHOR CONTRIBUTIONS

All authors were involved in designing the study, conducting the review, and preparing the manuscript.

## FUNDING

This work was conducted independently of a research grant.

## ACKNOWLEDGMENTS

We thank Maja Carlsson, specialist librarian in the Social Sciences Faculty Library at Lund University in Sweden, for her guidance in sourcing data. We also thank Mikael Thastum and Daniel Bach Johnsen, Department of Psychology and Behavioural Sciences at Aarhus University in Denmark, for their suggestions about measuring outcome.

## REFERENCES

- Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.
- Achenbach, T. M. (1991c). *Manual for the Youth Self-Report and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.
- Amai, K. (2020). Variables affecting the school adaptation of secondary-school students who do not seek help: attachment, coping style, positivity, and prospects. *Int. J. Adolesc. Youth* 25, 687–702. doi: 10.1080/02673843.2020.1717559
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn. Washington, DC: Author.
- American Psychological Association (2006). Presidential Task Force on Evidence-Based Practice. Evidence-based practice in psychology. *Am. Psychol.* 61, 271–285. doi: 10.1037/0003-066X.61.4.271
- \*Anderson, J., King, N., Tonge, B., Rollings, S., Young, D., and Heyne, D. (1998). Cognitive-behavioural intervention for an adolescent school refuser: a comprehensive approach. *Behav. Change* 15, 67–73. doi: 10.1017/s0813483900003181
- Arksey, H., and O'Malley, L. (2005). Scoping studies: towards a methodological framework. *Int. J. Soc. Res. Methodol.* 8, 19–32. doi: 10.1080/1364557032000119616
- Atkinson, L., Quarrington, B., and Cyr, J. J. (1985). School refusal: the heterogeneity of a concept. *Am. J. Orthopsychiatry* 55, 83–101. doi: 10.1111/j.1939-0025.1985.tb03423.x
- Attwood, G., and Croll, P. (2006). Truancy in secondary school pupils: prevalence, trajectories and pupil perspectives. *Res. Pap. Educ.* 20, 467–484. doi: 10.1080/02671520600942446
- Balu, R., and Ehrlich, S. B. (2018). Making sense out of incentives: a framework for considering the design, use, and implementation of incentives to improve attendance. *J. Educ. Stud. Placed Risk* 23, 93–106. doi: 10.1080/10824669.2018.1438898
- \*Beidas, R. S., Crawley, S. A., Mychailyszyn, M. P., Comer, J. S., and Kendall, P. C. (2010). Cognitive-behavioral treatment of anxious youth with comorbid school refusal: clinical presentation and treatment response. *Psychol. Top.* 19, 255–271.
- Bennett, K., Manassis, K., Walter, S., Cheung, A., Wilansky-Traynor, P., Diaz-Granados, N., et al. (2013). Cognitive behavioral therapy age effects in child and adolescent anxiety: an individual patient data meta-analysis. *Depress. Anxiety* 30, 829–841. doi: 10.1002/da.22099
- \*Berg, I., and Jackson, A. (1985). Teenage school refusers grow up: a follow-up study of 168 subjects, ten years on average after in-patient treatment. *Br. J. Psychiatry* 147, 366–370. doi: 10.1192/bjp.147.4.366
- \*Berney, T., Kolvin, I., Bhate, S. R., Garside, R. F., Jeans, J., Kay, B., et al. (1981). School phobia: a therapeutic trial with clomipramine and short-term outcome. *Br. J. Psychiatry* 138, 110–118. doi: 10.1192/bjp.138.2.110

- \*Bernstein, G. A., Borchardt, C. M., Perwien, A. R., Crosby, R. D., Kushner, M. G., Thuras, P. D., et al. (2000). Imipramine plus cognitive-behavioral therapy in the treatment of school refusal. *J. Am. Acad. Child Adolesc. Psychiatry* 39, 276–283. doi: 10.1097/00004583-200003000-00008
- \*Bernstein, G. A., Garfinkel, B. D., and Borchardt, C. M. (1990). Comparative studies of pharmacotherapy for school refusal. *J. Am. Acad. Child Adolesc. Psychiatry* 29, 773–781. doi: 10.1097/00004583-199009000-00016
- \*Bernstein, G. A., Hektner, J. M., Borchardt, C. M., and McMillan, M. H. (2001). Treatment of school refusal: one-year follow-up. *J. Am. Acad. Child Adolesc. Psychiatry* 40, 206–213. doi: 10.1097/00004583-200102000-00015
- Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S., and Baugher, M. (1999). Psychometric properties of the screen for child anxiety related emotional disorders (SCARED): a replication study. *J. Am. Acad. Child Adolesc. Psychiatry* 38, 1230–1236. doi: 10.1097/00004583-199910000-00011
- \*Blagg, N. R., and Yule, W. (1984). The behavioural treatment of school refusal: a comparative study. *Behav. Res. Ther.* 22, 119–127. doi: 10.1016/0005-7967(84)90100-1
- Blöte, A. W., Miers, A. C., Heyne, D. A., and Westenberg, P. M. (2015). “Social anxiety and the school environment of adolescents,” in *Social Anxiety and Phobia in Adolescents: Development, Manifestation and Intervention Strategies*, eds K. Ranta, A. M. La Greca, L.-J. Garcia-Lopez, and M. Marttunen (Switzerland: Springer International Publishing), 151–181. doi: 10.1007/978-3-319-16703-9\_7
- Brouwer-Borghuis, M. L., Heyne, D., Sauter, F. M., and Scholte, R. H. (2019). The Link: an alternative educational program in the Netherlands to reengage school-refusing adolescents with schooling. *Cogn. Behav. Pract.* 26, 75–91. doi: 10.1016/j.cbpra.2018.08.001
- Bryce, G., and Baird, D. (1986). Precipitating a crisis: family therapy and adolescent school refusers. *J. Adolesc.* 9, 199–213. doi: 10.1016/s0140-1971(86)80003-3
- \*Buitelaar, J. K., van Andel, H., Duyx, J. H. M., and van Strien, D. C. (1994). Depressive and anxiety disorders in adolescence: a follow-up study of adolescents with school refusal. *Acta Paedopsychiatr.* 56, 249–253.
- \*Chhabra, T., and Puar, S. (2016). Social support in psychotherapy of panic disorder with agoraphobia in school setting. *Indian J. Health Wellbeing* 7, 648–650.
- \*Chorpita, B. F., Albano, A. M., Heimberg, R. G., and Barlow, D. H. (1996). A systematic replication of the prescriptive treatment of school refusal behavior in a single subject. *J. Behav. Ther. Exp. Psychiatry* 27, 281–290. doi: 10.1016/s0005-7916(96)00023-7
- \*Chu, B. C., Rizvi, S. L., Zendegui, E. A., and Bonavitaola, L. (2015). Dialectical behavior therapy for school refusal: treatment development and incorporation of web-based coaching. *Cogn. Behav. Pract.* 22, 317–330. doi: 10.1016/j.cbpra.2014.08.002
- \*Conoley, J. C. (1987). Strategic family intervention: three cases of school-aged children. *Sch. Psychol. Rev.* 16, 469–486.
- Egger, H. L., Costello, J. E., and Angold, A. (2003). School refusal and psychiatric disorders: a community study. *J. Am. Acad. Child Adolesc. Psychiatry* 42, 797–807. doi: 10.1097/01.chi.0000046865.56865.79
- Elliott, J. G., and Place, M. (2019). Practitioner review: school refusal: developments in conceptualisation and treatment since 2000. *J. Child Psychol. Psychiatry* 60, 4–15. doi: 10.1111/jcpp.12848
- Finning, K., Ukoumunne, O. C., Ford, T., Danielson-Waters, E., Shaw, L., Romero De Jager, I., et al. (2019). The association between anxiety and poor attendance at school: a systematic review. *Child Adolesc. Ment. Health* 24, 205–216. doi: 10.1111/camh.12322
- \*Flakierska, N., Lindström, M., and Gillberg, C. (1988). School refusal: a 15–20-year follow-up study of 35 Swedish urban children. *Br. J. Psychiatry* 152, 834–837. doi: 10.1192/bjp.152.6.834
- Flakierska-Praquin, N., Lindström, M., and Gillberg, C. (1997). School phobia with separation anxiety disorder: a comparative 20- to 29-year follow-up study of 35 school refusers. *Compr. Psychiatry* 38, 17–22. doi: 10.1016/s0010-440x(97)90048-1
- Fredricks, J. A., and McColskey, W. (2012). “The measurement of student engagement: a comparative analysis of various methods and student self-report instruments,” in *Handbook of Research on Student Engagement*, eds S. Christenson, A. Reschly, and C. Wylie (Boston, MA: Springer), 763–782. doi: 10.1007/978-1-4614-2018-7\_37
- Furuichi, Y., and Tamaki, H. (1994). Factors contributing to the enjoyment of junior high school life. *Bull. Faculty Educ. Okayama Univ.* 96, 105–113.
- Gallé-Tessoneau, M., and Heyne, D. (2020). Behind the SCREEN: identifying school refusal themes and sub-themes. *Emot. Behav. Diff.* 25, 139–154. doi: 10.1080/13632752.2020.1733309
- Gallé-Tessoneau, M., and Gana, K. (2019). Development and validation of the school refusal evaluation scale for adolescents. *J. Pediatr. Psychol.* 44, 153–163. doi: 10.1093/jpepsy/jsy061
- García-Fernández, J. M., Inglés, C. J., Martínez-Monteaquedo, M. C., Marzo, J. C., and Estévez, E. (2011). Inventario de Ansiedad Escolar: validación en una muestra de estudiantes de Educación Secundaria. *Psicothema* 23, 301–307.
- García-Fernández, J. M., Inglés, C. J., Marzo, J. C., and Martínez-Monteaquedo, M. C. (2014). Psychometric properties of the school anxiety inventory-short version in Spanish secondary education students. *Psicothema* 26, 286–292. doi: 10.7334/psicothema2013.288
- Gershenson, S., Jacknowitz, A., and Brannegan, A. (2017). Are student absences worth the worry in US primary schools? *Educ. Finance Policy* 12, 137–165. doi: 10.1162/edfp\_a\_00207
- \*Goh, C. W. (1989). School refusal: clinical features and treatment outcome. *Singapore Med. J.* 30, 550–552.
- Goodman, R. (1999). The extended version of the strengths and difficulties questionnaire as a guide to child psychiatric caseness and consequent burden. *J. Child Psychol. Psychiatry* 40, 791–801.
- \*Gosschalk, P. O. (2004). Behavioral treatment of acute onset school refusal in a 5-year old girl with separation anxiety disorder. *Educ. Treat. Child.* 27, 150–160.
- Gottfried, M. A. (2014). Chronic absenteeism and its effects on students’ academic and socioemotional outcomes. *J. Educ. Stud. Placed Risk* 19, 53–75. doi: 10.1080/10824669.2014.962696
- Gowers, S. G., Harrington, R. C., Whitton, A., Lelliott, P., Wing, J., Beevor, A., et al. (1999). Brief scale for measuring the outcomes of emotional and behavioural disorders in children: health of the Nation outcome scales for children and adolescents (HoNOSCA). *Br. J. Psychiatry* 174, 413–416. doi: 10.1192/bjp.174.5.413
- \*Gutiérrez-Maldonado, J., Magallón-Neri, E., Rus-Calafell, M., and Peñaloza-Salazar, C. (2008). Virtual reality exposure therapy for school phobia. *Anuario Psicol.* 40, 223–236.
- Hadi, S. R. A., Razak, A. H. A., Kalil, E. Z., and Ismail, W. S. W. (2014). Parental issues contributing to school refusal: a case report. *ASEAN J. Psychiatry* 15, 83–85.
- \*Hagopian, L. P., and Slifer, K. J. (1993). Treatment of separation anxiety disorder with graduated exposure and reinforcement targeting school attendance: a controlled case study. *J. Anxiety Disord.* 7, 271–280. doi: 10.1016/0887-6185(93)90007-8
- \*Hargett, M. Q., and Webster, R. E. (1996). Treatment integrity and acceptability with families: a case study of a child with school refusal. *Psychol. Sch.* 33, 319–324. doi: 10.1002/(sici)1520-6807(199610)33:4<319::aid-pits6>3.0.co;2-h
- Havik, T., Bru, E., and Ertesvåg, S. K. (2015). Assessing reasons for school non-attendance. *Scand. J. Educ. Res.* 59, 316–336. doi: 10.1080/00313831.2014.904424
- Head, G., and Jamieson, S. (2006). Taking a line for a walk: including school refusers. *Pastoral Care Educ.* 24, 32–40. doi: 10.1111/j.1468-0122.2006.00377.x
- Heubeck, B. G., and Neill, J. T. (2000). Confirmatory factor analysis and reliability of the Mental Health Inventory for Australian adolescents. *Psychol. Rep.* 87, 431–440. doi: 10.2466/PRO.87.6.431-440
- Heyne, D. (2006). “School refusal,” in *Practitioner’s Guide to Evidence-Based Psychotherapy*, eds J. E. Fisher and W. T. O’Donohue (New York, NY: Springer), 599–618.
- Heyne, D. (2019). Developments in classification, identification, and intervention for school refusal and other attendance problems: introduction to the special series. *Cogn. Behav. Pract.* 26, 1–7. doi: 10.1016/j.cbpra.2018.12.003
- Heyne, D., Gren-Landell, M., Melvin, G., and Gentle-Genitty, C. (2019). Differentiation between school attendance problems: why and how? *Cogn. Behav. Pract.* 26, 8–34. doi: 10.1016/j.cbpra.2018.03.006
- Heyne, D., Kearney, C. A., and Finning, K. (in press). “Introduction,” in *Mental Health and Attendance at School*, eds T. Ford, K. Finning, and D. Moore (Cambridge, MA: Cambridge University Press).



- Heyne, D., King, N., Tonge, B., Rollings, S., Pritchard, M., Young, D., et al. (1998). The self-efficacy questionnaire for school situations: development and psychometric evaluation. *Behav. Change* 15, 31–40. doi: 10.1017/S08134839000588X
- Heyne, D., and King, N. J. (2004). “Treatment of school refusal,” in *Handbook of Interventions That Work with Children and Adolescents: Prevention and Treatment*, eds P. M. Barrett and T. H. Ollendick (Chichester: John Wiley & Sons, Ltd), 243–272.
- \*Heyne, D., King, N. J., Tonge, B. J., Rollings, S., Young, D., Pritchard, M., et al. (2002). Evaluation of child therapy and caregiver training in the treatment of school refusal. *J. Am. Acad. Child Adolesc. Psychiatry* 41, 687–695. doi: 10.1097/00004583-200206000-00008
- Heyne, D., Maric, M., and Westenberg, P. M. (2007). *Self-Efficacy Questionnaire for Responding to School Attendance Problems. Unpublished Measure*. Leiden: Leiden University.
- Heyne, D., and Rollings, S. (2002). *School Refusal*. Malden: Blackwell.
- Heyne, D., and Sauter, F. M. (2013). “School Refusal,” in *The Wiley-Blackwell Handbook of The Treatment of Childhood and Adolescent Anxiety*, eds C. A. Essau, and T. H. Ollendick (Hoboken, NJ: John Wiley & Sons), 471–517.
- Heyne, D., Sauter, F. M., and Maynard, B. R. (2015). “Moderators and mediators of treatments for youth with school refusal or truancy,” in *Moderators and Mediators of Youth Treatment Outcomes*, eds M. Maric, P. J. M. Prins, and T. H. Ollendick (Oxford: Oxford University Press), 230–266. doi: 10.1093/med:psych/9780199360345.003.0010
- \*Heyne, D., Sauter, F. M., Ollendick, T. H., Van Widenfelt, B. M., and Westenberg, P. M. (2014). Developmentally sensitive cognitive behavioral therapy for adolescent school refusal: rationale and case illustration. *Clin. Child Fam. Psychol. Rev.* 17, 191–215.
- \*Heyne, D., Sauter, F. M., Van Widenfelt, B. M., Vermeiren, R., and Westenberg, P. M. (2011). School refusal and anxiety in adolescence: non-randomized trial of a developmentally sensitive cognitive behavioral therapy. *J. Anxiety Disord.* 25, 870–878. doi: 10.1016/j.janxdis.2011.04.006
- Hobbs, A., Kotlaja, M., and Wylie, L. (2018). Absenteeism interventions: an approach for common definitions in statewide program evaluations. *Justice Eval. J.* 1, 215–232. doi: 10.1080/24751979.2018.1517584
- Hughes, E. K., Gullone, E., Dudley, A., and Tonge, B. (2010). A case-control study of emotion regulation and school refusal in children and adolescents. *J. Early Adolesc.* 30, 691–706. doi: 10.1177/0272431609341049
- Hutt, E. L., and Gottfried, M. A. (2019). “Conclusion,” in *Absent from School: Understanding and Addressing Student Absenteeism*, eds M. A. Gottfried and E. L. Hutt (Cambridge, MA: Harvard Education Press), 213–220.
- Inglés, C. J., González-Maciá, C., García-Fernández, J. M., Vicent, M., and Martínez-Monteagudo, M. C. (2015). Current status of research on school refusal. *Eur. J. Educ. Psychol.* 8, 37–52.
- Ingul, J. M., Havik, T., and Heyne, D. (2019). Emerging school refusal: a school-based framework for identifying early signs and risk factors. *Cogn. Behav. Pract.* 26, 46–62. doi: 10.1016/j.cbpra.2018.03.005
- Ingul, J. M., and Nordahl, H. M. (2013). Anxiety as a risk factor for school absenteeism: What differentiates anxious school attenders from non-attenders?. *Ann. Gen. Psychiatry* 12:25. doi: 10.1186/1744-859X-12-25
- \*Kearney, C. A. (2002a). Case study of the assessment and treatment of a youth with multifunction school refusal behaviour. *Clin. Case Stud.* 1, 67–80. doi: 10.1177/1534650102001001006
- Kearney, C. A. (2002b). Identifying the function of school refusal behavior: a revision of the school refusal assessment scale. *J. Psychopathol. Behav. Assess.* 24, 235–245. doi: 10.1023/A:1020774932043
- Kearney, C. A. (2008). An interdisciplinary model of school absenteeism in youth to inform professional practice and public policy. *Educ. Psychol. Rev.* 20, 257–282. doi: 10.1007/s10648-008-9078-3
- Kearney, C. A. (2016). *Managing School Absenteeism at Multiple Tiers: An Evidence-Based and Practical Guide for Professionals*. Oxford: Oxford University Press.
- Kearney, C. A., and Albano, A. M. (2000). *When Children Refuse School: A Cognitive-Behavioral Therapy Approach, Parent Workbook*. San Antonio, TX: The Psychological Corporation.
- Kearney, C. A., González, C., Graczyk, P. A., and Fornander, M. J. (2019a). Reconciling contemporary approaches to school attendance and school absenteeism: toward promotion and nimble response, global policy review and implementation, and future adaptability (Part 1). *Front. Psychol.* 10:2222. doi: 10.3389/fpsyg.2019.02222
- Kearney, C. A., González, C., Graczyk, P. A., and Fornander, M. J. (2019b). Reconciling contemporary approaches to school attendance and school absenteeism: toward promotion and nimble response, global policy review and implementation, and future adaptability (Part 2). *Front. Psychol.* 10:2605. doi: 10.3389/fpsyg.2019.02605
- Kearney, C. A., and Graczyk, P. (2014). A response to intervention model to promote school attendance and decrease school absenteeism. *Child Youth Care Forum* 43, 1–25. doi: 10.1007/s10566-013-9222-1
- \*Kearney, C. A., Pursell, C., and Alvarez, K. (2001). Treatment of school refusal behavior in children with mixed functional profiles. *Cogn. Behav. Pract.* 8, 3–11. doi: 10.1016/s1077-7229(01)80037-7
- \*Kearney, C. A., and Silverman, W. K. (1990). A preliminary analysis of a functional model of assessment and treatment for school refusal behavior. *Behav. Modif.* 14, 340–366. doi: 10.1177/01454455900143007
- Kearney, C. A., and Silverman, W. K. (1993). Measuring the function of school refusal behavior: the school refusal assessment scale. *J. Clin. Child Psychol.* 22, 85–96. doi: 10.1207/s15374424jccp2201\_9
- Keppens, G., and Spruyt, B. (2020). The impact of interventions to prevent truancy: a review of the research literature. *Stud. Educ. Eval.* 65:100840. doi: 10.1016/j.stueduc.2020.100840
- Keppens, G., Spruyt, B., and Dockx, J. (2019). Measuring school absenteeism: administrative attendance data collected by schools differ from self-reports in systematic ways. *Front. Psychol.* 10:2623. doi: 10.3389/fpsyg.2019.02623
- Kidscreen Group Europe (2006). *The KIDSCREEN Questionnaires: Quality of Life Questionnaires for Children and Adolescents*. Lengerich: Pabst Science Publishers.
- \*King, N., Tonge, B. J., Heyne, D., Turner, S., Pritchard, M., Young, D., et al. (2001). Cognitive-behavioural treatment of school-refusing children: maintenance of improvement at 3- to 5-year follow-up. *Scand. J. Behav. Therapy* 30, 85–90.
- \*King, N. J., Tonge, B. J., Heyne, D., Pritchard, M., Rollings, S., Young, D., et al. (1998). Cognitive-behavioral treatment of school-refusing children: a controlled evaluation. *J. Am. Acad. Child Adolesc. Psychiatry* 37, 395–403. doi: 10.1097/00004583-199804000-00017
- \*King, N. J., Tonge, B. J., Turner, S., Heyne, D., Pritchard, M., Rollings, S., et al. (1999). Brief cognitive-behavioural treatment for anxiety-disordered children exhibiting school refusal. *Clin. Psychol. Psychother.* 6, 39–45. doi: 10.1002/(sici)1099-0879(199902)6:1<39::aid-cpp178>3.0.co;2-w
- Kirkham, J. J., Gorst, S., Altman, D. G., Blazeby, J. M., Clarke, M., Devane, D., et al. (2016). Core outcome set - STAndards for reporting: the COS-STAR statement. *PLoS Med.* 13:e1002148. doi: 10.1371/journal.pmed.1002148
- Kirkham, J. J., Gorst, S., Altman, D. G., Blazeby, J. M., Clarke, M., Tunis, S., et al. (2019). Core outcome set-STANDARDISED protocol items: the COS-STAP statement. *Trials* 20:116. doi: 10.1186/s13063-019-3230-x
- Kolubinski, D. C., Frings, D., Nikèevia, A. V., Lawrence, J. A., and Spada, M. M. (2018). A systematic review and meta-analysis of CBT interventions based on the Fennell model of low self-esteem. *Psychiatry Res.* 267, 296–305. doi: 10.1016/j.psychres.2018.06.025
- Kovacs, M. (1992). *Children's Depression Inventory*. New York, NY: Multi-Health Systems, Inc.
- Krebs, P., Norcross, J. C., Nicholson, J. M., and Prochaska, J. O. (2018). Stages of change and psychotherapy outcomes: a review and meta-analysis. *J. Clin. Psychol.* 74, 1964–1979. doi: 10.1002/jclp.22683
- Kurita, H. (1991). School refusal in pervasive developmental disorders. *J. Autism. Dev. Disord.* 21, 1–15. doi: 10.1007/bf02206993
- \*Last, C. G., Hansen, C., and Franco, N. (1998). Cognitive-behavioral treatment of school phobia. *J. Am. Acad. Child Adolesc. Psychiatry* 37, 404–411. doi: 10.1097/00004583-199804000-00018
- Law, D., and Jacob, J. (2015). *Goals and Goal Based Outcomes (GBOs): Some Useful Information*, third Edn. London: CAMHS Press.
- \*Lepola, U., Leinonen, E., and Koponen, H. (1996). Citalopram in the treatment of early-onset panic disorder and school phobia. *Pharmacopsychiatry* 29, 30–32. doi: 10.1055/s-2007-979539
- Lomholt, J. J., Johnsen, D. B., Silverman, W. K., Heyne, D., Jeppesen, P., and Thastum, M. (2020). Feasibility study of Back2School, a modular cognitive behavioral intervention for youth with school attendance problems. *Front. Psychol.* 11:586. doi: 10.3389/fpsyg.2020.00586

- Lyneham, H. J., Street, A. K., Abbott, M. J., and Rapee, R. M. (2008). Psychometric properties of the school anxiety scale-teacher report (SAS-TR). *J. Anxiety Disord.* 22, 292–300. doi: 10.1016/j.janxdis.2007.02.001
- \*Maeda, N. (2012). The effects of active parental intervention on school refusal behaviour. *Prog. Soc. Welfare Res.* 7, 15–25.
- \*Maeda, N. (2016). School-based behavioural consultation for school-refusal behaviour. *Int. J. Psychol. Stud.* 8, 40–52.
- \*Maeda, N. (2017). Effectiveness of Morita Therapy-based consultation for a school-refusing adolescent with psychogenic fever. *J. Educ. Train. Stud.* 5, 1–7.
- \*Maeda, N., Hatada, S., Sonoda, J., and Takayama, I. (2012). School-based intensive exposure therapy for school refusal behavior. *Clin. Case Stud.* 11, 299–311. doi: 10.1177/1534650112457456
- Malcolm, H., Wilson, V., Davidson, J., and Kirk, S. (2003). *Absence from School: A Study of its Causes and Effect in Seven LEAs*. Nottingham, UK: Queen's Printers.
- \*Mansdorf, I. J., and Lukens, E. (1987). Cognitive-behavioral psychotherapy for separation anxious children exhibiting school phobia. *J. Am. Acad. Child Adolesc. Psychiatry* 26, 222–225. doi: 10.1097/00004583-198703000-00018
- March, J. S. (2013). *Multidimensional Anxiety Scale for Children*, 2nd Edn. Toronto: Multi-Health Systems.
- March, J. S., Parker, J. D. A., Sullivan, K., Stallings, P., and Conners, C. K. (1997). The multidimensional anxiety scale for children (MASC): factor structure, reliability, and validity. *J. Am. Acad. Child Adolesc. Psychiatry* 36, 554–565. doi: 10.1097/00004583-199704000-00019
- Maric, M., Heyne, D. A., de Heus, P., van Widenfelt, B. M., and Westenberg, P. M. (2012). The role of cognition in school refusal: an investigation of automatic thoughts and cognitive errors. *Behav. Cogn. Psychother.* 40, 255–269. doi: 10.1017/s1352465811000427
- Martin, A. J. (2014). *The Motivation and Engagement Scale*, 14th Edn. Sydney: Lifelong Achievement Group.
- Maynard, B. R., Heyne, D., Brendel, K. E., Bulanda, J. J., Thompson, A. M., and Pigott, T. D. (2018). Treatment for school refusal among children and adolescents: a systematic review and meta-analysis. *Res. Soc. Work Pract.* 28, 56–67. doi: 10.1177/1049731515598619
- Maynard, B. R., McCrear, K. T., Pigott, T. D., and Kelly, M. S. (2013). Indicated truancy interventions for chronic truant students: a campbell systematic review. *Res. Soc. Work Pract.* 23, 5–21. doi: 10.1177/1049731512457207
- \*McCune, N., and Hynes, J. (2005). Ten year follow-up of children with school refusal. *Irish J. Psychol. Med.* 22, 56–58. doi: 10.1017/s0790966700008946
- \*McKay-Brown, L., McGrath, R., Dalton, L., Graham, L., Smith, A., Ring, J., et al. (2019). Reengagement with education: a multidisciplinary home-school-clinic approach developed in Australia for school-refusing youth. *Cogn. Behav. Pract.* 26, 92–106. doi: 10.1016/j.cbpra.2018.08.003
- \*McShane, G., Bazzano, C., Walter, G., and Barton, G. (2007). Outcome of patients attending a specialist educational and mental health service for social anxiety disorders. *Clin. Child Psychol. Psychiatry* 12, 117–124. doi: 10.1177/135910450701090
- \*Melvin, G. A., Dudley, A. L., Gordon, M. S., Klimkeit, E., Gullone, E., Taffe, J., et al. (2017). Augmenting cognitive behavior therapy for school refusal with Fluoxetine: a randomized controlled trial. *Child Psychiatry Hum. Dev.* 48, 485–497. doi: 10.1007/s10578-016-0675-y
- Melvin, G. A., and Gordon, M. S. (2019). Antidepressant medication: is it a viable and valuable adjunct to cognitive-behavioral therapy for school refusal? *Cogn. Behav. Pract.* 26, 107–118. doi: 10.1016/j.cbpra.2018.07.005
- \*Meyer, E. A., Hagopian, L. P., and Paclawsky, J. (1999). A function-based treatment for school refusal behavior using shaping and fading. *Res. Dev. Disabil.* 20, 401–410. doi: 10.1016/s0891-4222(99)00021-9
- \*Moffitt, C. E., Chorpita, B. F., and Fernandez, S. N. (2003). Intensive cognitive-behavioral treatment of school refusal behaviour. *Cogn. Behav. Pract.* 10, 51–60. doi: 10.1016/s1077-7229(03)80008-1
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Prisma Group, R. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Med.* 6:e1000097. doi: 10.1371/journal.pmed1000097
- Obbarius, A., van Maasakkers, L., Baer, L., Clark, D. M., Crocker, A. G., de Beurs, E., et al. (2017). Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM Depression and Anxiety Working Group. *Qual. Life Res.* 26, 3211–3225. doi: 10.1007/s11136-017-1659-5
- OFSTED (2019). *The Education Inspection Framework*. Manchester: Ofsted.
- Ollendick, T. H. (1983). Reliability and validity of the Revised Fear Survey Schedule for Children (FSSCR). *Behav. Res. Ther.* 21, 685–692. doi: 10.1016/0005-7967(83)90087-6
- Ollendick, T. H., and King, N. J. (1998). Assessment practices and issues with school-refusing children. *Behav. Change* 15, 16–30. doi: 10.1017/S0813483900005878
- Olson, D. H., Portner, J., and Bell, R. Q. (1982). *Family Adaptability and Cohesion Evaluation Scale II*. St. Paul, MN: University of Minnesota.
- Ouzzani, M., Hammady, H., Fedorowicz, Z., and Elmagarmid, A. (2016). Rayyan - a web and mobile app for systematic reviews. *Syst. Rev.* 5:210. doi: 10.1186/s13643-016-0384-384
- Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., and Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *Int. J. Evid. Based Healthc.* 13, 141–146. doi: 10.1097/xeb.0000000000000050
- Pham, M. T., Rajiæ, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., and McEwen, S. A. (2014). A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Res. Synth. Methods* 5, 371–385. doi: 10.1002/jrsm.1123
- \*Phillips, D., and Wolpe, S. (1981). Multiple behavioral techniques in severe separation anxiety of a twelve-year-old. *J. Behav. Ther. Exp. Psychiatry* 12, 329–332. doi: 10.1016/0005-7916(81)90073-2
- Pina, A. A., Zerr, A. A., Gonzales, N. A., and Ortiz, C. D. (2009). Psychosocial interventions for school refusal behavior in children and adolescents. *Child Dev. Perspect.* 3, 11–20. doi: 10.1111/j.1750-8606.2008.00070.x
- Place, M., Hulsmeier, J., Davis, S., and Taylor, E. (2002). The coping mechanisms of children with school refusal. *J. Res. Spec. Educ. Needs* 2, 1–10. doi: 10.1111/j.1471-3802.2002.00167.x
- \*Prabhuswamy, M., Srinath, S., Girimaji, S., and Seshadri, S. (2007). Outcome of children with school refusal. *Indian J. Pediatr.* 74, 375–379. doi: 10.1007/s12098-007-0063-5
- \*Preece, D., and Howley, M. (2018). An approach to supporting young people with autism spectrum disorder and high anxiety to re-engage with formal education - the impact on young people and their families. *Int. J. Adolesc. Youth* 23, 468–481.
- Reynolds, C. R., and Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale: Second Edition*. Los Angeles: Western Psychological Services.
- Reynolds, C. R., and Richmond, B. O. (2008). *Revised Children's Manifest Anxiety Scale*. Los Angeles, CA: Western Psychological Services.
- \*Roberts, D. (1998). The use of hypnosis and brief strategic therapy with a case of separation anxiety and school refusal. *Contemp. Hypnosis* 15, 219–222. doi: 10.1002/ch.138
- Rodriguez-Quintana, N., and Lewis, C. C. (2019). Ready or not? Transitions of depressed adolescents during acute phase of treatment. *Child Psychiatry Hum. Dev.* 50, 950–959. doi: 10.1007/s10578-019-00895-4
- \*Rollings, S., King, N., Tonge, B., Heyne, D., and Young, D. (1998). Cognitive-behavioural intervention with a depressed adolescent experiencing school attendance difficulties. *Behav. Change* 15, 87–97. doi: 10.1017/S081348390000320X
- Schoeneberger, J. A. (2012). Longitudinal attendance patterns: developing high school dropouts. *Clear. House J. Educ. Strateg. Issues Ideas* 85, 7–14. doi: 10.1080/00098655.2011.603766
- Silverman, W. K., and Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV, Child and Parent Versions*. San Antonio, TX: Psychological Corporation.
- Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behav. Res. Ther.* 36, 545–566. doi: 10.1016/s0005-7967(98)00034-5
- Spence, S. H. (2018). Assessing anxiety disorders in children and adolescents. *Child Adolesc. Ment. Health* 23, 266–282. doi: 10.1111/camh.12251
- Spielberger, C. D., Edwards, C. D., Lushene, R., Montuori, J., and Platzek, D. (1973). *State-Trait Anxiety Inventory for Children: Sampler Set, Manual, Test Booklet, Scoring Key*. Palo Alto, CA: Mind Garden.
- Steinhausen, H.-C., Müller, N., and Metzke, C. W. (2008). Frequency, stability and differentiation of self-reported school fear and truancy in a community sample. *Child Adolesc. Psychiatry Ment. Health* 2:17.

- \*Takata, Y. (2002). Supporting by a nurse teacher in a school infirmary using collage therapy. *Psychiatry Clin. Neurosci.* 56, 371–379. doi: 10.1046/j.1440-1819.2002.01025.x
- Thornton, M., Darmody, M., and McCoy, S. (2013). Persistent absenteeism among Irish primary school pupils. *Educ. Rev.* 65, 488–501. doi: 10.1080/00131911.2013.768599
- \*Tolin, D. F., Whiting, S., Maltby, N., Diefenbach, G. J., Lothstein, M. A., Hardcastle, S., et al. (2009). Intensive (daily) behavior therapy for school refusal: a multiple baseline case series. *Cogn. Behav. Pract.* 16, 332–344. doi: 10.1016/j.cbpra.2009.02.003
- Tonge, B. J., and Silverman, W. K. (2019). Reflections on the field of school attendance problems: for the times they are a-changing? *Cogn. Behav. Pract.* 26, 119–126. doi: 10.1016/j.cbpra.2018.12.004
- Torrens Armstrong, A. M., McCormack Brown, K. R., Brindley, R., Coreil, J., and McDermott, R. J. (2011). Frequent fliers, school phobias, and the sick student: school health personnel's perceptions of students who refuse school. *J. Sch. Health* 81, 552–559. doi: 10.1111/j.1746-1561.2011.00626.x
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann. Intern. Med.* 169, 467–473.
- Ura, S. K., Castro-Olivo, S. M., and d'Abreu, A. (2019). Outcome measurement of school-based SEL intervention follow-up studies. *Assess. Effect. Intervent.* doi: 10.1177/1534508419862619
- \*Valles, E., and Oddy, M. (1984). The influence of a return to school on the long-term adjustment of school refusers. *J. Adolesc.* 7, 35–44. doi: 10.1016/0140-1971(84)90046-0
- Van Widenfelt, B. M., Treffers, P. D. A., De Beurs, E., Siebelink, B. M., and Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clin. Child Fam. Psychol. Rev.* 8, 135–147. doi: 10.1007/s10567-005-4752-1
- Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., et al. (2017). What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. *Am. Psychol.* 72, 79–117. doi: 10.1037/a0040360
- Whittemore, R., and Knafl, K. (2005). The integrative review: updated methodology. *J. Adv. Nurs.* 52, 546–553. doi: 10.1111/j.1365-2648.2005.03621.x
- Wolpert, M., Görzig, A., Deighton, J., Fugard, A. J. B., Newman, R., and Ford, T. (2015). Comparison of indices of clinically meaningful change in child and adolescent mental health services: difference scores, reliable change, crossing clinical thresholds and 'added value' – an exploration using parent scores on the SDQ. *Child Adolesc. Ment. Health* 20, 94–101. doi: 10.1111/camh.12080
- \*Wu, X., Liu, F., Cai, H., Huang, L., Li, Y., Mo, Z., et al. (2013). Cognitive behaviour therapy combined Fluoxetine treatment superior to cognitive behaviour therapy alone for school refusal. *Int. J. Pharmacol.* 9, 197–203. doi: 10.3923/ijp.2013.197.203

\* Publications reviewed in the current study.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Heyne, Strömbeck, Alanko, Bergström and Ulriksen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## APPENDIX A | SEARCH TERMS

“school refus\*” (school refusal / school refusing / school refuser) OR “school phobi\*” (school phobia / school phobic) OR “anxiety-based school refus\*” OR “anxiety-based absen\*” (anxiety-based absence / anxiety-based absentee / anxiety-based absenteeism) OR “anx\* school refus\*” (anxious school refus\* / anxious school refusing) OR “school reluctan\*” (school reluctance / school reluctant; as long as the youth are not identified as truanting or showing some other attendance problem) OR “school absen\*” (school absence / school absentee / school absenteeism; as long as the youth are not identified as truanting or showing some other attendance problem) OR “school avoidan\*” (school avoider / school avoidance / school avoidant; as long as the youth are not identified as truanting or showing some other attendance problem) OR “emotional\* absen\*” (emotional absen\* / emotionally absen\*) OR “school refus\* behav\*” (school refus\* behaviour / school refus\* behaviour) OR “separation anx\*” (separation anxiety / separation anxious; as long as this was conceptually linked to absence from school or difficulty going to school) AND [Child\* (child / childhood / children) OR adolescen\* (adolescence / adolescent) OR student\* (student / students) OR youth\* (youth / youths) OR young\* (young / youngster) OR teenage\* (teenage / teenager) OR “lower school” OR “upper school” OR “grammar school” OR “high school”] AND [RCT OR experiment\* OR “quasi experiment\*” OR “longitudinal\*” OR “control group\*” OR “case control” OR “quasi-experiment\*” OR “case stud\*” (case study / case studies) OR “follow-up” OR interven\* (intervene / intervening / intervention) OR program\* (program / programme) OR outcome\* OR eval\* (evaluate / evaluated / evaluation)].

## APPENDIX B | PUBLICATIONS USING DATA FROM A SAMPLE REPORTED IN AN EARLIER PUBLICATION, EXCLUDED AT STEP 2 OF SCREENING

- Elsherbiny, M. M. (2017). Using a preventive social work program for reducing school refusal. *Children Sch.* 39, 81–88.
- Flakierska-Praquin, N., Lindström, M., and Gillberg, C. (1997). School phobia with separation anxiety disorder: a comparative 20- to 29-year follow-up study of 35 school refusers. *Comprehens. Psychiatry* 38, 17–22.
- Layne, A. E., Bernstein, G. A., Egan, E. A., and Kushner, M. G. (2003). Predictors of treatment response in anxious-depressed adolescents with school refusal. *J. Am. Acad. Child Adol. Psychiatry* 42, 319–326.
- Nishida, A., Sugiyama, S., Aoki, S., and Kuroda, S. (2004). Characteristics and outcomes of school refusal in Hiroshima, Japan: proposals for network therapy. *Acta Med. Okayama* 58, 241–249.
- Nuttall, C., and Woods, K. (2013). Effective intervention for school refusal behaviour. *Educ. Psychol. Practice* 29, 347–366.
- Reissner, V., Jost, D., Krahn, U., Knollmann, M., Weschenfelder, A.-K., Neumann, A., et al. (2015). The treatment of school avoidance in children and adolescents with psychiatric illness: a randomized controlled trial. *Deutsc. Ärzteb. Int.* 112, 655–662.
- Sibeoni, J., Orri, M., Podlipski, M. A., Labey, M., Campredon, S., Gerardin, P., et al. (2018). The experience of psychiatric care of adolescents with anxiety-based school refusal and of their parents: a qualitative study. *J. Canad. Acad. Child Adol. Psychiatry* 27, 39–49.
- Walter, D., Hautmann, C., Minkus, J., Petermann, M., Lehmkuhl, G., Goertz-Dorten, A., et al. (2013). Predicting outcome of inpatient CBT for adolescents with anxious-depressed school absenteeism. *Clin. Psychol. Psychother.* 20, 206–215.

## APPENDIX C | PUBLICATIONS REPORTING CASE STUDIES THAT WERE EXCLUDED AT STEP 3 OF SCREENING

- Chabrol, H., Fouraste, R., Morón, P., and Zorzano, M. T. (1995). Father-child hospitalization in the separation anxiety. *Actas Luso Exp. Neurol. Psiquiatr. Cienc. Afines* 23, 223–226.
- Christogiorgos, S., and Giannakopoulos, G. (2014). School refusal and the parent-child relationship: a psychodynamic perspective. *J. Infant Child Adol. Psychother.* 13, 182–192.
- Coulter, S. (1995). School refusal, parental control and wider systems: lessons from the management of two cases. *Irish J. Psychol. Med.* 12, 146–149.
- Ding, J., Gadit, A. M., and Peer, S. B. (2014). School refusal in adolescent young man: could this be an idiopathic amotivational syndrome? *BMJ Case Rep.* 2014:bcr2013203508.
- Hadi, S. R. A., Razak, A. H. A., Kalil, E. Z., and Ismail, W. S. W. (2014). Parental issues contributing to school refusal: a case report. *ASEAN J. Psychiatry* 15, 83–85.
- Oner, O., Yurtbasi, P., Er, A., and Basoglu, N. (2014). The inpatient treatment process for severe school refusal. *Klinik Psikofarm. Büll. Bul. Clin. Psychopharmacol.* 24, 176–179.
- Paccione, D. M. R., and Contessakislus, M. A. (1987). School phobia: Identification of subtypes as a prerequisite to treatment intervention. *Adolescence* 22, 377–384.
- Sherman, J., and Formanek, R. (1985). School phobia in a multiphobic family: the family that phobes together. *Child Adol. Soc. Work J.* 2, 114–124.
- Wetchler, J. L. (1986). Family therapy of school-focused problems: a macrosystemic perspective. *Contemp. Fam. Ther.* 8, 224–240.