

Video Feedback to Enhance Fidelity in Practice-Based Interventions for Young People, a Small-n Study

Author(s)

Boendermaker, Leonieke; Landman, Erwin

DOI

[10.1177/21582440211069387](https://doi.org/10.1177/21582440211069387)

Publication date

2022

Document Version

Final published version

Published in

Sage Open

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Boendermaker, L., & Landman, E. (2022). Video Feedback to Enhance Fidelity in Practice-Based Interventions for Young People, a Small-*n* Study. *Sage Open*, 12(1), 1-16.
<https://doi.org/10.1177/21582440211069387>

**General rights**


It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the library: <https://www.amsterdamuas.com/library/contact/questions>, or send a letter to: University Library (Library of the University of Amsterdam and Amsterdam University of Applied Sciences), Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Video Feedback to Enhance Fidelity in Practice-Based Interventions for Young People, a Small-*n* Study

SAGE Open
January-March 2022: 1–16
© The Author(s) 2022
DOI: 10.1177/21582440211069387
journals.sagepub.com/home/sgo


Leonieke Boendermaker^{1,2}  and Erwin Landman³

Abstract

Competent delivery of interventions in child and youth social care is important, due to the direct effect on client outcomes. This is acknowledged in evidence-based interventions (EBI) when, post-training, continued support is available to ensure competent delivery of the intervention. In addition to EBI, practice-based interventions (PBI) are used in the Netherlands. The current paper discusses to what extent competent delivery of PBI can be influenced by introducing supervision for professionals. This study used a mixed-method design: (1) A small-*n* study consisting of six participants in a non-concurrent multiple baseline design (MBL). Professionals were asked to record conversations with clients during a baseline period (without supervision) and an intervention period (with supervision). Visual inspection, the non-overlap of all pairs (NAP), and the Combinatorial Inference Technique (CIT) scores were calculated. (2) Qualitative interviews with the six participants, two supervisors, and one lead supervisor focused on the feasibility of the supervision. Four of six professionals showed improvement in treatment fidelity or one of its sub-scales. Had all participants shown progress, this could have been interpreted as an indication that targeted support of professionals contributes to increasing treatment integrity. Interviews have shown that supervision increased the professionals' enthusiasm, self-confidence, and awareness of working with the core components of the intervention. The study has shown that supervision can be created for PBI and that this stimulates professionals to work with the core components of the intervention. The heterogeneous findings on intervention fidelity can be the result of supervision being newly introduced.

Keywords

fidelity, interventions, young people, competency building, social work, social sciences, supervision, organizational behavior, management

Introduction

Competent delivery of interventions is related to positive outcomes for children and families (Goense et al., 2016; Lyon et al., 2018; Schoenwald et al., 2009). Evidence-based interventions (EBI) work with continued consultation and support following (initial) training to enable competent delivery of interventions on a long-term basis and to sustain the effectiveness of the intervention (Dorsey et al., 2017; Lyon et al., 2018; Ogden et al., 2005). Supervision (Dorsey et al., 2013; Cunningham et al., 2006; Schoenwald, 2016), also referred to as (expert) coaching (Akin, 2016; Kretlow & Bartholomew, 2010; Lyon et al., 2011), is recognized as a key strategy for building and sustaining competencies among professionals for a proper implementation of interventions (Fixen et al., 2019; Kretlow & Bartholomew, 2010; Wandersman et al., 2012). Attention to qualitative execution of interventions and, thereby, the positive implementation of interventions is key in EBI.

Much less attention is paid to these factors in community practice contexts (i.e., “usual care”—see definition below) (Lyon et al., 2018). Likewise, research attention for the support of professionals working in usual care is limited (Dorsey et al., 2018). This is surprising as both interventions (1) have the same aim—attaining positive outcomes for clients, (2) are known to include many components of evidence-based interventions (Garland et al., 2010; Lyon et al., 2018), and (3) organizations and professionals in children's services are

¹Amsterdam University of Applied Sciences, The Netherlands

²University of Amsterdam, The Netherlands

³NHL/Stenden University of Applied Sciences, Leeuwarden, The Netherlands

Corresponding Author:

Leonieke Boendermaker, Faculty of Applied Social Sciences and Law, Amsterdam University of Applied Sciences, Wibautstraat 5a, Amsterdam 1091 GH, The Netherlands.

Email: l.boendermaker@hva.nl



under high (financial) pressure to deliver effective care (Connell et al., 2009; Friele et al., 2018).

The current paper will discuss to what extent competent delivery of interventions in usual care in The Netherlands can be influenced by introducing supervision to support professionals in this sector.

Usual Care

We begin by clarifying our definition of “usual care” as it exists in the Netherlands where three separate systems are in place: (1) Child and youth *health care* (infant health system, school health system), (2) Child and youth *social care*, based on (a) easily accessible, preventive neighborhood teams and (b) indicated specialized help by local and regional organizations for youth and parenting assistance, as well as organizations for child protection and youth probation, and (3) Child and youth *mental health care*. Since the introduction of the Youth Act in 2015 Dutch municipalities are responsible for providing appropriate services and care for children, young people and families in need of assistance, including mental health provisions (Netherlands Youth Institute, 2019). Due to this development, greater collaboration amongst the three systems is taking place. This research project addresses child and youth *social care* and, more specifically, indicated, specialized help for youth and parenting assistance.

Evidence-Based and Practice-Based Interventions

In Dutch child and youth social care, organizations roughly provide two types of interventions: evidence-based interventions (EBI) and practice-based interventions (PBI). EBI is typically imported from abroad and is applied to the main internalizing and externalizing problem groups. Examples are Parent Management Training Oregon (PMTO) (Forgatch & Patterson, 2010), Triple P (Sanders et al., 2003), Aggression Replacement Training (ART) (Goldstein et al., 1998) or Multi System Therapy (MST) (Henggeler et al., 2009). These EBIs are theoretically based, well documented, protocolled or structured, contain a manual and have gained empirical support in at least two randomized controlled trials (Veerman & van Yperen, 2007). PBIs, on the other hand, are developed within organizations for child and youth social care, usually for client populations with combinations of complex internalizing and externalizing problems. These PBIs make use of a combination of known, effective components of EBI and parts of recognized EBIs which overlap (Leijten et al., 2015; Visscher et al., 2018)

A national program for effective youth care was introduced around the turn of the last century. It was aimed at the bottom-up development of PBI toward EBI, following distinct evidence levels (Veerman & van Yperen, 2007): (1) descriptive and theoretical evidence; (2) indicative evidence; and (3) causal evidence. Each level stems from different types of research (see Brady et al., 2016, for an example of

this developmental process). These are well-known steps in the development of interventions (Bartholomew et al., 2006) which were logically the basis of the currently well-known and worldwide implemented EBI (Schoenwald & Hoagwood, 2001).

Linked to the bottom-up development of interventions, a national accreditation panel was initiated along with a national database of effective interventions, whereby only interventions that are at least accredited on the level of descriptive and theoretical evidence are included. The database supports organizations and local governments in their choices on investments in intervention and care.

Inclusion in the database stimulates organizations to engage in research concerning the interventions they carry out. Admission is valid for 5 years, after which time reassessment is necessary to remain in the databank. Deliverance of research is required to acquire the next evidence level during reassessment. The performance of implementation research combined with research on intervention outcomes is a minimum requirement (Netherlands Youth Institute, n.d.).

Performance of Intervention

During research addressing the implementation of interventions, focus on the quality of the performance of the intervention is crucial. This is indicated by the terms “treatment integrity” or “treatment fidelity,” meaning the degree to which intervention is carried out as intended (McLeod et al., 2013; Perepletchikova, 2011; Schoenwald, Garland, Chapman, et al., 2011). To determine treatment integrity it is necessary to know the components contained in an intervention. Additionally, a distinction is made between common and specific factors of an intervention. Common factors of interventions apply to all target groups regardless of presenting issue. Examples are a good relationship between caregiver and client or, creating positive expectancies (Andrews et al., 1996; McLeod et al., 2013). Specific factors are components of interventions which are effective by specific target groups (i.e., with specific presenting issues). This concerns both therapeutic and conversational therapy techniques as well as intervention format (i.e., group or individual sessions) (Blase & Fixsen, 2013; Van Yperen et al., 2015).

The application of the different intervention components is central in determining treatment integrity or treatment fidelity. This refers not only to the application of the intervention components (therapist adherence), but also competent application (therapist competence) (Perepletchikova, 2011). Therapist competence refers to the technical skills and clinical acumen (timing and appropriateness) in delivering the intervention components (Barber et al., 2006; Barber, Sharpless et al., 2007; Barber, Triffleman et al., 2007). Finally, treatment differentiation in interventions is considered: the degree to which components other than those in the original intervention are applied (often to meet individual client needs) (McLeod et al., 2013; Perepletchikova, 2007).

Here we will refer to the term “fidelity” to indicate professionals’ adherence and competence concerning common as well as specific ingredients of intervention.

Supervision

Supervision (Milne, 2009, p. 15) is “the formal provision, by approved supervisors, of a relationship-based education and training that is work-focused and which manages, develops, and evaluates the work of colleagues.” Supervision has different functions: foremost, it is directed at the progress of the client, reflection on previous sessions, and anticipation of upcoming sessions (Cunningham et al., 2006; Goense et al., 2015). Milne (2009) defines this as the *normative function* of supervision (oversight of quality control and client safety issues). The other two functions are a *restorative function* (fostering emotional support for professionals) and a *formative function* (supervising skill development) (Bearman et al., 2017; Milne, 2009).

Active learning methods, such as role play and other forms of rehearsal are essential for competency building in supervision (Bearman et al., 2013, 2017; Goense, Boendermaker, & Van Yperen, 2016; Nadeem et al., 2013). Additionally, working with video feedback (VF) is considered a powerful (Murray & Leadbetter, 2018; Topor et al., 2017) and effective active learning method, especially in combination with an observation form of the skills to be learned (Fukkink et al., 2011). Despite the fact that reviewing actual practice by using audio or video recordings is considered one of the “gold standard techniques” in supervision (next to symptom monitoring, fidelity assessment, clinical behavioral rehearsal in supervision, and supervisor modeling) (Milne, 2009), only few interventions regularly work with VF as an active learning method in supervision (Dorsey et al., 2018). Only PMTO is known for incorporating VF on a regular basis (Akin, 2016; Ogden et al., 2005; Thijssen & De Ruiter, 2010).

Monitoring Fidelity

In efficacy studies, the measurement of fidelity is vital for linking the studied intervention with the outcomes. In EBI for young people with behavioral problems—and their families—different instruments to assess therapist or trainer fidelity are in use (Goense et al., 2018; Schoenwald, Garland, SouthamGerow, et al., 2011). Instruments for research purposes typically focus on intervention adherence (Goense et al., 2014; Perepletchikova et al., 2007), whereas in supervision, competence assessment will likely be considered more important (McLeod et al., 2016). In EBI for behavioral problems (Goense et al., 2018) and in cognitive behavioral therapy for youth anxiety (McLeod et al., 2016) observational instruments are used to assess fidelity. An example of such an instrument, which is used in regular supervision with VF, is the fidelity of implementation rating system (FIMP) by the Parent Management Training Oregon (Forgatch &

Patterson, 2010). To our knowledge, this is the only example of an instrument that is integrated in an online system and used on a regular basis in video-based supervision (Goense et al., 2018).

Identifying and operationalizing the core components of the intervention is a crucial step in developing such an instrument (Feely et al., 2018; Kirk, 2020). It is important to not only describe the practice elements that must be applied, but also to describe how to implement them (McLeod et al., 2016), allowing an intervention to be “teachable, learnable, doable and assessable in practice” (Fixen et al., 2019, p. 71). Input from professionals working with the intervention should be organized and used to describe the implementation of the core components (Feely et al., 2018; Goense, 2018). The core components of an intervention are those elements that are expected (based on theory) or known (based on empirical research) to contribute to client outcomes. These can be seen as the cause of the intervention’s effect.

The Study

This study describes a small-scale project where supervision with a focus on fidelity was introduced to support professionals of two PBIs in the Netherlands. The central question of this study was: to what extent does supervision—based on a combination of video feedback, role play and a fidelity monitoring instrument as active learning methods—influence the level of fidelity with which professionals conduct the intervention?

The study aims to deliver support in the charting and influence of intervention fidelity of PBI in the Netherlands, as part of the bottom-up development of PBI to EBI.

Method

Design

This 2-year study (2016–2018) used a mixed-method design and consisted of two parts. (1) A small-*n* study consisting of six participants in a non-concurring multiple baseline design (MBL) was used to explore the effect of newly introduced supervision using video feedback, role play, and a fidelity monitoring instrument. Participants filmed a selection of sessions with clients during a 6-month baseline period (without supervision) and an 8-month intervention period (with supervision). These recordings were used in the supervision sessions and were coded by the research team to score fidelity. (2) Qualitative interviews with the six participants, two supervisors, and one lead supervisor on the supervision process were conducted, examining the learning experiences and the feasibility of the supervision. Data were gathered simultaneously, with the qualitative data having a complementary function (Palinkas et al., 2011).

The small-*n* design was chosen in consultation with both participating organizations. As Dutch municipalities have been responsible for providing services and care since the

introduction of the Youth Act in 2015, both financial pressure and workload have increased. No margin exists for a study with the contribution of large groups of professionals.

Context

Two organizations for child and youth social care participated in the supervision project. Both are large organizations that offer a variety of interventions and special education. Both were invited to select a practice-based intervention that (1) was accredited as “theoretically well-grounded” by the Dutch national accreditation committee for youth interventions and formally included in the national database of effective interventions and (2) worked regularly with individual meetings/coaching sessions with young people focused on behavior change that would serve as the point of reference of the supervision.

Site 1 decided to involve the individual counseling sessions of *School2Care* in the supervision project. *School2Care* is an integrated 12-hour day program for young people between the ages of 12 and 17 with externalizing and internalizing problems; it offers special education, care, and support during leisure time. After a 6- to 9-month period the youths resume their regular activities (school, internship, employment) while still receiving outpatient counseling for 6 months. Employees function as coaches: they provide all program components, including individual counseling (Pronk, 2016).

Site 2 decided to involve the individual coaching sessions of *Julia*, a program consisting of 24 weekly meetings ranging from an hour to a half-day which, on average, is completed in half a year. Participants are girls between the ages of 13 and 19 who have encountered sexually inappropriate behavior, and their parents. They experience problems with love, relationships, limits, coercion, and sexuality. In the program six modules are covered in individual coaching sessions, dealing with themes such as self-image, resilience, or the use of social media and internet (Cleef, 2015).

Professionals in both PBIs are expected to use a mixture of techniques in the individual coaching sessions to address (1) non-specific intervention ingredients (Blase & Fixsen, 2013), such as establishing contact, stimulating motivation, and establishing clear goals with clients; and (2) specific therapeutic solution-focused (Gingerich & Peterson, 2013) techniques (such as exception-, scaling, or wonder-questions) and cognitive behavioral (Andrews & Bonta, 2003; Goldstein et al., 1998) techniques (such as cognitive restructuring) in non-protocolled individual sessions with young people.

Intervention

The supervision included video feedback, role play, and application of an instrument to monitor fidelity. Supervision was organized as group supervision and developed by the

Table 1. Content of the Group Supervision With Video Feedback, Role-Play, and Monitoring Instrument.

Start of the supervision session
Welcome
Presentation of the agenda for the supervision (based on the questions of professionals and the video fragments that will be discussed)
Reference to agreements made in the last session
Transition to discussion of video fragment
The professional explains the fragment (what was filmed and why, what core component is to be analyzed)
The question of the professional is clarified
An observation assignment is formulated based on the monitoring instrument
The team watches the fragment
The professional reflects on his/her fragment (which behavior is shown, and what is the connection with the core component discussed)
In-depth questions are asked by supervisor and/or team (for example: what exactly happened, what did you do, what was the effect?)
Participants reflect only on the behavior shown in the fragment.
The team discusses how the professional could further develop positive behavior
The team asks what the professional needs in order to do that
The question of the professional is answered
The feedback is finished with a positive remark on something that went well
Transition to rehearsal (in role play)
It is determined which core component will be practiced in the role play
Non-participants get an observation assignment based on the monitoring instrument
The role play starts with an instruction (description of the situation, number of minutes, rules for interrupting)
Feedback is given based on the 5:1 principle
Closing of the meeting
Summary of the meeting
Evaluation: what was learned, how will this be applied in the coming days/weeks, and could this be recognized in the behavior of colleagues
Agreements for the next meeting

research team based on previous projects (Boendermaker et al., 2012; Goense et al., 2015) and inspired by the organization of supervision in PMTO. An expert of the license holder organization of PMTO in the Netherlands functioned as lead supervisor and trained both supervisors in two 1-day training sessions. Substantiation of the content was outlined in a manual (Goense & Ruitenbergh, 2018).

Supervision sessions followed a fixed structure: specific topics discussed at the start of the session, transition to the discussion of a video fragment, transition to rehearsal in role play, and the closing of the meeting (see Table 1). During the 8 month intervention period, seven 2.5-hour group supervision meetings took place for site 1 and five for site 2. Both

supervisors used an intervention log to report on the supervision content applied. Both used the topics of the start of the session, although it took the supervisor of site 2 two sessions to work with reference to the agreements of the former sessions (see Table 1). Both applied all parts of the discussion of a fragment. Only the supervisor of site 1 worked with rehearsal in role play and both worked with the topics of the closure of the meeting (see Table 1). Supervisors had no disciplinary duties.

As for professionals, supervisor support is necessary to achieve an effective implementation of supervision (Cunningham et al., 2006; Goense et al., 2015; Milne & Reiser, 2016). Likewise, as in supervision content, supervisor training was developed based on previous projects and in close collaboration with the licensed expert, and outlined in a manual (Goense & Ruitenbergh, 2018). The training for supervisors focused on agenda setting, providing feedback, working with action planning, role play, use of the fidelity monitoring instrument, and working with a fixed schedule in the supervision sessions (see Table 1). Following, five support meetings for the supervisors took place between late 2016 and late 2017. One supervisor participated in all meetings, the second supervisor (due to practical circumstances) in three. Supervisor support was concluded with an evaluation meeting (end 2017).

At both sites professionals focus on stimulating the strengths of young people. In the same way, supervision is focused on stimulating the professionals' strengths (Murray & Leadbetter, 2018). Also at both sites, the supervisor applied the same techniques in supervision sessions as professionals were expected to use with their clients. For example, professionals are supposed to use role play with their clients. Likewise, professionals were engaged in role play to experience how this helps in applying new behavior.

Participants

Supervisors. Two (female) supervisors (one from each site) participated in the pilot. One is a supervisor/trainer with a bachelor-level degree and more than 10 years of experience in supervision and the other is a master-level psychologist with about 2 years of experience in supervision. Mean age was 35.

Professionals. Six professionals (three male, three female) participated in the pilot (site 1: $n=4$; site 2: $n=2$). All completed a bachelor-level social work education. Their ages varied between 27 and 38 years, and excepting one participant they all had more than 10 years' work experience in youth care.

Selection. Both organizations had two teams working with the selected intervention. Teams were approached by a researcher of their own organization who, together with the authors and a third researcher, formed the research team in this project. All aspects of the pilot were discussed during

a team meeting (including the recording of sessions with clients, the use of the recordings in group supervision sessions, the coding of the recordings by the research team, the need for written informed consent of all clients to record sessions, as well as consent of the parents with clients under 16). Teams were free to decline on participation in the pilot. In both organizations one team volunteered to participate in the supervision and provided informed consent for participation in interviews. The other teams (one in each organization) were facing a period of instability due to staff turnover and sick leave and decided not to participate. Both teams have experience with non-protocolled support sessions, concerning personal dilemmas in cases or with implementing the intervention.

Video Recordings

The six professionals participated in a short training on recording client conversations using an instruction film (Meerhoek et al., 2015), and received instructions on informed consent to record conversations with young people.

Participating professionals were asked to record four conversations with clients during the baseline period and four during the intervention period (6 professionals \times 4 recordings \times 2 periods) in different phases of their work with clients. In multiple baseline designs, the recommended measurement standard is to include at least five data points in each baseline and intervention phase in at least three participants (Kratowill et al., 2013). In this study the number of participants (six) is more than sufficient. Since acquiring consent and making the videos in daily practice is demanding, four recordings during each phase were planned. When multiple baseline designs include at least three data points in each phase this is considered to meet standards with reservations.

Fidelity Rating

Development fidelity instrument. At both sites the fidelity instrument was developed based on previous work (Goense & Ruitenbergh, 2015). In line with the steps for identifying core components of intervention as described in Feely et al. (2018) and Kirk (2020) the core components were based on the intervention descriptions available in the Dutch national database of effective interventions (Cleef, 2015; Pronk, 2016). According to the steps of developing a practice profile (Metz, 2016) work sessions with professionals and their supervisor were used at each site to define core components and competent delivery (Boendermaker & Goense, 2017). This process is described in a manual (Goense, 2018).

The core components defined were: *contact*, *application of techniques*, and *goal orientation*. Contact consisted of seven (site 1) and six (site 2) elements respectively, such as: "being a good listener," "being patient," "being reliable," "connecting with the youth," and "being a positive role

model.” At both sites five elements were distinguished for application of techniques: “giving feedback,” “providing insight into emotions, thoughts, and behavior,” “utilizing motivational conversation techniques,” “stimulating and practicing alternative behavior,” and “working in a solution-oriented manner.” In both interventions one element was distinguished for goal orientation. Because of the small difference in the description of contact in both interventions, the total number of elements was 13 at site 1 and 12 at site 2. This small difference is acceptable from the viewpoint of implementation; in this way the fidelity instrument is applicable and fitting for the specific situation in practice (Damschroder et al., 2009; Fixen et al., 2019). Also, due to the fact that by consideration of fidelity we assume the total score of the three core components.

Each element was further defined as to what constituted a competent execution. “Being a good listener,” for example, means that the professional “takes their time and gives the conversation their full attention,” “is an active listener (makes eye contact, uses friendly intonation and facial expressions, nods, and/or agrees),” “speaks in a calm and clear manner, knows how to incorporate silence,” “summarizes what has been discussed,” and “is flexible when encountering resistance (sees youth’s point of view, repeats arguments).” In this manner, the three core components were operationalized in professionals’ behavior on the 13 (site 1) and 12 (site 2) elements (Appendix).

The operationalized elements were combined into an observational questionnaire, to be employed to monitor the execution of the intervention. In the project, this instrument was used in two ways: as a supervision instrument to *reflect* on the execution and as an instrument for the researchers to *measure* the quality of the execution (Goense et al., 2015).

Fidelity Rating

The observational questionnaire was completed for each video recording by a researcher of the research team, all of whom were familiar with the interventions and helped develop the instruments. Each recording was viewed in its entirety, with researchers checking for the three core components (operationalized in 12/13 elements) regarding they were applied, and if so, how competently. Each item was assigned a rating between 1 and 9, where a score of 7 to 9 means “good work,” a score of 4 to 6 means “acceptable,” and a score of 1 to 3 suggests that more work is needed on the specific component. Scores were combined to determine the total score.

Inter-rater Reliability

At the start of both coding processes agreements were made on the interpretation and rating of the recordings. Subsequently, two researchers coded the same recording. The inter-rater reliability of this was satisfactory (site 1: ICC

.62) to good (site 2: ICC .80). Subsequently, two randomly selected recordings were coded by two researchers as well. These tests also yielded satisfactory and good inter-rater reliability scores respectively (site 1: resp. ICC .64 and .88; site 2: ICC .80 twice).

Qualitative Interviews

To gather qualitative information on the supervision process, professionals (6) as well as supervisors (2) and the lead supervisor (1) were approached for an interview near the end of the intervention period. Interviews were between 30 and 45 minutes. All respondents received a transcript of their interview for validation. Interviews focused on the content of the supervisory sessions and (learning) experiences during the supervision (as participant, supervisor, or lead).

Analyses

Fidelity measures. In this project each of the six professionals was treated as a single case. Single-subject experimental design data analysis started with a visual inspection of the alliance, application of techniques, goal orientation, and total fidelity score (Barlow et al., 2009). After visual inspection the non-overlap of all pairs (NAP) and the Combinatorial Inference Technique (CIT) scores were calculated. Both scores compare the scores between the baseline and the intervention period. A so-called NAP score is used in single-subject studies to calculate the differences in scores between the baseline and the intervention period. The NAP score expresses the likelihood that a random score from the intervention period is higher than a random score from the baseline period (in this case, the period with and without coaching). The NAP score is a value between 0 and 1 in which .50 means no difference between both time periods. As a rule of thumb, in this pilot we interpreted .65 as “meaningful” (Parker & Vannest, 2009).

The so-called CIT score was used to establish the level of change relative to the maximum change achievable based on the baseline score. This method was derived from the Combinatorial Inference Technique, which compares all possible changes with the maximum change (Rouanet et al., 2000)

Interview data. Data were analyzed using Maxqda 11 by a master level, junior researcher in collaboration with the first author; working from a directed content analyses approach the focus was on identifying themes or patterns in the data (Hsieh & Shannon, 2005). Basic codes were (learning) experiences with the distinct parts of the supervision protocol (VF, role play and the fidelity instrument). Due to the limited number of interviews analysis consisted of one round of coding.

Results are per site, combined with the results of the quantitative analysis, discussed with participants for validation.

Table 2. Mean Scores on the Monitoring Instrument of Each Professional (Range 1–9).

	Baseline (no supervision)	Number of recordings	Intervention (supervision)	Number of recordings
Professional 1				
Contact	6.7	2	6.7	7
Use of techniques	6.3	1	4.5	7
Goal orientation	5.6	2	6.0	7
Total	6.2	2	5.7	7
Professional 2				
Contact	6.1	2	7.0	3
Use of techniques	1.9	3	5.3	3
Goal orientation	3.8	3	6.3	3
Total	3.9	3	6.3	3
Professional 3				
Contact	5.9	1	7.2	7
Use of techniques	3.2	1	2.8	6
Goal orientation	3.5	1	6.8	7
Total	4.2	1	5.6	7
Professional 4				
Contact	8.3	3	6.9	4
Use of techniques	5.3	3	5.7	2
Goal orientation	6.3	3	6.8	4
Total	6.6	3	6.5	4
Professional 5				
Contact	7.7	6	8.3	1
Use of techniques	5.8	6	6.8	1
Goal orientation	6.0	6	8.0	1
Total	6.8	6	7.7	1
Professional 6				
Contact	7.2	5	8.0	3
Use of techniques	5.4	4	7.4	3
Goal orientation	4.4	5	5.0	3
Total	6.3	5	7.5	3

Results

Fidelity Ratings

Mean fidelity scores. Producing video recordings proved challenging, especially during the baseline period. Professionals were not always successful in attaining permission to record sessions, therefore the number of recordings in the baseline and the intervention period varied. In total, 20 baseline recordings were made and 25 recordings during the intervention phase (a total of 45 available recordings out of the requested 48). The desired distribution of recordings over the different phases of treatment was, with some exceptions, maintained.

Sessions with clients were sometimes difficult, especially those with young people entering a program, where the focus is on engagement and motivation. Therefore some recordings were very short; recordings varied in length between 5 and 96 minutes (three recordings were 5 minutes in length, the duration of most were around 25–30 minutes). All recordings were accepted as the goal was competency building and learning and with the consideration that a 5-minute session

with an unmotivated young person may offer significant learning possibilities.

As the focus of the sessions differed, so did the number of items in the core element *contact*. It was decided to present the outcomes as averages of the three core components of both interventions: *contact*, *application of techniques*, and *goal orientation* (Table 2). The column “number of recordings” shows that not all three components were applied in all conversations. According to the professionals, these did not always fit the purpose of the individual session.

Fidelity change. Visual inspection in single case research consists of evaluating the graphs with respect to six elements: level, variability, trend, overlap, intercept gap, and consistency (Vannest & Ninci, 2015). Assessing the level change means evaluating whether treatment fidelity increased from baseline to intervention. Variability is evaluated by judging whether fidelity was stable between measurement moments within phases. Trend occurs when increasing or decreasing patterns are visible within phases. For example, increasing fidelity within the baseline phase would indicate trend and

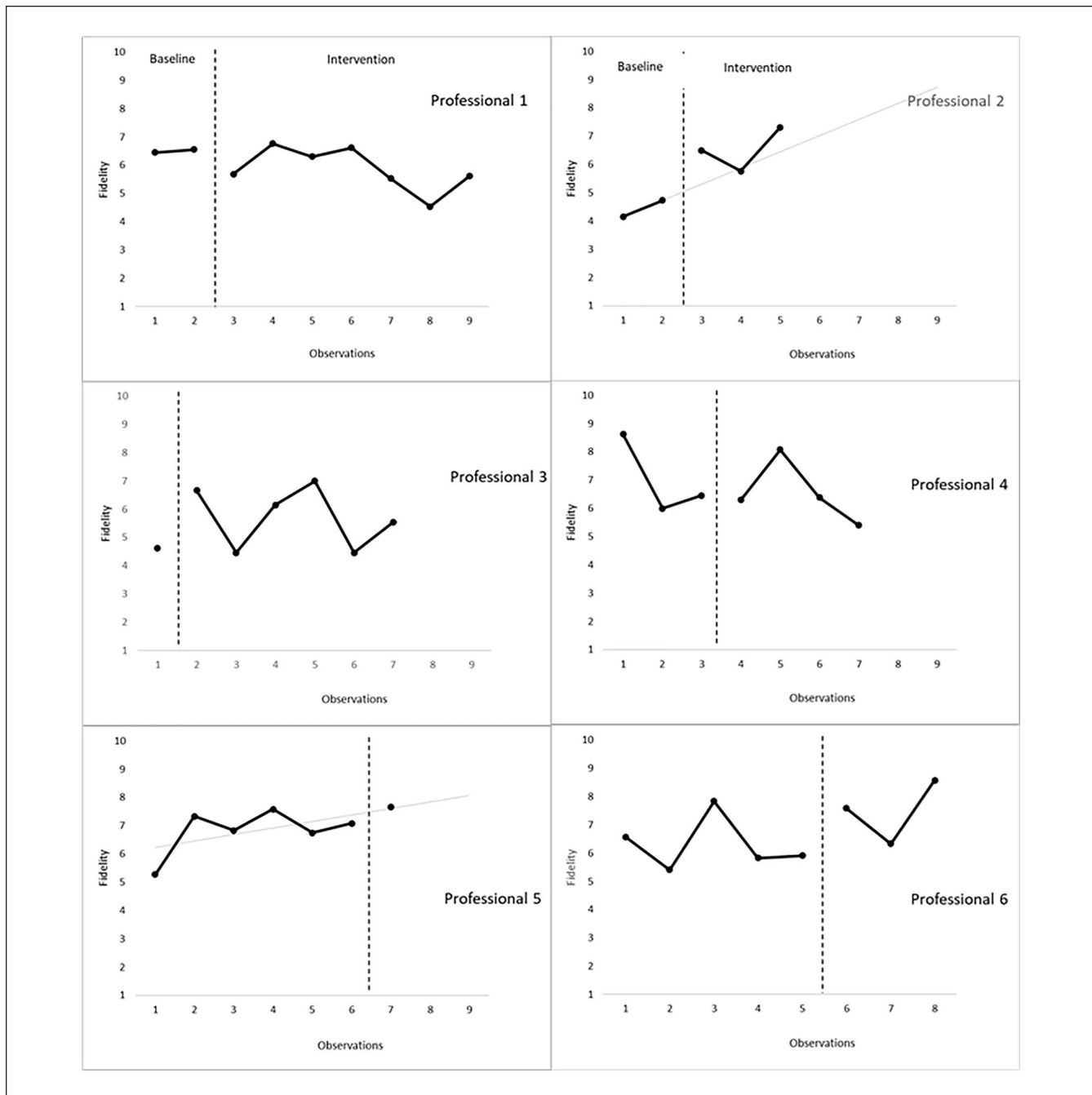


Figure 1. Development of the treatment fidelity during baseline and intervention phases in six professionals.

thus complicate assigning change to the intervention. Overlap is assessed by evaluating the difference in range of fidelity in both phases. It indicates to what degree data points in the baseline phase fall in the range of the data seen in the intervention. Intercept gap evaluates whether change occurs immediately when intervention is started, or is delayed and cumulative. Lastly, consistency is evaluated by examining whether patterns across cases are the same. Desirable change would mean consistent increasing level with a clear intercept

gap and no overlap between phases, without signs of high variability or problematic trend in one of the phases.

The visual inspection of the development of treatment fidelity over time for the six professionals (Figure 1) consistently shows no desirable change in each of the six cases (Table 3). Professional 2 and 5 seems to have higher levels of fidelity in the intervention phase, but trend in the baseline is problematic. In the other professionals there is no evident level-change between the two phases. Neither immediate or

Table 3. Visual Inspection of the Fidelity Score by Assessing the Development of Six Professionals With Respect to Six Elements (Vannest & Ninci, 2015).

	Level	Variability ^a	Trend	Overlap	Intercept ^b	Consistency ^c
Prof 1	Decreasing	Stable	Not problematic	Much	No gap	No evidence
Prof 2	Increasing	Stable	Problematic	None	Yes	No evidence
Prof 3	Increasing	Stable	Not problematic	Partly	Yes	No evidence
Prof 4	Decreasing	Stable	Not problematic	Much	No gap	No evidence
Prof 5	Increasing	Stable	Problematic	None	No gap	No evidence
Prof 6	Increasing	Stable	Not problematic	Much	No gap	No evidence

^aNo more than 3 points difference between two subsequent data points within the phases is coded as stable.

^bMore than 1 point difference between last data point in baseline and start of intervention phase denotes an intercept gap.

^cDevelopment within the professional that would support the effectiveness of the intervention are marked.

cumulative change occurs between phases and much overlap between data from baseline and intervention phase exists.

The subscales follow a similar pattern (Figure 2). Participants 1, 2, 3, 5, and 6 show improvement over time in some or all of the subscales. Due to positive trend during baseline, lack of consistency between the professionals and phases with few data points, the patterns render no evidence of an effect in subscales of treatment fidelity. With respect to contact, professionals 2, 3, 5, and 6 show improvement, however, there is only one data point in one of the phases in two of the cases (p3 and p5). The implementation of techniques grows only in professional 2, but there a positive trend during baseline already existed. Lastly, with respect to goal orientation, professional 1, 2, 3, and 5 show improvement. However, lack of data in one of the phases (p3 and p5) and positive trend in the baseline (p1 and p2) are present.

Evaluation of changes in fidelity. As the NAP scores (Table 4) show, professional 2 shows meaningful change in application of techniques, goal orientation, and total score. Professional 3 shows meaningful change in the elements contact and goal orientation. Professional 5 makes meaningful progress in all elements and professional 6 makes meaningful progress on those items that fall under contact and application of techniques and on total score. Professionals 1 and 4 show no meaningful change. Based on CIT changes that were found were moderate to strong: The professionals that show meaningful change make between 44% and 96% progress on the different parts of the fidelity measure (Table 4). These findings do not determine that a functional relation exists between video feedback and treatment fidelity since not all replications show the same pattern. Some desirable changes occur, but whether these changes are due to the video feedback is doubtful at least.

Interview Supervision

Participants responded enthusiastically and saw advantages in working with the organized supervision. It ensures that supervision is goal-oriented and helps to formulate [and

practice?] concrete (learning) questions during discussion of the video fragments. The fixed closure of a discussion of the fragment creates a focus for the professional for the next session with the client. “You see that what has been previously said continues to play a role in the following sessions and that video recordings are made of this” (supervisor).

Recording and viewing of videos. The professionals had to accustom themselves to producing videos. Requesting permission (also from parents) was discussed and practiced in the first supervisory session. Watching the recordings was very positive for all participants. They felt that the recordings clearly showed how the professional performed and how the youth responded. Performance was more evident than by the oral description of a situation during peer supervision. Through observation by others and through the questions and suggestions by colleagues and the supervisor, professionals learn alternative working methods. They become more aware of their techniques and the intervention in which they work. The strength-based approach used in the coaching sessions was considered a welcome change from standard case discussions, thereby creating more enthusiasm and energy and increasing the professionals’ confidence. This allowed for the manifestation of a safe atmosphere and a learning culture. As one professional commented, “This is educational for the entire organization.”

Fidelity instrument. The fidelity instrument helps to view the recordings in a focused way and to provide concrete and specific feedback. It contributes to “speaking the same language” regarding the invention being used.

Professionals stated that using the fidelity instrument led to an increased awareness of the core components of the intervention which changed their approach to the work. One professional said: “Normally I just try to have a good conversation, but now I give much more thought to what I want to achieve with the conversation.” Another professional: “I would start a conversation, certain topics would present themselves along the way, we would discuss these and that would be it. Now I have been able to structure that. I am more aware of it, and I am working on incorporating this.”

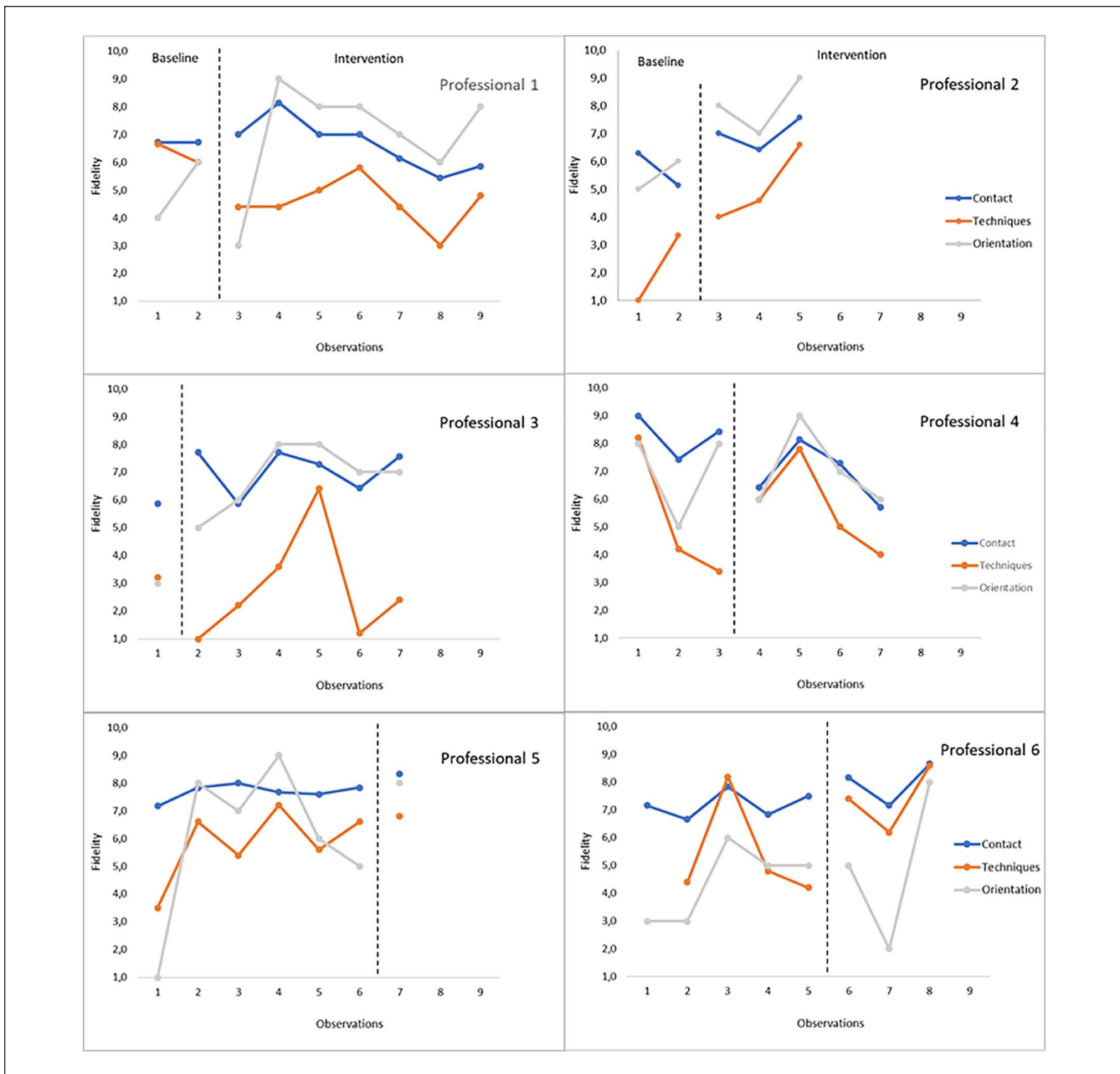


Figure 2. Development of the three core elements contact, application of techniques and goal orientation during baseline and intervention phases in six professionals.

Several professionals shared their realization that when a coaching session did not go as envisioned, it was not always because of the client: “(. . .) Professionals tend to adopt a certain style and then claim afterwards that they can’t work with a certain youth, because they feel no connection. You cannot really say that (. . .) You can look at yourself: What exactly am I doing (. . .) You can better yourself and adjust your approach.”

Role play. After discussing the contributed fragment rehearsal (in role play) was also seen as a positive learning

experience, because tips could be immediately applied. The professional who had contributed the fragment chose which tips to practice. One example: “For example, we practised using solution-focused questions . . . you really have to formulate your questions differently. If you’ve done it once, then you’ll try it again the next time” (professional).

Supervisors. Providing feedback required a great deal of creativity from both supervisors. Incorporating different methods for providing feedback (flipchart, oral feedback, post-its) helped to keep it fun and varied. Both supervisors

Table 4. NAP and CIT Scores of the Six Pilot Professionals (≥ 0.65 bold).

	Alliance	Use of techniques	Goal orientation	Total
NAP scores				
Professional 1	.446	.271	.580	.424
Professional 2	.584	.771	.813	.662
Professional 3	.659	.377	.929	.598
Professional 4	.260	.567	.573	.419
Professional 5	.690	.683	.750	.693
Professional 6	.710	.740	.533	.689
CIT scores				
Professional 1	—	—	—	—
Professional 2	—	.44	.70	.60
Professional 3	.75	—	.55	—
Professional 4	—	—	—	—
Professional 5	.81	.60	.96	.82
Professional 6	.76	.81	—	.78

learned a great deal from the supervisory sessions where they discussed their own recordings of the supervision sessions with the professionals.

They recognized the value of working with a strength-based approach in their own supervision and that of the professionals as a parallel process of working with the clients of both interventions. The suggestion was made in the interviews to create reminders to help professionals incorporate the tips and core components in daily practice. Busy work schedules cause these elements to quickly fade into the background as professionals cannot find the time to apply the agreed-upon core elements.

Preconditions. For this type of supervision to stand a chance of success, several organizational preconditions must be met. Quality equipment must be readily available to record sessions. Professionals should be able to view recordings anywhere (making a secure online environment for uploading and downloading a prerequisite). As acquiring clients' permission to record conversations proved challenging, a well-formulated text or pamphlet appropriate for the target group could be helpful. Finally, adequate time must be made available for the lengthy nature of supervision, addressing recording sessions, viewing of the recordings, and preparing the supervision. It is expected that recording enough videos will not be an issue. For the purpose of the project, eight recordings were needed per professional, which was considered a burden.

Discussion

The goal of this project was to develop supervision for practice-based interventions and to test the effect on intervention fidelity. Applied to a small-scale project of professionals working with young people "in unprotocolled" coaching sessions in two interventions, group supervision occurred using

video feedback, role play, a fidelity monitoring instrument and supervision of the supervisors.

Although it took some time for the professionals to become comfortable in asking youths' (and sometimes their parents') permission to record conversations, as well as making the actual recordings, ultimately working with video feedback in group supervision has increased the professionals' enthusiasm, self-confidence, and awareness of working with the core components of the intervention.

Unfortunately, this positive picture from the interviews is not reflected in the quantitative data. Four out of six professionals showed improvement in treatment fidelity or one of its subscales. Had all participants shown progress, this could have been interpreted as an indication that targeted support of professionals contributes to increasing treatment integrity. However, not all project participants showed meaningful progress and progress did not take place in all core components. Moreover, other aspects of the data such as trend, overlap, and short phases, complicate assigning the improvement to the supervision. Two professionals showed a decline in quality in one component. As this was not a meaningful change it can be concluded that these professionals' fidelity scores remained more or less the same; these professionals showed no development. All in all, some desirable changes occurred, but whether these changes were due to the video feedback is doubtful at least.

Influence of Age and Experience

There are several possible explanations for the heterogeneous results. First, other research shows that age and experience might play a part. For example, interviews with supervisors have shown that professionals with less experience benefit more from support (Goense et al., 2015), whereas experimental research by Bearman et al. (2013) shows that practice and modeling (in the pilot: modeling using video recordings)

mostly impact older employees. However, differences in age and experience were not that substantial in this pilot, making attribution difficult. Moreover, one could argue that all professionals started the pilot with relatively high fidelity, making progress in the short time of the pilot difficult and perhaps not equally necessary in all cases.

Supervision Frequency

Monthly supervision occurred during this study. Over a limited period of time (estimate around 20 hours per person, based on 12.5–17.5 hours of group supervision over 8 months, plus time for producing recordings and preparing supervision) a great deal was accomplished.

In other EBIs there is a higher frequency of supervision. In PMTO, for example, twice monthly supervision is employed (Akin, 2016). In Multisystem Therapy (Henggeler et al., 1993) weekly supervision is the norm. Research in the application of supervision in evidence-based intervention shows that in practice supervision occurs less frequently than prescribed in the intervention. A high workload, which was also seen in our project, plays a role in this (Dorsey et al., 2017). It is possible that a higher frequency of supervision in our project would have led to other results.

Implementation of Supervision

This project has shown that supervision can be created for practice-based intervention. Interviews show that supervision based on a strength-oriented approach and learning culture stimulates professionals, resulting in active learning and skill building. These are important core functions of supervision in the implementation of interventions (Akin, 2016). As stated, the participants were previously unfamiliar with group supervision with a focus on fidelity. This form of support is to be seen as an implementation strategy, geared toward the competent delivery of practice-based interventions. It should not be underestimated that the adaptation of a new work method such as this is in itself an implementation process. It is for good reason that several organizational aspects, which are known barriers in implementation processes, are mentioned in the interviews (Damschroder et al., 2009; Fixen et al., 2019).

In a next phase of supervision at both sites it is important to: (1) pay greater attention during group supervision to those core elements not yet fully applied, such as the application of techniques; (2) to pay attention to the use of rehearsal in role play during supervision; and (3) ascertain what consequences the development in treatment integrity has for client outcomes.

Limitations and Future Directions

The intervention period of this project was 8 months. During this time, each participant could only ask a learning question

twice, and these (basic) learning questions mostly focused on the core component “contact.” Outcomes show that in a next phase supervisors could focus more on the application of specific core components (such as “techniques” or “goal orientation”) which now show (too) little progress or show a decline in quality.

Secondly, recording sessions with clients was experienced as challenging, especially during the baseline period. It took professionals some time to acclimate to asking client permission. In one organization, supervision did not start at the agreed upon moment, even though one of the professionals had already made arrangements with her clients to record sessions. Because of these practicalities, the number of recordings made during the baseline and during the intervention period differed. One professional only made one recording during the intervention period. This may have influenced the outcomes. In addition, recording situations were not randomly assigned, as getting permission and recording were novel aspects and professionals and clients had to get used to these. In a next phase of the supervision and research, random assignment of sessions to be recorded would be a possible solution to getting too many recordings during baseline or intervention phase.

Thirdly, the baseline as well as the intervention period were relatively short in the non-concurrent multiple baseline design. The design would have benefited from a longer baseline and intervention period, during which a larger number of measures could have been taken and the design could have been brought into greater alignment with the standards of multiple-baseline designs. In practice, it was challenging to acquire four measurements in each phase, resulting in phases consisting of less than three observations in four cases. Also, in a non-concurrent multiple baseline design, ideally the start of the baseline is simultaneous. When conditions change in one case only the data pattern in that case should change. If the data patterns in other cases change as well at the same time, there might be an external cause. This could not be controlled in the project as observations were executed in different timeframes. If this project is replicated, obtaining phase lengths of at least five data points within at least three cases that receive supervision in the same timeframe is recommended. Additionally, the fact that participation in the pilot was voluntary could influence outcome. It is possible that these professionals were already eager to learn. However, as two professionals showed no development this might not be the case. Also the fact that the intervention at site 2 only works with girls, compared to boys and girls in site 1, might be influencing the results.

Finally, evaluation of fidelity was done using an instrument developed for use in the supervision sessions. This instrument has not yet been validated and may pay insufficient attention to certain components. The appraisal in general was quite “crude.” Every recording was viewed in its entirety, regardless of length. It is probable that a 5-minute recording will be scored differently than a recording lasting 55 minutes.

Appendix

Table AI. Operationalization of Three Core Components at Two Sites in Elements and Example Behavior.

Core component	Element	Example behavior
Contact	1. The professional is a good listener	Takes adequate time for the session and gives full attention Has an active listening attitude (friendly intonation/facial expression, nodding) Speaks calmly and clearly, is comfortable with silences Summarizes what was discussed
	2. The professional is involved and empathetic	Asks about the young person's different areas of life Inquires about experiences, thoughts, and feelings Empathizes with the young person's perspective Responds to and empathizes with what the young person is saying Gives compliments and magnifies success Names emotions and behaviors
	3. The professional is predictable and consistent	Is clear about expectations within the conversation Is consistent in statements Is transparent about what information is shared with whom Is open regarding consequences and uses them consistently
	4. The professional connects with the young person	Adjusts attitude, tone, and language to the young person Begins at a low-threshold/is light-hearted and uses humor Considers the (im)possibilities of the young person Checks if the young person and professional understand each other
	5. The professional has an open and accepting attitude	Doesn't judge or accuse Doesn't show disappointment in the young person Relates, uses humor
	6. The professional acts as a positive role model	Is optimistic and shows confidence Demonstrates pro-social values (arrives on time, looks neat, does not smoke in view of the student)
	7. The professional is direct (extra element site 1)	Holds the young person accountable for his/her own responsibility where necessary Is realistic about the young person's capabilities Repeatedly inquires about wishes and ideas
Application of techniques	1. The professional uses feedback	Lists the pros and cons of behavior and its consequences Gives feedback on the behavior, not on the person Gives sincere compliments
	2. The professional provides insight into feelings, thoughts, and behavior	Investigates patterns in thoughts, feelings, and behavior with the 5G method Names helpful and disruptive thoughts Works with the top-10 cognitive distortions Explores with the young person the consequences of his/her behavior and alternatives
	3. The professional stimulates the young person to show alternative behavior	Practice alternative behavior with the student (role play) Make concrete agreements about how the student will self-practice Provides feedback on student behavior Explains, draws or makes behavior visual
	4. The professional applies solution-focused questions	Asks expectation questions Asks scaling questions Asks wonder-questions Focuses on solutions and strengths (re)Formulates plans in small steps
	5. The professional uses modeling (site 1) 5. The professional uses motivational interviewing techniques (site 2)	Names experiences or examples from his/her own life or work Shows alternative behavior, for example, being able to receive feedback Uses open questions Summarizes Reacts to resistance by taking the young person's position and repeats arguments Provokes change language
Goal orientation	The professional works in a goal-oriented fashion	Determines the goal together at the start of a sessions Returns to previously set goals Asks targeted questions Guides the young person back to target when he/she strays Draws up concrete actions and plans together

Acknowledgments

The authors would like to thank the staff working with interventions for young people at site 1 and 2, and the researchers and supervisors working at both organizations for their contribution to the practical implementation of the project. We thank (former) colleagues Pauline Goense, Inge Ruitenbergh, Claire Bernaards, Peter Kemper and Marinus Spreen for their contribution to the project.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by ZonMW under grant number 729102005 (2015).

ORCID iD

Leonieke Boendermaker  <https://orcid.org/0000-0002-7869-3439>

References

- Akin, B. A. (2016). Practitioner views on the core functions of coaching in the implementation of an evidence-based intervention in child welfare. *Children and Youth Services Review, 68*, 159–168. <https://doi.org/10.1016/j.childyouth.2016.07.010>
- Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Anderson Publishing Co.
- Andrews, D. A., Zinger, I., Hoge, R. D., & Bonta, J. (1996). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. In D. F. Greenberg (Ed.), *Criminal careers* (Vol. 2, pp. 437–472). Dartmouth Publishing Company.
- Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Beth Connolly Gibbons, M. (2006). The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the national institute drug abuse collaborative cocaine treatment study. *Psychotherapy Research, 16*(2), 229–240. <https://doi.org/10.1080/10503300500288951>
- Barber, J. P., Sharpless, B. A., Klostermann, S., & McCarthy, K. S. (2007). Assessing intervention competence and its relation to therapy outcome: A selected review derived from the outcome literature. *Professional Psychology Research and Practice, 38*(5), 493–500. <https://doi.org/10.1037/0735-7028.38.5.493>
- Barber, J. P., Triffleman, E., & Marmar, C. (2007). Considerations in treatment integrity: Implications and recommendations for PTSD research. *Journal of Traumatic Stress, 20*(5), 793–805. <https://doi.org/10.1002/jts.20295>
- Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Pearson.
- Bartholomew, L. K., Parcel, G. S., Kok, G., Gottlieb, N. H., Schaalma, H., Markham, C., & Partida, S. (2006). *Planning health promotion programs: An intervention mapping approach* (2nd ed.). Jossey-Bass.
- Bearman, S. K., Schneiderman, R. L., & Zoloth, E. (2017). Building an evidence base for effective supervision practices: An analogue experiment of supervision to increase EBT fidelity. *Administration and Policy in Mental Health, 44*(2), 293–307. <https://doi.org/10.1007/s10488-016-0723-8>
- Bearman, S. K., Weisz, J. R., Chorpita, B. F., Hoagwood, K., Ward, A., Ugueto, A. M., & Bernstein, A. (2013). More practice, less preach? The role of supervision processes and therapist characteristics in EBP implementation. *Administration and Policy in Mental Health, 40*(6), 518–529. <https://doi.org/10.1007/s10488-013-0485-5>
- Blase, K., & Fixsen, D. (2013). *Core intervention components: Identifying and operationalizing what makes programs work. ASPE research brief*. US Department of Health and Human Services.
- Boendermaker, L., Boomkens, C., Goense, P., & Steffens, H. (2012). *De implementatie van MDFT bij spirit in amsterdam. procesevaluatie van de periode 2010–2011*. (Implementation of MDFT by Spirit in Amsterdam. Process evaluation of the period 2010–2011). Hogeschool van Amsterdam.
- Boendermaker, L., & Goense, P. B. (2017). Meten van behandelintegriteit (Measurement of treatment integrity). In T. Van Yperen, J. W. Veerman, & B. Bijl (Eds.), *Zicht op effectiviteit* (Insight on effectivity) (pp. 375–397). Lemniscaat b.v.
- Brady, B., Canavan, J., & Redmond, S. (2016). Bridging the gap: Using Veerman and Van Yperen's (2007) framework to conceptualise and develop evidence informed practice in an Irish youth work organisation. *Evaluation and Program Planning, 55*, 128–133. <https://doi.org/10.1016/j.evalprogplan.2016.01.003>
- Cleef, M. (2015). *Julia*. Spirit Jeugd en Opvoedhulp.
- Connell, R., Fawcett, B., & Meagher, G. (2009). Neoliberalism, new public management and the human service professions: Introduction to the special issue. *Journal of Sociology, 45*(4), 331–338. <https://doi.org/10.1177/1440783309346472>
- Cunningham, P. B., Randall, J., Henggeler, S. W., & Schoenwald, S. K. (2006). Multisystemic therapy supervision: A key component of quality assurance. In T. K. Neill (Ed.), *Helping others help children: Clinical supervision of child psychotherapy* (pp. 137–160). American Psychological Association.
- Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science, 4*(1), 50.
- Dorsey, S., Kerns, S. E. U., Lucid, L., Pullmann, M. D., Harrison, J. P., Berliner, L., Thompson, K., & Deblinger, E. (2018). Objective coding of content and techniques in workplace-based supervision of an EBT in public mental health. *Implementation Science, 13*(1), 19. <https://doi.org/10.1186/s13012-017-0708-3>
- Dorsey, S., Pullmann, M. D., Deblinger, E., Berliner, L., Kerns, S. E., Thompson, K., Unutzer, J., Weisz, J. T., & Garland, A. F. (2013). Improving practice in community-based settings: A randomized trial of supervision-study protocol. *Implementation Science, 8*(1), 1–11.
- Dorsey, S., Pullmann, M. D., Kerns, S. E. U., Jungbluth, N., Meza, R., Thompson, K., & Berliner, L. (2017). The juggling act of supervision in community mental health: Implications for

- supporting evidence-based treatment. *Administration and Policy in Mental Health*, 44, 838–852. <https://doi.org/10.1007/s10488-017-0796-z>
- Feely, M., Seay, K. D., Lanier, P., Auslander, W., & Kohl, P. L. (2018). Measuring fidelity in research studies: A field guide to developing a comprehensive fidelity measurement system. *Child and Adolescent Social Work Journal*, 35, 139–152.
- Fixen, D. L., Blase, K. A., & Van Dyke, M. K. (2019). *Implementation, practice & science*. Active Implementation Research Network.
- Forgatch, M. S., & Patterson, G. R. (2010). Parent management training—Oregon model: An intervention for antisocial behavior in children and adolescents. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed., pp. 159–177). Guilford Press.
- Friele, R. D., Bruning, M. R., Bastiaanse, I. L. W., Boer, R. D., Bux, A. J. E. H., De Groot, J. F., & Hageraats, R. (2018). *Eerste evaluatie jeugdwet*. ZonMW.
- Fukink, R., Huijbregts, S., & Todd, L. (2011). Effects of video enhanced reflective practice (VERP). Current evidence and future challenges. In H. Kennedy, M. Landor, & L. Todd (Eds.), *Video enhanced reflective practice. Professional development through attuned interactions* (pp. 242–254). Jessica Kingsley Publishers.
- Garland, A. F., Brookman-Frazee, L., Hurlburt, M. S., Accurso, E. C., Zoffness, R. J., Haine-Schlagel, R., & Ganger, W. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists' offices. *Psychiatric Services*, 61(8), 788–795.
- Gingerich, W. J., & Peterson, L. T. (2013). Effectiveness of solution-focused brief therapy: A systematic qualitative review of controlled outcome studies. *Research on Social Work Practice*, 23(3), 266–283.
- Goense, P. B. (2018). *Effective reflection, developing a reflection tool*. Amsterdam University of Applied Sciences (AUAS).
- Goense, P. B., & Ruitenbergh, I. (2015). *Toolkit videofeedback in de zorg voor jeugd* (Tool kit video feedback in child and youth social care). Eburon.
- Goense, P. B., & Ruitenbergh, I. (2018). *Effective reflection, supervision with video recordings and role-play*. Amsterdam University of Applied Sciences.
- Goense, P. B., Assink, M., Stams, G. J., Boendermaker, L., & Hoeve, M. (2016). Making 'what works' work: A meta-analytic study of the effect of treatment integrity on outcomes of evidence-based interventions for juveniles with antisocial behavior. *Aggression and Violent Behavior*, 31, 106–115.
- Goense, P. B., Boendermaker, L., & Van Yperen, T. (2016). Support systems for treatment integrity. *Research on Social Work Practice*, 26(1), 69–73.
- Goense, P. B., Boendermaker, L., & Van Yperen, T. (2018). Measuring treatment integrity. *The Journal of Behavioral Health Services & Research*, 45(3), 469–488. <https://doi.org/10.1007/s11414-018-9600-4>
- Goense, P. B., Boendermaker, L., Van Yperen, T. A., Stams, G. J., & Van Laar, J. (2014). Implementation of treatment integrity procedures. *Zeitschrift Fur Psychologie / Journal of Psychology*, 222(1), 12–21. <https://doi.org/10.1027/2151-2604/a000161>
- Goense, P. B., Pronk, S., Boendermaker, L., Bakker, R., Ruitenbergh, I., & Bertling, L. (2015). *Leren op de werkvloer*. (Learning in the workplace). Eburon.
- Goldstein, A. P., Glick, B., & Gibbs, J. C. (1998). *Aggression replacement training: A comprehensive intervention for aggressive youth* (rev ed.). Research Press.
- Henggeler, S. W., Melton, G. B., Smith, L. A., Schoenwald, S. K., & Hanley, J. H. (1993). Family preservation using multisystemic treatment: Long-term follow-up to a clinical trial with serious juvenile offenders. *Journal of Child and Family Studies*, 2(4), 283–293.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (2009). *Multisystemic therapy for antisocial behavior in children and adolescents*. Guilford Press.
- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Kirk, M. A. (2020). Adaptation. In P. Nilsen & S. A. Birken (Eds.), *Handbook on implementation science* (pp. 317–332). Edward Elger Publishing.
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, 34(1), 26–38.
- Kretlow, A. G., & Bartholomew, C. C. (2010). Using coaching to improve the fidelity of evidence-based practices: A review of studies. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, 33(4), 279–299. <https://doi.org/10.1177/0888406410371643>
- Leijten, P., Dishion, T. J., Thomaes, S., Raaijmakers, M. A. J., Orobio de Castro, B., & Matthys, W. (2015). Bringing parenting interventions back to the future: How randomized micro-trials may benefit parenting intervention efficacy. *Clinical Psychology Science and Practice*, 22(1), 47–57. <https://doi.org/10.1111/cpsp.12087>
- Lyon, A. R., Stanick, C., & Pullmann, M. D. (2018). Toward high-fidelity treatment as usual: Evidence-based intervention structures to improve usual care psychotherapy. *Clinical Psychology Science and Practice*, 25(4), e12265.
- Lyon, A. R., Stirman, S. W., Kerns, S. E., & Bruns, E. J. (2011). Developing the mental health workforce: Review and application of training approaches from multiple disciplines. *Administration and Policy in Mental Health*, 38, 238–253. <https://doi.org/10.1007/s10488-010-0331-y>
- McLeod, B. D., Southam-Gerow, M. A., Rodríguez, A., Quinoy, A. M., Arnold, C. C., Kendall, P. C., & Weisz, J. R. (2016). Development and initial psychometrics for a therapist competence instrument for CBT for youth anxiety. *Journal of Clinical Child & Adolescent Psychology*, 47(1), 47–60.
- McLeod, B. D., Southam-Gerow, M. A., Tully, C. B., Rodríguez, A., & Smith, M. M. (2013). Making a case for treatment integrity as a psychosocial treatment quality indicator for youth mental health care. *Clinical Psychology Science and Practice*, 20(1), 14–32.
- Meerhoek, K., Tareé, R., Ruitenbergh, I., Van Aniel, H., & Goense, P. B. (2015). *Instructiefilm opnemen videomateriaal* (instructional recording video material). <https://www.youtube.com/watch?v=C8rPIfmVFqo>
- Metz, A. (2016). *Practice profiles: A process for capturing evidence and operationalizing innovations*. white paper. National Implementation Research Network.
- Milne, D. L. (2009). *Evidence-based clinical supervision: Principles and practice*. John Wiley & Sons.

- Milne, D. L., & Reiser, R. (2016). Supporting our supervisors: Sending out an SOS. *Cognitive Behaviour Therapy, 9*, e19.
- Murray, S., & Leadbetter, J. (2018). Video enhanced reflective practice (VERP): Supporting the development of trainee educational psychologists' consultation and peer supervision skills. *Educational Psychology in Practice, 34*(4), 397–411.
- Nadeem, E., Gleacher, A., & Beidas, R. S. (2013). Consultation as an implementation strategy for evidence-based practices across multiple contexts: Unpacking the black box. *Administration and Policy in Mental Health, 40*(6), 439–450. <https://doi.org/10.1007/s10488-013-0502-8>
- Netherlands Youth Institute. (2019). *Reform of the Dutch system for child and youth care 4 years later*. Netherlands Youth Institute.
- Netherlands Youth Institute. (n.d.). *Databank effectieve jeugdinterventies (databases effective youth interventions)*. <https://www.nji.nl/interventies>
- Ogden, T., Forgatch, M. S., Askeland, E., Patterson, G. R., & Bullock, B. M. (2005). Implementation of parent management training at the national level: The case of Norway. *Journal of Social Work Practice, 19*(3), 317–329. <https://doi.org/10.1080/02650530500291518>
- Palinkas, L. A., Aarons, G. A., Horwitz, S., Chamberlain, P., Hurlburt, M., & Landsverk, J. (2011). Mixed method designs in implementation research. *Administration and Policy in Mental Health, 38*, 44–53. <https://doi.org/10.1007/s10488-010-0314-z>
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357–367.
- Perepletchikova, F. (2007). Treatment integrity in treatment outcome research (2000-2004): Analysis of the studies and examination of the associated factors. *Dissertation Abstracts International: Section B. The Sciences and Engineering, 68*(6-B), 4139.
- Perepletchikova, F. (2011). On the topic of treatment integrity. *Clinical Psychology Science and Practice, 18*(2), 148–153. <https://doi.org/10.1111/j.1468-2850.2011.01246.x>
- Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*(6), 829–841. <https://doi.org/10.1037/0022-006X.75.6.829>
- Pronk, S. (2016). *School2Care*. Altra Onderwijs en Jeugdhulp. <https://www.nji.nl/system/files/2021-04/Uitgebreide-beschrijving-School2Care.pdf>
- Rouanet, H., Bernard, J. M., Bert, M. C., Lecoutre, B., Lecoutre, M. P., & Le Roux, B. (2000). *New ways in statistical methodology: From significance tests to Bayesian inference* (2nd ed.). Peter Lang AG.
- Sanders, M. R., Cann, W., & Markie-Dadds, C. (2003). The triple P-positive parenting programme: A universal population-level approach to the prevention of child abuse. *Child Abuse Review, 12*(3), 155–171. <https://doi.org/10.1002/car.798>
- Schoenwald, S. K. (2016). The multisystemic therapy® quality assurance/quality improvement system. In W. O'Donohue & A. Maragakis (Eds.), *Quality improvement in behavioral health* (pp. 169–192). Springer.
- Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health, 38*(1), 32–43.
- Schoenwald, S. K., Garland, A. F., SouthamGerow, M. A., Chorpita, B. F., & Chapman, J. E. (2011). Adherence measurement in treatments for disruptive behavior disorders: Pursuing clear vision through varied lenses. *Clinical Psychology Science and Practice, 18*(4), 331–341.
- Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services, 52*(9), 1190–1197.
- Schoenwald, S. K., Sheidow, A. J., & Chapman, J. E. (2009). Clinical supervision in treatment transport: Effects on adherence and outcomes. *Journal of Consulting and Clinical Psychology, 77*(3), 410–421.
- Thijssen, J., & De Ruiter, C. (2010). Strenge bewaking van kwaliteit bij PMTO. experimenteel onderzoek naar PMTO in Nederland. (Strict security of the quality of PMTO, Experimental research in PMTO in The Netherlands). *Kennis, 4*(4), 8–14.
- Topor, D. R., AhnAllen, C. G., Mulligan, E. A., & Dickey, C. C. (2017). Using video recordings of psychotherapy sessions in supervision: Strategies to reduce learner anxiety. *Academic Psychiatry, 41*(1), 40–43.
- Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling & Development, 93*(4), 403–411.
- Van Yperen, T., Veerman, J. W., & Van Den Berg, G. (2015). *Elementen die er toe doen: Overzicht van begrippen over werkzame elementen en een voorstel voor een indeling. (Elements that matter: Overview of concepts of effective elements and suggestion for layout)*. Nederlands Jeugdinstituut.
- Veerman, J. W., & van Yperen, T. A. (2007). Degrees of freedom and degrees of certainty: A developmental model for the establishment of evidence-based youth care. *Evaluation and Program Planning, 30*(2), 212–221.
- Visscher, L., Evenboer, K. E., Jansen, D. E., Scholte, R. H., Knot-Dickscheit, J., Veerman, J. W., Reijneveld, S. A., & van Yperen, T. A. (2018). Identifying practice and program elements of interventions for families with multiple problems: The development of a taxonomy. *Children and Youth Services Review, 95*, 64–70. <https://doi.org/10.1016/j.childyouth.2018.10.030>
- Wandersman, A., Chien, V. H., & Katz, J. (2012). Toward an evidence-based system for innovation support for implementing innovations with quality: Tools, training, technical assistance, and quality assurance/quality improvement. *American Journal of Community Psychology, 50*(3–4), 445–459.