Assessment of the implementation fidelity of early interventions: methods of data collection

Colombe Lemire
Université du Québec à Trois-Rivières

Carmen Dionne
Université du Québec à Trois-Rivières

Michel Rousseau
Université du Québec à Trois-Rivières

Abstract

There is a lack of standards regarding methods used to collect and assess data on implementation fidelity of early childhood intervention practices.

Objectives: The purpose of this article is to document and discuss data collection methods that researchers could use for the complex process of evaluating implementation fidelity.

Methods: We conducted a critical analysis of 27 empirical articles published between 1998 and 2019 on implementation of early childhood intervention practices, using the *Matrix Method* (Garrard, 2014).

Results: Eight of the 26 studies used a log or implementation report. In 50 % of the studies (13/26), questionnaires collected information about quality of the intervention delivered, satisfaction and participant responsiveness. Direct or indirect observation (using video) was the most common collection method utilized. Most studies (16/26) used video observation.

Conclusion: Using several methods rather than a single one is advantageous and allows for data triangulation. Different aspects of implementation fidelity must be considered, and information collected through a variety of methods, from different sources and on many occasions.

Keywords: implementation fidelity, early interventions, methods

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

1. Introduction

In the field of early intervention, many authors (Eames, Daley, Hutchings, Hughes, Jones, et al., 2008; Knoche, Sheridan, Edwards, & Osborn, 2010; Luze & Peterson, 2004; Thomas & Marvin, 2016) recognize the importance of assessing implementation fidelity of practices designed to support early child development. An assessment of implementation fidelity looks at whether a program or intervention practices are being used as intended (Carroll, Patterson, Wood, Booth, Rick, & Balain, 2007; Pence, Justice, & Wiggins, 2008; Snyder, Hemmeter, McLean, Sandall, & McLaughlin, 2013). Assessing implementation fidelity is useful and needed for many reasons.

First, without data on the program implementation, accurate interpretation of outcomes would not be possible (Durlak & DuPre, 2008). Fidelity assessment helps determine whether the absence of or weak effects of an intervention can be explained by ineffective intervention practices or an implementation problem (Bond & Drake, 2019; Mowbray, Holter, Teague, & Bybee, 2003). Moreover, fidelity assessment plays an important role in identifying the active ingredients of a program or intervention practices. “Active ingredients” are the most important elements or combination of elements associated with positive outcomes (Dunst, Trivette, & Raab, 2013).

Another reason that validates the importance of assessing fidelity are the documented links between high levels of fidelity and positive intervention effects (Bond & Drake, 2019).

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Conversely, lower levels of fidelity is linked to weaker intervention effects (Chang, Shire, Shih, Gelfand, & Kasari, 2016; Durlak & Dupré, 2008).

When shifting from a research context to implementation in practice settings, data on new interventions used as intended can be collected during the process (Fixsen, Blase, Metz, & Dyke, 2013). Implementation fidelity data are collected with a goal of continuous improvement (Fixsen et al., 2013; Loyd, Supplee & Materra, 2013).

Moreover, in the past few years, various authors have recommended assessing two aspects of fidelity: fidelity of practice implementation (professional development practices which include training, coaching, mentoring) and intervention practice fidelity (Barton & Fettig, 2013; Dunst et al., 2013; Snyder, Hemmeter, Fox, Bishop, & Miller, 2013). From this perspective, fidelity assessment examines if implementation practices are used as intended, with a goal to ensure fidelity of intervention practices will yield intended outcomes (Dunst et al., 2013). For all these reasons, there is a consensus about the importance of assessing implementation fidelity.

However, there are still too few studies that assess implementation fidelity in education and early intervention (Barton & Fettig, 2013; Ledford & Wolery, 2013; O'Donnell, 2008; Ruble, McGrew, & Toland, 2013). Very few studies have focused comprehensively on implementation fidelity (Domitrovich, Gest, Jones, Gill, & DeRousie, 2010). Indeed, studies mostly document quantity of interventions and, more rarely, quality (Caron, Bérubé, & Paquet, 2017; Hamre, Justice, Pianta, Kilday, Sweeney, et al., 2010). In addition, according to Snyder, Hemmeter, Meeker, Kinder, Pasia and McLaughlin (2012) as well as Sheridan, Edwards,

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Marvin and Knoche (2009), there is not enough information about professional development practices to support the application of new intervention practices.

There is also a lack of consensus regarding how to assess implementation fidelity (Proctor, Silmere, Raghavan, Hovmand, Aarons, et al., 2011). In early intervention, as described by Kaiser and Hemmeter, the "standards for treatment fidelity are relatively modest" (2013, p. 81). More work is needed to ensure that assessments of implementation fidelity of early intervention practices are consistently and fully carried out (Domitrovich, Gest, Jones, Gill, & DeRousie, 2010; Kaiser & Hemmeter, 2013). This article documents and discusses data collection methods that researchers could use for the complex but essential process of evaluating implementation fidelity. The research question is as follows: What methods are used in early intervention to assess fidelity of implementation practices and of intervention practices?

2. Method

2.1 Search for publications

To answer this question, a literature review was conducted using the ERIC, PsycINFO and Medline databases. Keywords used were *fidelity*, *implement**, and *early intervention*. The search criteria were the date of publication (1998 to 2019) and language (French and English). The literature review initially identified 319 publications, but the number fell to 251 after duplicates were removed.

2.2 Inclusion and exclusion criteria

Inclusion criteria were as follows: 1) population (preschool-aged children); 2) intervention designed to support young children's development; and 3) context of the intervention (daycare, preschool class). The exclusion criterion was type of publication

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

(literature reviews, systematic reviews, expert opinions, editorials, book chapters, conference or research reports, and theses). In total, 27 articles were retained after inclusion and exclusion criteria were applied. Two articles (Upshur, Wenz-Gross, & Reed, 2013; Wenz-Gross & Upshur, 2012) examined the same study; therefore, those two articles were treated as one study.

2.3 Article synthesis and analysis

The synthesis and critical analysis of those 27 articles are based on the *matrix method* (Garrard, 2014). The matrix method is a simple way to organize, synthesize and analyze the information of a literature review (Klopper, Lubbe & Rugbeer, 2007). The contents of all 27 articles were organized and structured using three matrices. A first matrix drew a profile of each study (context and study design; research questions or objectives; implementation fidelity assessment targets). A second matrix documented data collection methods used to assess fidelity of implementation. After the two matrices were completed, and considering the significance of the observation method's representation, a third matrix was added to focus specifically on the observation method (e.g., frequency, time points, measurement tools). A synthesis of the information gleaned from these matrices is presented below. This article is organized around three sections: 1) Description of studies: fidelity assessment targets; 2) Results: data collection methods used to assess fidelity of implementation practices and of early intervention practices; 3) Discussion: advantages and disadvantages associated with the various methods.

3. Description of studies: fidelity assessment target

In 16 studies included in the review, intervention practices are the target of fidelity assessment. Most of those studies focus on the effects of early interventions and document

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

fidelity from this perspective. For example, Clements and Sarama (2008) looked at the effects of a mathematics skills development program in preschool. Bierman, Nix, Domitrovich, Welsh and Gest (2014) investigated the effects of a curriculum targeting language development, literacy, and socioemotional learning.

Three studies (Burns, Kidd, Nasser, Aier, & Stechuk, 2012; Conroy, Sutherland, Algina, Wilson, Martinez, & Whalon, 2015; Fox, Hemmeter, Snyder, Binder, & Clarke, 2011) assess the effects of professional development on fidelity of intervention practices. The three studies considered only fidelity of intervention practices and not fidelity of professional development practices.

In this review, it is observed that starting in 2012, fidelity of implementation practices is more frequently assessed. Seven studies (Chazin, Barton, Ledford, & Pokorski, 2018; Dunst et al., 2013; Lambert, Gallagher, & Abbott-Shim, 2015; Powell & Diamond, 2013; Ruble, 2013; Snyder, Hemmeter, McLean, Sandall, McLaughlin, & Algina, 2018; Wenz-Gross & Upshur, 2012) assessed fidelity of implementation practices and of intervention practices, as well as developmental improvements in children. In most of those studies, the researchers addressed the statistical relationship of two aspects of fidelity and the effects on children. For example, the objective of Dunst et al. (2013) was to implement with fidelity naturalistic instructional procedures using *PALS: An Evidence-Based Approach to Professional Development* (Dunst & Trivette, 2009). An analysis with a structural equation model was carried out using data for fidelity of professional development practices and naturalistic instructional procedures, and for developmental improvements in children. This review also highlights for the same period (after 2012) the publication of studies on psychometric validation of tools designed to assess

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

implementation fidelity (Boavida, Akers, McWilliam, & Jung, 2015; Snyder, Hemmeter, Fox, et al., 2013; Sutherland, McLeod, Conroy, Abrams, & Smith, 2014). The Results section below describes fidelity assessment methods and tools used in the studies.

4. Results

Table 1 presents the study methods by fidelity assessment targets, that is, implementation or intervention practices, as reported by their respective authors: log or implementation report; self-reported checklist; questionnaire; compilation of contacts or attendance; duration of the intervention; observation (on-site or video); audio recording and interview.

4.1 Log or implementation report

In 8 of 26 studies, logs or implementation reports were used to collect data on fidelity of implementation practices (n = 2) and of intervention practices (n = 6). Data collected were almost exclusively quantitative (7 of 8 studies). One study collected qualitative data using open-ended questions in implementation reports: Bingham, Culatta, and Hall-Kenyon's (2016) study on the implementation of an early literacy intervention in kindergarten. The open-ended questions aimed to document the challenges and successes linked to the intervention practices implemented.

For implementation practices, data were collected using coaches' logbooks (Snyder et al., 2018) and summaries of coaching sessions (Powell & Diamond, 2013).

In 6 studies (Bierman et al., 2014; Bingham et al, 2016; Domitrovich et al., 2010; Guo, Dynia, Logan, Justice, Breit-Smith, & Kaderavek, 2016; Hamre, Justice, Pianta, Kilday, Sweeney, et al., 2010; Shire, Chang, Shih, Bracaglia, Kodjoe, & Kasari, 2017), self-reports

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

(logs, fidelity reports) provide information on the implementation of intervention practices as intended or on quantity or frequency of use.

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)*
https://doi.org/10.1016/j.evalprogplan.2020.101870

Table 1

Data Collection Methods

	Fidelity	Written documents					Observations				
		Log, impl. report	**Checklist**	**Quest.**	**Contacts or att. comp.**	**Duration**	**On site**	**Video**	**Audio**	**Interv.**	**Total**
1-Clements & Sarama (2008)	Int.						**X**				**1**
2-Pence (2008)	Int.			**X**			**X**	**X**			**3**
3-Domitrovich et al. (2010)	Int	**X**					**X**				**2**
4-Hamre et al. (2010)	Int.	**X**			**X**	**X**		**X**			**4**
5-Fox et al. (2011)	Int.			**X**				**X**		**X**	**3**
6-Strain & Bovey (2011)	Int.			**X**			**X**				**2**
7-Burns et al. (2012)	Int.						**X**				**1**

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)*
https://doi.org/10.1016/j.evalprogplan.2020.101870

8a-Upshur et al. (2013); 8b-Wenz-Gross et Upshur (2012)	*Imp.*/int.		**X**	***V***	***V***			**X**			**4**
9-Dunst et al. (2013)	*Imp.*/Int.		***V*/X**		***V***			**X**			**3**
10-Powell et Diamond (2013)	*Imp*/Int.	***V***	***V*/X**			***V***		**X**	**X**	***V***	**6**
11-Ruble et al. (2013)	*Imp.*/Int.		***V*/X**								**1**
12-Snyder, Hemmeter, Fox, et al. (2013)	Valid.								**X**		**1**
13-Zucker, Solari, Landry, et Swank (2013)	Int.								**X**		**1**
14-Bierman et al. (2014)	Int.	**X**		**X**				**X**			**3**
15-Sutherland et al. (2014)	Valid.							**X**	**X**		**2**
16-Boavida et al. (2015)	Valid.							**X**	**X**		**2**

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)*
https://doi.org/10.1016/j.evalprogplan.2020.101870

17-Conroy et al. (2015)	Int.						**X**	**X**				**2**
18-Lambert et al. (2015)	*Imp.*/Int.			***V***	***V***	***V***	**X**		***V***	***V***		**6**
19-Bingham et al. (2016)	Int.	**X**					**X**					**2**
20-Chang, Shire, Shih, Gelfand, et Kasari (2016)	Int.							**X**				**1**
21-Guo et al. (2016)	Int.	**X**		**X**				**X**				**3**
22-Shire et al. (2017)	Int.	**X**		**X**	**X**			**X**				**4**
23-Snyder et al. (2018)	*Imp.*/Int.	***V***	**X**	***V***	***V***		***V***	**X**				**6**
24-Dunlap, Strain, Lee, Joseph, & Leech (2017)	Int.		**X**	**X**								**2**
25-Chazin et al. (2018)	*Imp.*/int.							***V*/X**				**1**

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)*
https://doi.org/10.1016/j.evalprogplan.2020.101870

26-Holzinger, Laister, Vivanti, Barbaresi, et Fellinger (2019)	Int.			**X**	**X**	**X**		**X**			**4**
		2*V*/6X	**3*V*/6X**	**3*V*/8X**	**4*V*/3X**	**2*V*/2X**	**1*V*/14X**	**1*V*/15X**	**2*V***	**1*V*/1X**	

Note. Int. = intervention = X ; Imp. = implementation = V; 1- Log, impl. report = Log and implementation report; 3- Quest. = Questionnaire; 4- Contacts or att. comp. = Contacts or attendance compilation; 9- Interv. = Interview

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

4.2 Checklist or rating scale

Three studies used checklists completed by teachers to assess fidelity of implementation practices. In the studies by Dunst et al. (2013) and Ruble et al. (2013), teachers completed checklists designed to assess if professional training and coaching were implemented as intended. In Powell and Diamond's (2013) study, teachers filled in a rating scale to indicate the level of usefulness of the interventions and the coaches' attitudes.

Six of the 26 studies used checklists to evaluate fidelity of intervention practices without links to a predefined observation method. For instance, in the study by Wenz-Gross and Upshur (2012) on implementing intervention practices targeting socioemotional development, teachers completed daily lesson checklists to identify the activities covered and not covered. A checklist was also used to document teacher satisfaction with the interventions. Snyder et al. (2018) employed the *Learning Targets Rating Scale–Research Version* 2.0 (LTRS; Snyder, Crowe, Hemmeter, Sandall, McLean, & Crow, 2009). This unpublished instrument assesses the quality (e.g., relevance, usefulness, measurability) of embedded learning opportunities selected by teachers. In a study by Ruble et al. (2013), the quality of intervention plans was evaluated by independent raters, who completed the *Individual Education Program (IEP) quality* (Ruble, McGrew, Dalrymple, & Jung, 2010). Those checklists or rating scales used a Likert scale or a dichotomous scale.

4.3 Questionnaire

Three studies collected data on fidelity of implementation practices using questionnaires. First, in the study by Wenz-Gross and Upshur (2012), preschool teachers were trained and

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

supported for two years to implement a curriculum to promote children's socioemotional skills. At the end of each teacher training or coaching session, teachers filled out a satisfaction questionnaire. The study by Lambert et al. (2015) concerns the effects of a mentoring program on the socioemotional intervention practices of Head Start teachers. Mentors and their protégés filled out a questionnaire on the quality of their relationship at the year's end. Lastly, the study by Snyder et al. (2018) assesses the effects of professional development (on-site coaching and self-coaching via website) on use of naturalistic intervention procedures. Teachers completed questionnaires to assess the social validity of training and coaching.

With respect to fidelity of intervention practices, in two studies (Bierman et al., 2014; Pence et al., 2008) teachers completed a questionnaire on the quality of implementation of their practices. In 6 other studies (Dunlap et al., 2017; Fox et al., 2011; Guo et al., 2016; Holzinger et al., 2019; Shire et al., 2017; Strain & Bovey, 2011) a questionnaire was used to examine the degree of satisfaction or teachers' perceived relevance of intervention practices.

4.4 Compilation of contacts or attendance

Data on duration or number of contacts was collected in 5 studies that look at fidelity of professional development. Lambert et al. (2015) documented mentors' contacts (number and duration) with teachers. In the study by Snyder et al. (2018), one group of teachers engaged in self-coaching via a website. Those teachers received weekly email messages prompting them to self-coach. Snyder et al. (2018) counted the number of emails delivered to and read by the teachers.

Concerning fidelity of intervention practices, Shire et al. (2017) assessed the proportion of young autistic children's contacts with the program *Joint Attention, Symbolic Play,*

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Engagement, and Regulation. Holzinger et al. (2019) measured duration of early intervention activities for children with autism spectrum disorders as well as participant retention. In a study by Hamre et al. (2010) on implementation of a curriculum addressing literacy and language development, teachers recorded the duration of their interventions as well as the number of days each week in which they completed the activities.

4.5 Observation (video or on site)

First, only two studies (Chazin et al., 2018; Snyder et al., 2018) used observation to assess fidelity of implementation practices. In Snyder et al.'s (2018) study, an independent observer watched the training and coaching sessions, and scored a fidelity checklist. Chazin et al. (2018) collected data on fidelity of implementation practices using a direct observation coding system to assess coaching sessions.

On-site (15/26) or video (16/26) observation is the most common collection method utilized. In assessing fidelity of intervention or implementation practices, three trends are apparent for time points and frequency of observations: 1) eight studies (Bierman et al., 2014; Chazin et al., 2018; Domitrovich et al., 2010; Guo et al., 2016; Hamre et al., 2010; Powell & Diamfond, 2013; Snyder et al., 2018; Wenz-Gross & Upshur, 2012) carried out repeated observations (for example, once a month or during each intervention session); 2) in almost a quarter of the studies, observations were conducted at three time points during periods varying from two months to a year—before or at onset of intervention implementation, during the intervention and at the maintenance stage. Two other studies (Clements & Sarama, 2008; Hamre et al., 2010) performed observations at three time points during the school year: fall, winter and spring; 3) four studies collected data through observation at two time points: before and during

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

(Powell & Diamond, 2013; Strain & Bovey, 2011) as well as before and after (Burns et al., 2012; Zucker et al., 2013) intervention implementation.

In 11 articles, implementation fidelity measurement tools were available and information on psychometric qualities accessible. Table 2 matches the tools with the studies. Some tools were used in more than one study.

Table 2

Fidelity Measurement Tools

Tools	Studies
Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008)	Bierman et al. (2014) Domitrovich et al. (2010) Hamre et al. (2010)
Early Childhood Classroom Observation Measure (OCCOM ; Stypek & Byler, 2004),	Burns et al. (2012)
Language and Literacy Subscale of the Early Language and Literacy Classroom Observation (ELLCO) *Pre-K Tool* (Smith, Brady, & Anastasopoulos, 2008)	Powell and Diamond (2013)
BEST in CLASS Adherence and Competence Scale (BiCACS)	Conroy et al. (2015) Sutherland et al. (2014)*
Teaching Pyramid Observation Tool-Pilot Version (TPOT-P)	Fox et al. (2011)*
Language-focused curriculum fidelity checklist (LFC)	Pence et al. (2008)*
LEAP's Quality Program Indicators (QPIs)	Strain and Bovey (2011)*
Routines-Based Interview Implementation Checklist (RBI)	Boavida et al. (2015)*

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Note. The studies referenced are the ones associated with the tool.

Other, less accessible tools (unpublished manuscript, thesis, etc.) were included in studies to support observations on intervention practice fidelity—for instance, the *Assessment Profile for Early Childhood Programs: Research Edition II* (Abbott-Shim & Sibley, 1998) in the study by Lambert et al. (2015). Most studies used checklists or Likert scales to assess fidelity of intervention practices through observation.

Other observation techniques were also employed in studies included in this review to show intervention practice fidelity. Conroy et al. (2015) utilized the *Teacher-Child Interaction Direct Observation System – Research Version 2.1* (TCIDOS–RV2.1), a partial-interval behavioral observation coding system used to record and code observed practices. Observation periods last 10 seconds (with an audible signal), with an additional 5 seconds to note down whether or not the behavior occurred. Shire et al. (2017) produced random 10-minute videos at three time points (at the beginning, at the end and at the maintenance stage). Snyder et al. (2018) employed a continuous-event observational coding system—the *Embedded Instruction Observation System* (EIOS; Snyder et al., 2009)—to quantify occurrence of learning opportunities provided to children.

4.6 Audio recording and interview

Audio recordings are used to collect data on fidelity of implementation practices. Powell and Diamond (2013) recorded two out of four coaching sessions; research assistants transcribed the verbal interactions. Transcribed data were used to corroborate information contained in the coaches' implementation reports. In Lambert et al.'s study (2015), audio-recorded standardized interviews were carried out to document coaches' communication skills. The interviews

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

supplemented the *Carkhuff Scale* (Carkhuff, 1969). Fox et al. (2011) conducted structured interviews with teachers to assess fidelity of a comprehensive model for promoting young children's social competence.

5. Discussion

The discussion of this literature review's results focuses on the advantages and disadvantages of the various data collection methods.

5.1 Log or implementation report

Self-reported logs or implementation reports are useful to itemize and compile completed interventions (Fortin, 2016). For instance, in the study by Domitrovich et al. (2010), teachers recorded information in weekly logs about the interventions implemented and the number of intervention units delivered during a given week. This method can be employed to easily collect data on a regular basis. Another advantage of teacher-reported logs is the access they provide to components of the intervention that are difficult to observe (Sutherland, McLeod, Conroy, & Cox, 2013); one example is a component in which frequency of execution is low. In addition, self-reported implementation reports or logs can be useful to bring to light qualitative information, as in the study by Bingham et al. (2016). In that study, teachers documented the number of interventions delivered in addition to answering open-ended questions about challenges linked to practices used. In these ways, valuable information about implementation can be collected from the teachers' perspectives.

However, the time and attention required to write those reports or logs can present challenges (Friesner & Hart, 2005). The social desirability effect is also a disadvantage (Dane & Schneider, 1998; Domitrovich et al., 2010). According to Dane and Schneider (1998), practitioners

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

would not emphasize changes made to a program, or intervention practices that were not used; they simply would not want to appear to be uncooperative. Also, some studies (Hamre et al., 2010; Lillehoj, Griffin, & Spoth, 2004; Resnicow, Davis, Smith, Lazarus-Yaroch, Baranowski, et al., 1998) noted differences or non-significant relationships between self-reported implementation data and observational data. However in the study by Snyder et al. (2018), agreement between data on coaching sessions and observational data was very high.

5.2 Checklist and rating scale

This review notes that checklists without observation methods are used in studies (Dunst et al., 2013; Powell & Diamond, 2013; Ruble et al., 2013) that focus on the complexity of the implementation process by documenting the two aspects of fidelity (implementation and intervention). Checklists filled out by teachers document fidelity of implementation practices (training or coaching).

Those checklists also provide information about intervention practice fidelity. For example, Ruble et al. (2013) used a checklist—the IEP quality measurement tool (Ruble et al., 2010)—to collect data on quality of intervention plans. An independent observer rated the intervention plans using eight of the nine quality indicators established in the *Individuals With Disabilities Education Act* (2004). Most interesting here is the pertinence of the information provided about the field of early intervention. Indeed, quality of intervention plans is linked to developmental improvements in children (Ruble et al., 2013). In a three-year study by Wenz-Gross and Upshur (2012), teachers filled out a fidelity checklist (key tasks and activities) for each intervention, 89 in total. For Wenz-Gross and Upshur (2012), checklists serve as memory aids or

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

self-assessment tools for those implementing the intervention. Checklists are a way to continually improve implementation fidelity; they can be used to track the implementation more easily and at a lesser cost. These advantages are worth mentioning given that the successful implementation of new practices in the field can take many years, as Fixsen, Naoom, Blase, Friedman and Wallace (2005) point out.

However, checklists provide limited information when they are dichotomous and used to list interventions (Sutherland et al., 2013). They do not allow data collection on how an intervention is conducted. For this reason, Sutherland et al. (2013) recommend using checklists with a Likert-type scale.

5.3 Questionnaire

In the current literature review, the questionnaire method is the most frequently used after observation. Questionnaires provide insight into participants' perceptions and opinions (Fortin, 2016). They are also low-cost and simple to administer (Charron, 2004). The fact that they can be completed anonymously is an advantage (Charron, 2004).

Although questionnaires are simple to administer, their development requires time and a systematic approach to ensure a cohesive process. Concepts to measure must be chosen, clear questions formulated and organized, a pretest developed, and validity and fidelity of results verified (Gauthier, Bourgeois, Forget, & Turgeon, 2016). Moreover, as in the case for other self-reported methods, social desirability may also be an issue.

5.4 Compilation of contacts or attendance

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Recording attendance and evaluating the duration of an intervention are used to document the number of treatments received (Lambert et al., 2015; Shire et al., 2017). To have an effect, it is not enough to deliver the expected intervention, as noted by Sutherland et al. (2013); participant engagement and appropriate intervention timing are also necessary.

In terms of limits, these methods do not provide information on what Sutherland et al. (2013) call relational factors (relationships between teachers and children, degree of responsiveness and commitment). This information is important, given that without participant responsiveness, even interventions shown to be effective and implemented with fidelity cannot achieve the desired effects (Sutherland et al., 2013).

5.5 Observation (video or on-site)

Observation is a valued method because it provides objective information (Sutherland et al., 2013). There is less of a social desirability effect on observation data than on self-reported measures (Dane & Schneider, 1998). As for disadvantages, if observers are not blinded to conditions (experimental or control), this can affect objectivity (Conroy et al., 2015).

Observations can be carried out on site or using video. On-site observation allows behavior descriptions within the context of the intervention (Fortin, 2016). According to Boavida et al. (2015), on-site observation delivers better access to body language compared with video recordings. At the same time, the advantage of using video is that it is less intrusive than a personal presence (Boavida et al., 2015). Some studies (Boavida et al., 2015; Pence et al., 2008) combined in-class observation with video. This approach is advantageous since it provides more reliable data.

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

In some studies (Domitrovich et al., 2010; Hamre et al., 2010; Wenz-Gross & Upshur, 2012) in this review, observation time points (video or on-site) were carried out on a regular basis—once a month, for example. This frequency tracks the evolution of implementation while increasing data reliability (Domitrovich et al., 2010). However, some researchers (Domitrovich et al., 2010; Sutherland et al., 2013) point out the intrusive and costly aspects (time and resources) of frequent observations. Then again, using a restricted number of video samples to evaluate implementation limits generalization of results (Hamre et al., 2010). In fact, high frequency of observation time points is a feature of some study designs focused on effects of an intervention. For example, a single-subject study design, such as the one used by Chazin et al. (2018), is characterized by the systematic use of repeated observation measures (Fortin, 2016; Lanovaz, 2013).

Using observational tools with good psychometric properties also helps collect data with reliability. Studies included in this literature review have examined the psychometric qualities of some tools, for example, the *Teaching Pyramid Observation Tool-Pilot* (Snyder, Hemmeter, Fox, et al., 2013) and the *Classroom Assessment Scoring System* (Pianta et al., 2008). However, a number of checklist-like tools to evaluate the implementation in early intervention are developed "in-house" (Caron et al., 2017). A number of studies (Domitrovich et al., 2010; Dunst et al., 2013; Hulleman, Rimm-Kaufmann, & Tashia, 2013; Sutherland et al., 2013) evoke the importance of ascertaining the validity and fidelity of data produced by those checklists.

Another way to collect observational data identified in this review is the partial-interval recording technique. This technique is recommended to evaluate frequency of behaviors (Meany-

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Daboul, Roscoe, Bourret, & Ahearn, 2007), more specifically when behaviors occur at a high frequency (University of Kansas, 2019). It is also useful when the beginning and end of an intervention is difficult to determine (Magerotte & Willaye, 2010). Conroy et al. (2015) note that it is possible to over- or underestimate implementation of practices. It is important to know that partial-interval observation records only intervention occurrence or non-occurrence over a determined period—5, 10 or 15 seconds for example (Magerotte & Willaye, 2010). Partial-interval observation does not record all interventions observed during the time interval, unlike continual observation.

Moreover, many studies use on-site or video observation to improve implementation fidelity. In a professional coaching process, videos are useful for giving feedback (Conroy et al., 2015; Powell & Diamond, 2013). In the study by Holzinger et al. (2019), videos were produced to support supervision of the certification process when using an early intervention program for children with autism. Snyder et al. (2018) assessed two types of professional development: self-coaching via a website and on-site meetings with a coach. Teachers in the self-coaching group filmed themselves for 16 weeks and analyzed their use of the practices. In the other group, coaches observed teachers in class and provided performance feedback. Powell and Diamond's (2013) coaching program employed different methods to collect data on use of intervention practices with a goal of continuous improvement: a checklist, on-site observation and videos recorded by the teachers. These authors point out disadvantages to using videos. In their study, coaching was delivered to teachers at distance. To give feedback, teachers made videos of the practices implemented and then sent them to the coaches. Powell and Diamond report that

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

teachers had taped the videos several times to record good performances. Here, the social desirability effect is a disadvantage.

5.6 Audio recording and interview

Audio recordings and structured or standardized interviews are both underrepresented in the studies included in the review. Audio recordings are used to verify data fidelity of other methods such as questionnaires (Ruble et al., 2013) and self-reporting (Powell & Diamond, 2013). The audio recording method proves to be less intrusive than direct observation. Being able to listen to a recording several times is an advantage. Yet, this method can be time-consuming due to the time required to transcribe the tapes (Powell & Diamond, 2013).

One advantage of the structured or standardized interview method is that it ensures consistency when comparing participant responses (Fortin, 2010). However, these interviews are time-consuming and costly, and social desirability continues to be an issue (Charron, 2004).

6. Implications for practice

In light of the above information, various findings emerge concerning evaluation methods of implementation fidelity of early childhood intervention practices. First, depending on their advantages and disadvantages, some methods, used alone or combined, may be more appropriate to assess one or several dimensions of implementation fidelity.

Second, using several methods rather than a single one is advantageous (Guo et al., 2016). Methods should include a compensatory component. In other words, the strengths of a method (e.g., objectivity of observation) should counterbalance the weaknesses of another (e.g., desirability effect of self-reporting). In the studies reviewed, three different data collection

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

methods, including observation, predominated when assessing fidelity. The data collected also came from various sources: independent observers, coaches, participants. This allowed for triangulation of data when the methods focused on the same dimension of fidelity. Data triangulation is interesting given the reliability issues present in any method (Halle, Metz, & Martinez-Beck, 2013).

Third, on-site or video observation is a preferred method to assess implementation fidelity. And yet, the way observations are carried out should align with the study objectives. If the study objective is to assess the effects of intervention practices, observational data on fidelity could be collected frequently by independent observers blinded to the study conditions. In larger scale studies targeting long-term adoption by practice settings, the way assessment of implementation fidelity is conducted must take into consider that particular context. Our review highlights the relevance of self-reported data collection methods (checklists or logs). To support continuous improvement, those methods function as self-evaluations and checklists. Another interesting method involves teachers self-recording their practices. Videos enable teachers to observe themselves in action and to reflect on their use of the new intervention practices. Whether combined or not, written and video-recorded self-reporting methods are particularly useful to sustain a professional coaching process.

7. Conclusion

Over the years, the objectives pursued and methods of evaluating implementation fidelity have changed. For example, over the past 8 years, using professional development practices as intended has become a target of fidelity assessment. In addition, studies like those included in

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

this review (Bingham et al., 2016; Conroy et al., 2015; Upshur et al., 2013; Wenz-Gross & Upshur, 2012) focus on implementation fidelity to enhance understanding of the effects of intervention practices based on other variables such as characteristics of the children, classes or teachers. Fidelity assessment is also part of a training process to quickly identify challenges and changes to make in program implementation (Chen, 2015; Fixsen et al., 2013; Loyd et al., 2013). In sum, fidelity assessment is very useful both in research and applied contexts (Loyd et al., 2013). But to ensure thorough assessment of implementation fidelity, different aspects must be considered, including study goal; objectives and targets of the fidelity assessment; evolving nature of implementation; and resources and characteristics of the research or intervention setting. Evaluating implementation fidelity of a program or practices is complex, but essential (Durlak, 2015). Such complexity must be considered (Century, Rudnick, & Freeman, 2010; Domitrovich et al., 2010; Guo et al., 2016).

Funding

This research was supported in part by grants from the Chaire de recherche du Canada (CRC) en intervention précoce, Fonds de recherche du Québec – Société et culture (FRQSC, 189836), and Social Sciences and Humanities Research Council (SSHRC, 767-2016-8004).

References

Abbott-Shim, M., & Sibley, A. (1998). *Assessment profile for early childhood programs*. Atlanta, GA: Quality Counts.

Barton, E. E., & Fettig, A. (2013). Parent-implemented interventions for young children with disabilities: A review of fidelity features. *Journal of Early Intervention, 35*(2), 194–219. doi:10.1177/1053815113504625

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Bierman, K. L., Nix, R. L., Domitrovich, C. E., Welsh, J. A., & Gest, S. D. (2014). The Head Start REDI project and school readiness. In A. J. Reynolds, A. J. Rolnick, & J. A. Temple (Eds), *Health and education in early childhood: Predictors, interventions, and policies* (pp. 208–233). New York, NY: Cambridge University Press.

Bingham, G. E., Culatta, B., & Hall-Kenyon, K. M. (2016). Examining the impacts of systematic and engaging early literacy (SEEL): Attention to teacher practices and classroom effects across the kindergarten year. *Journal of Research in Childhood Education, 30*(4), 494–512. doi:10.1080/02568543.2016.1216021

Boavida, T., Akers, K., McWilliam, R. A., & Jung, L. A. (2015). Rasch analysis of the routines-based interview implementation checklist. *Infants & Young Children, 28*(3), 237–247. doi:10.1097/IYC.0000000000000041

Bond, G. R., & Drake, R. E. (2019). Assessing the Fidelity of Evidence-Based Practices: History and Current Status of a Standardized Measurement Methodology. *Administration and Policy in Mental Health and Mental Health Services Research*, 1-11.

Burns, M. S., Kidd, J. K., Nasser, I., Aier, D. J., & Stechuk, R. (2012). An interaction, a conversation, often in the context of play: Constructing Intentional teaching in early childhood education. *NHSA Dialog, 15*(3), 272–285. doi:10.1080/15240754.2012.694496

Carkhuff, R. R. (1969). Carkhuff Scale for empathetic understanding in interpersonal processes. *American Journal of Community Psychology, 6*(6), 583–591.

Caron, V., Bérubé, A., & Paquet, A. (2017). Implementation evaluation of early intensive behavioral intervention programs for children with autism spectrum disorders: A

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

systematic review of studies in the last decade. *Evaluation and Program Planning, 62*, 1–8. doi:10.1016/j.evalprogplan.2017.01.004

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*(40), 1–9. doi:10.1186/1748-5908-2-40

Century, J., Rudnick, M., & Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *American Journal of Evaluation, 31*(2), 199-218.

Chang, Y.-C., Shire, S. Y., Shih, W., Gelfand, C., & Kasari, C. (2016). Preschool deployment of evidence-based social communication intervention: JASPER in the classroom. *Journal of Autism and Developmental Disorders, 46*(6), 2211–2223. doi:10.1007/s10803-016-2752-2

Charron, A. (2004). La description de pratiques d'orthographes approchées d'enseignantes du préscolaire en contexte québécois : une réflexion méthodologique. In E. Falardeau, C. Fisher, C. Simard, & N. Sorin (Eds), *Proceeding of 9th International Association ot the AIRDF* (pp. 1–14). Québec, QC: Presses de l'Université Laval. Retrieved from http://www.colloqueairdf.fse.ulaval.ca/actes/index.php?action=par_auteur

Chazin, K. T., Barton, E. E., Ledford, J. R., & Pokorski, E. A. (2018). Implementation and intervention practices to facilitate communication skills for a child with complex communication needs. *Journal of Early Intervention, 40*(2), 138–157. doi:10.1177/1053815118771397

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Chen, A. H. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective*. Thousand Oaks, CA: SAGE Publications.

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*(2), 443–494. doi:10.3102/0002831207312908

Conroy, M. A., Sutherland, K. S., Algina, J. J., Wilson, R. E., Martinez, J. R., & Whalon, K. J. (2015). Measuring teacher implementation of the "BEST in CLASS" intervention: Program and corollary child outcomes. *Journal of Emotional and Behavioral Disorders, 23*(3), 144–155. doi:10.1177/1063426614532949

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control?. *Clinical Psychology Review, 18*(1), 23–45. doi:10.1016/s0272-7358(97)00043-3

Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S., & DeRousie, R. M. S. (2010). Implementation quality: Lessons learned in the context of the Head Start REDI trial. *Early Childhood Research Quarterly, 25*(3), 284–298. doi:10.1016/j.ecresq.2010.04.001

Dunlap, G., Strain, P., Lee, J. K., Joseph, J., & Leech, N. (2017). A Randomized controlled evaluation of prevent-teach-reinforce for young children. *Topics in Early Childhood Special Education, 37*(4), 195–205. doi:10.1177/0271121417724874

Dunst, C. J., & Trivette, C. M. (2009). Let's be PALS: an evidence-based approach to professional development. *Infants and Young Children, 22*(3), 164-176. doi: 10.1097/IYC.0b013e3181abe169

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Dunst, C. J., Trivette, C. M., & Raab, M. (2013). An implementation science framework for conceptualizing and operationalizing fidelity in early childhood intervention studies. *Journal of Early Intervention, 35*(2), 85–101. doi:10.1177/1053815113502235

Durlak, J. A. (2015). Studying program implementation is not easy but it is essential. *Society for Prevention Research, 16*(8), 1123–1127. doi:10.1007/s11121-015-0606-3

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3-4), 327–350. doi:10.1007/s10464-008-9165-0

Eames, C., Daley, D., Hutchings, J., Hughes, J. C., Jones, K., Martin, P., et al., (2008). The leader observation tool: A process skills treatment fidelity measure for the Incredible Years Parenting Programme. *Child: Care, Health and Development, 34*(3), 391–400. doi:10.1111/j.1365-2214.2008.00828.x

Fixsen, D., Naoom, S., Blase, K., Friedman, R., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: The National Implementation Research Network.

Fixsen, D., Blase, K., Metz, A., & Dyke, M. V. (2013). Statewide implementation of evidence-based programs. *Exceptional Children, 79*(2), 213-230.

Fortin, M.-F. (2010). *Fondements et étapes du processus de recherche. Méthode quantitatives et qualitatives*. (2nd ed.). Montréal, QC: Chenelière Éducation inc.

Fortin, M.-F. (2016). *Fondements et étapes du processus de recherche. Méthode quantitatives et qualitatives*. (3rd ed.). Montréal, QC: Chenelière Éducation inc.

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Fox, L., Hemmeter, M., Snyder, P., Binder, D. P., & Clarke, S. (2011). Coaching early childhood special educators to implement a comprehensive model for promoting young children's social competence. *Topics in Early Childhood Special Education, 31*(3), 178–192. doi:10.1177/0271121411404440

Friesner, T., & Hart, M. (2005). Learning logs: Assessment or research method?. *Electronic Journal of Business Research Methods, 3*(2), 117–222.

Garrard, J. (2014). *Health sciences literature review made easy: The matrix method*. Burlington, MA: Jones & Bartlett Learning.

Gauthier, B., Bourgeois, I., Forget, A., & Turgeon, J. (2016). *Recherche sociale : de la problématique à la collecte des données*. (6th ed.). Québec, QC: Presses de l'Université du Québec.

Guo, Y., Dynia, J. M., Logan, J. A. R., Justice, L. M., Breit-Smith, A., & Kaderavek, J. N. (2016). Fidelity of implementation for an early-literacy intervention: Dimensionality and contribution to children's intervention outcomes. *Early Childhood Research Quarterly, 37*, 165–174. doi:10.1016/j.ecresq.2016.06.001

Halle, T., Metz, A., & Martinez-Beck, I. (2013). *Applying implementation science in early childhood programs and systems*. Baltimore, MD: Paul H. Brookes Publishing Co.

Hamre, B. K., Justice, L. M., Pianta, R. C., Kilday, C., Sweeney, B., Downer, J. T., et al., (2010). Implementation fidelity of my teaching partner literacy and language activities: Association with preschoolers' language and literacy growth. *Early Childhood Research Quarterly, 25*(3), 329–347. doi:10.1016/j.ecresq.2009.07.002

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Holzinger, D., Laister, D., Vivanti, G., Barbaresi, W. J., & Fellinger, J. (2019). Feasibility and outcomes of the Early Start Denver Model implemented with low intensity in a community setting in Austria. *Journal of Developmental & Behavioral Pediatrics: JDBP, 40*(5), 354–363. doi:10.1097/DBP.0000000000000675

Hulleman, C. S., Rimm-Kaufmann, S. E., & Tashia, A. (2013). Innovative methodologies to explore implementation. Whole-part-whole-construct validity, measurement, and analytical issues for intervention fidelity assessment in education research. In T. Halle, A. Metz, & I. Martinez-Beck (Eds), *Applying implementation science in early childhood programs and system* (pp. 65–94). Baltimore, MD: Paul H. Brookes Publishing Co.

Kaiser, A. P., & Hemmeter, M. L. (2013). Treatment fidelity in early childhood special education research: Introduction to the special issue. *Journal of Early Intervention, 35*(2), 79–84. doi:10.1177/1053815113515181

Klopper, R., Lubbe, S., & Rugbeer, H. (2007). The matrix method of literature review. *Alternation, 14*(1), 262-276.

Knoche, L. L., Sheridan, S. M., Edwards, C. P., & Osborn, A. Q. (2010). Implementation of a relationship-based school readiness intervention: A multidimensional approach to fidelity measurement for early childhood. *Early Childhood Research Quarterly, 25*(3), 299–313. doi:10.1016/j.ecresq.2009.05.003

Lambert, R., Gallagher, P. A., & Abbott-Shim, M. (2015). An evaluation of the intensity of mentoring: Child outcomes. *Early Child Development and Care, 185*(8), 1314–1330. doi:10.1080/03004430.2014.992426

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Lanovaz, M. J. (2013). L'Utilisation de devis expérimentaux à cas unique en psychoéducation. *Revue de psychoéducation, 42*(1), 161-183.

Ledford, J. R., & Wolery, M. (2013). Procedural fidelity: An analysis of measurement and reporting practices. *Journal of Early Intervention, 35*(2), 173–193. doi:10.1177/1053815113515908

Lillehoj, C. J., Griffin, K. W., & Spoth, R. (2004). Program provider and observer ratings of school-based preventive intervention implementation: Agreement and relation to youth outcomes. *Health Education & Behavior, 31*(2), 242–257. doi:10.1177/1090198103260514

Loyd, C. M., Supplee, L. H., & Materra, S. K. (2013). An eye to efficient and effective fidelity measurement for both research and practice. In T. Halle, A. Metz et I. Martinez-Beck (dir.), *Applying implementation science in early childhood programs and systems*. (pp. 139-154). Baltimore, MD: Paul H. Brookes Publishing Co.

Luze, G. J., & Peterson, C. A. (2004). Improving outcomes for young children by assessing intervention integrity and monitoring progress: "Am I doing it right and is it working?". *Young Exceptional Children, 7*(2), 20–29. doi:10.1177/109625060400700203

Magerotte, G., & Willaye, E. (2010). *Intervention comportementale clinique. Se former à l'A.B.A*. Louvain-la-Neuve, Belgique: De Boeck Supérieur.

Meany-Daboul, M. G., Roscoe, E. M., Bourret, J. C., & Ahearn, W. H. (2007). A comparison of momentary time sampling and partial-interval recording for evaluating functional relations. *Journal of Applied Behavior Analysis, 40*(3), 501–514. doi:10.1901/jaba.2007.40-501

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. American Journal of Evaluation, 24(3), 315-340.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78*(1), 33–84. doi:10.3102/0034654307313793

Pence, K. L., Justice, L. M., & Wiggins, A. K. (2008). Preschool teachers' fidelity in implementing a comprehensive language-rich curriculum. *Language, Speech, and Hearing Services in Schools, 39*(3), 329–341. doi:10.1044/0161-1461(2008/031)

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*. Baltimore, MD: Paul H. Brookes Publishing Co.

Powell, D. R., & Diamond, K. E. (2013a). Implementation fidelity of a coaching-based professional development program for improving Head Start teachers' literacy and language instruction. *Journal of Early Intervention, 35*(2), 102–128. doi:10.1177/1053815113516678

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al., (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research, 38*(2), 65–76. doi:10.1007/s10488-010-0319-7

Resnicow, K., Davis, M., Smith, M., Lazarus-Yaroch, A., Baranowski, T., Baranowski, J., et al., (1998). How best to measure implementation of school health curricula: A comparison of three measures. *Health Education Research, 13*(2), 239–250. doi:10.1093/her/13.2.239

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Ruble, L. A., McGrew, J., Dalrymple, N., & Jung, L. A. (2010). Examining the quality of IEPs for young children with autism. *Journal of Autism and Developmental Disorders, 40*(12), 1459–1470. doi:10.1007/s10803-010-1003-1

Ruble, L. A., McGrew, J. H., & Toland, M. D. (2013). Mechanisms of change in compass consultation for students with autism. *Journal of Early Intervention, 35*(4), 378–396. doi:10.1177/1053815114546495

Sheridan, S. M., Edwards, C. P., Marvin, C. A., & Knoche, L. L. (2009). Professional development in early childhood programs: Process issues and research needs. Early Education and Development, 20(3), 377-401. http://dx.doi.org/10.1080/10409280802582795

Shire, S. Y., Chang, Y. C., Shih, W., Bracaglia, S., Kodjoe, M., & Kasari, C. (2017). Hybrid implementation model of community-partnered early intervention for toddlers with autism: A randomized trial. *Journal of Child Psychology and Psychiatry*, *58*(5), 612–622. doi:10.1111/jcpp.12672

Smith, M. W., Brady, J. P., & Anastasopoulos, L. (2008). *User's guide to the early language & literacy classroom observation Pre-K Tool (ELLCO Pre-K)*. Baltimore, MD: Paul H. Brookes Publishing Co.

Snyder, P., Crowe, C., Hemmeter, M. L., Sandall, S., McLean, M., & Crow, R. (2009). *Embedded instruction for early learning observation system [Manual and training videos]*. Unpublished instrument, University of Florida, GA.

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

Snyder, P., Hemmeter, M. L., Meeker, K. A., Kinder, K., Pasia, C. and McLaughlin, T. (2012). Characterizing key features of the early childhood professional development literature. *Infants & Young Children, 25*(3), 188-212. doi: 10.1097/IYC.0b013e31825a1ebf

Snyder, P., Hemmeter, M. L., Fox, L., Bishop, C. C., & Miller, M. D. (2013). Developing and gathering psychometric evidence for a fidelity instrument: The teaching pyramid observation tool–pilot version. *Journal of Early Intervention, 35*(2), 150–172. doi:10.1177/1053815113516794

Snyder, P., Hemmeter, M. L., McLean, M. E., Sandall, S. R., & McLaughlin, T. (2013). Embedded instruction to support early learning in response to intervention frameworks. In V. Buysse & E. S. Peisner-Feinberg (Eds), *Handbook of response to intervention in early childhood* (pp. 283–298). Baltimore, MD: Paul H. Brookes Publishing Co.

Snyder, P., Hemmeter, M. L., McLean, M., Sandall, S., McLaughlin, T., & Algina, J. (2018). Effects of professional development on preschool teachers' use of embedded instruction practices. *Exceptional Children, 84*(2), 213–232. doi:10.1177/0014402917735512

Strain, P. S., & Bovey, E. H. (2011). Randomized, controlled trial of the LEAP Model of early intervention for young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 31*(3), 133–154. doi:10.1177/0271121411408740

Stipek, D., & Byler, P. (2004). The early childhood classroom observation measure. Early Childhood Research Quarterly, 19(3), 375-397.

Sutherland, K. S., McLeod, B. D., Conroy, M. A., & Cox, J. R. (2013). Measuring implementation of evidence-based programs targeting young children at risk for

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**

emotional/behavioral disorders: Conceptual issues and recommendations. *Journal of Early Intervention, 35*(2), 129–149. doi:10.1177/1053815113515025

Sutherland, K. S., McLeod, B. D., Conroy, M. A., Abrams, L. M., & Smith, M. M. (2014). Preliminary psychometric properties of the best in CLASS adherence and Competence Scale. *Journal of Emotional and Behavioral Disorders, 22*(4), 249–259. doi:10.1177/1063426613497258

Thomas, A. E., & Marvin, C. A. (2016). Program monitoring practices for teachers of the deaf and hard of hearing in early intervention. *Communication Disorders Quarterly, 37*(3), 184–193. doi:10.1177/1525740115597862

University of Kansas. (2019). *Partial interval recording*. Retrieved from http://www.specialconnections.ku.edu/?q=assessment/data_based_decision_making/teacher_tools/partial_interval_recording.

Upshur, C., Wenz-Gross, M., & Reed, G. (2013). A pilot study of a primary prevention curriculum to address preschool behavior problems. *The Journal of Primary Prevention, 34*(5), 309–327. doi:10.1007/s10935-013-0316-1

Wenz-Gross, M., & Upshur, C. (2012). Implementing a primary prevention social skills intervention in urban preschools: Factors associated with quality and fidelity. *Early Education and Development, 23*(4), 427–450. doi:10.1080/10409289.2011.589043

Zucker, T. A., Solari, E. J., Landry, S. H., & Swank, P. R. (2013). Effects of a brief tiered language intervention for prekindergartners at risk. *Early Education and Development, 24*(3), 366–392. doi:10.1080/1040928

Postprint version. La version publiée se trouve à Lemire, C., Dionne, C., & Rousseau, M. (2020). *Assessment of the implementation fidelity of early interventions: methods of data collection.* Evaluation and Program Planning. *83(101870)* **https://doi.org/10.1016/j.evalprogplan.2020.101870**