Developing Usable Innovations

Developing Usable Innovations

Karen A Blase, Dean L Fixsen, Melissa Van Dyke

Active Implementation Research Network
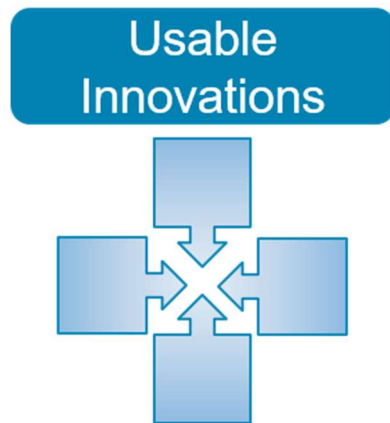
Usable innovations are operationalized so they are teachable, learnable, doable, and assessable in practice. Usable innovations are effective when they are used as intended and they have a way to detect the presence and strength of the innovation as it is used in everyday practice.

Innovations have been the focus of implementation efforts in all fields of endeavor in human services, social sciences, agriculture, business, computing, engineering, manufacturing, and so on. Standard practices are what we do every day. Innovations are new and are deviations from standard practice. Klein and Sorra (1996) define an innovation as "a technology or practice that an organization is using for the first time, regardless of whether other organizations have previously used the technology or practice." Nord and Tucker (1987) and Rogers (1995) have offered similar definitions of an innovation.

All innovations are not created equal. Some are complex (the internet), some are technical (the internal combustion engine), and some are simple (the plow). A big problem at the practice level is trying to figure out what is supposed to be done when attempting to use an innovation for the first time. Anyone who has read a manual to use an innovation or attempted to work with a child with autism or attempted to provide emergency obstetric services has experienced the problem. If you don't know what to do it will be difficult to do it and even if you are somehow successful it will be virtually impossible to replicate the innovation to repeatedly realize the intended outcomes. As described by Teague, Bond, and Drake (1998, p. 217), "without detailed descriptions of interventions, replication is difficult; without reliable measurement of interventions, conclusions about presence or absence of effects are questionable."

For the past few decades, evidence-based practices have been highlighted as effective innovations. The emphasis has been on the rigor of the research methods used to produce evidence in support of an innovation. Good evidence is important. If an organization is going to select an innovation for use, there needs to be some indication that the effort will be "worth it" in terms of outcomes. However, rigor is insufficient to support the everyday use of innovations.

Usable Innovation criteria

While debates about evidence-based programs have sharpened the focus on research methods and internal validity, the definition of an innovation has not been part of the discussion. From an implementation point of view, the definition of an innovation is critical so that it is teachable, learnable, doable, and assessable in practice. To produce socially significant outcomes, a Usable Innovation meets four criteria:

1. Clear *description* of the innovation
   a. Philosophy, Values, and Principles
      i. The philosophy, values, and principles that underlie the innovation provide guidance for all innovation-related decisions and evaluations; and are used to promote consistency, integrity, and sustainable effort across all organization units.
      ii. For example, the Teaching-Family Model philosopy emphasizes ecologically appropriate treatment (family, peer, school, community); promotes values that include providing care and treatment that is humane, effective, individualized, satisfactory to recipients and consumers, cost efficient, and replicable; and uses principles derived from applied behavior analysis concerning teaching appropriate alternative behavior, positive motivation, and self government (Phillips, Phillips, Fixsen, & Wolf, 1974; Wolf, Kirigin, Fixsen, Blase, & Braukmann, 1995)
   b. Inclusion and exclusion criteria that define the population for which the innovation is intended
      i. The criteria define who is most likely to benefit when the innovation is used as intended, and who is not likely to benefit.
      ii. For example, Multisystemic Therapy (MST) includes youths who are a) serious juvenile offenders b) at imminent risk of placement in residential care (incarceration) and c) living with at least one parent or adult caregiver (Schoenwald, Brown, & Henggeler, 2000)
      iii. In other examples, Chinman, Imm, and Wandersman (2004) stated it is important to be as specific as possible when describing whom the innovation is intended to serve. Feldstein and Glasgow (2008) suggested the targeted populations' range of characteristics must be considered, such as age, gender, socioeconomic status, literacy, native language, and culture. Recipients often have competing demands for their attention, and pre-existing health conditions or family or work demands may make it physically challenging to follow through with encouraged actions. Thus, relevant aspects of intended recipients are a characteristic of the definition of an innovation.
2. Clear description of the *essential functions* that define the innovation

a. Essential functions are the features that must be present to say that an innovation exists in a given location (sometimes called core intervention components, active ingredients, or practice elements). Embry and Biglan (2008, p. 75) describe essential functions as evidence-based kernels that are "fundamental units of behavioral influence that appear to underlie effective prevention and treatment for children, adults, and families. A kernel is a behavior–influence procedure shown through experimental analysis to affect a specific behavior and that is indivisible in the sense that removing any of its components would render it inert."

b. For example, McHugo, Drake, Teague, and Xie (1999, p. 820) describe the nine essential components of Assertive Community Treatment as community locus, assertive engagement, high intensity, small caseload, continuous responsibility, staff continuity, team approach, multidisciplinary staff, and work closely with support systems.

c. In other examples, Sandler et al. (2005) noted that it is a challenge to describe an innovation in enough detail to get an accurate understanding of its core components in practice. Unlike physical products where drawings or simple models of the product can be produced, human service descriptions rely on written concept statements and outlines. Elwyn, Taubert, and Kowalczuk (2007) warned against causal ambiguity where the precise reasons for successful outcomes are not understood. Problems using an innovation in new settings will be difficult to resolve if the essential functions of the innovation itself are not clear.

Damschroder et al. (2009), Kilbourne, Neumann, Pincus, Bauer, and Stall (2007), Szulanski (1996), and others relate adaptability and refinement of innovations to the clarity of the core components, "the essential and indispensable elements" of the innovation itself. Greenhalgh, Robert, MacFarlane, Bate, and Kyriakidou (2004) call the core components the "hard core" and the adaptable elements the "soft periphery." These authors agree that adaptation is meaningful only when the core elements are known and understood. As Kilbourne et al. (2007) stated, "having the core elements detailed, while also providing options for implementing these core elements, is vital for optimizing both fidelity to the intervention and flexibility in its implementation."

Blase and Fixsen (2013) state that "rather than being based on hunches and best guesses, intervention programs are increasingly expected to be evidence-based. However, when evidence-based programs are replicated or scaled up, it is critical not only to know whether a program works, but which program elements are essential in making the program successful. To date, though, few programs have had hard data about which program features are critical —core components‖ and which features can be adapted without jeopardizing outcomes."

3. *Operational definitions* of the essential functions

a. Practice profiles (Blase, Metz, Bartley, & Fixsen, 2011; Tilly III, 2008) describe the essential functions in terms of activities that allow an innovation to be teachable, learnable, doable, and assessable in practice; and promote consistency across practitioners at the level of actual service delivery (also known as innovation configurations; Hall & Hord, 2011; Hord, Rutherford, Huling-Austin, & Hall, 1987).

b. For example, teaching appropriate alternative behavior is an essential function of the Teaching-Family Model (Braukmann, Kirigin Ramp, Braukmann, Willner, & Wolf, 1983; Phillips et al., 1974); a teaching interaction is operationally defined as:

a. Qualitative Components
   i. Use a calm, caring speaking voice
   ii. Be enthusiastic and positive when praising
   iii. Be calm and matter of fact when offering corrective feedback
   iv. Stay in close proximity
   v. Use polite and pleasant requests (please..., would you...)

b. Behavior Components
   i. Initial positive statement
      1. statement of praise, empathy, affection
      2. set a positive tone for the interaction
   ii. Name the skill (use a concept label)
      1. the focus for the interaction
   iii. Describe the inappropriate behavior (reactive teaching only)
      1. specific description (a replay of what happened or was omitted)
      2. demonstrate what cannot be described (facial expressions, gestures)
      3. no blaming or mocking the youth (be non-judgemental)
   iv. Describe the negative consequence (reactive teaching only)
      1. loss of access to privileges, points, checkmarks
      2. positive correction statement (earn back half the loss by practicing now)
   v. Describe the appropriate behavior
      1. restate the skill label
      2. specific description (exactly what is expected)
      3. demonstrate what cannot be described (voice tone, facial expressions)
   vi. Give a rationale
      1. brief, personal, believable statement
      2. point out short-term natural benefits or harms of the skill
      3. link skill label, behaviors, and outcomes
   vii. Request acknowledgement
      1. check for understanding

viii. Practice
1. state skill label
2. describe/demonstrate appropriate behavior components
3. set up the practice "scene"
4. act out the scene

ix. Practice feedback
1. effective praise for the parts of practice that went well
2. corrective feedback for the parts that need improvement
3. repractice to criterion/youth comfortable with new skill
4. positive consequences (positive correction: earn back half of any loss)

x. General praise
1. descriptive praise for engaging in the interaction (e.g. cooperation)
2. encourage the effort to learn

Aarons, Hurlburt, and Horwitz (2011) advocate for a "high degree of procedural specificity in work activities." Chinman et al. (2004) and Greenhalgh et al. (2004) recommend the development of "innovation configurations" as described by Hall and Hord (2011). Innovation configurations (also known as practice profiles) provide specific examples of expected, developmental, and unacceptable behavior related to each essential function of an innovation.

4. A practical assessment of fidelity
   a. A fidelity assessment relates to the innovation philosophy, values, and principles; essential functions; and core activities specified in the practice profiles; and a fidelity assessment is practical and can be done repeatedly in the context of typical human service systems.
   b. A fidelity assessment provides evidence that the innovation is effective when used as intended; that is, the fidelity assessment is highly correlated with intended outcomes with evidence that shows high fidelity use of an innovation produces desirable outcomes and low fidelity use does not.
   c. For example, Forgatch and DeGarmo (2011, p. 238) outline a fidelity measure (called FIMP) for the Parent Management Training Oregon model (PMTO) that assesses:
      i. Knowledge: Demonstrated understanding of PMTO content and theoretical principles.
      ii. Structure: Ability to accomplish agenda activities and goals while addressing family issues. Includes maintaining orderly flow, leading without dominating, responsiveness to family, good transitions, and sensitive timing and pacing.
      iii. Teaching: Proficiency in strategies that promote parents' mastery and use of PMTO practices. Verbal teach includes standard pedagogical tactics (give information, make suggestions); active teach engages families in the learning process by brainstorming, role playing, and eliciting solutions.

  iv. Process: Provides support that promotes a safe and supportive learning context. Includes questioning that leads to insight, maintaining balance among participants, encouraging skill development, joining family's storyline.

  v. Overall Development: Promotes family's growth in PMTO use. Includes likelihood that family can/will use procedures, family's apparent satisfaction, likelihood of continuing, managing unique/difficult aspects of contexts/issues.

  FIMP ratings are based on time samples of therapy sessions in which two core parenting practices are delivered: skill encouragement and limit setting. Two sessions are rated for each component, one introducing the component and another troubleshooting that component. For practtitioner certification purposes, full sessions are rated. For research and reliability assessment, segments of approximately 10 min are sampled from video-recorded family intervention sessions. To identify segments for rating, trained assistants spot-check tapes labeled with topics, seeking segments of approximately 10 min with content on a relevant component (i.e., skill encouragement or limit setting) and teaching activity (e.g., debriefing home practice, role playing, brainstorming for incentives or negative consequences). FIMP raters are required to be certified PMTO practitioners and are familiar with PMTO manuals and practices.

 d. Graham et al. (2006) noted it is important to define what constitutes knowledge (the core components, operationalized) so that it can be measured in applications. A fidelity measure can be used to determine whether knowledge-based innovations have been sufficient to bring about the desired change or whether new interventions may be required. Klein and Sorra (1996) point to the ultimate criterion – the extent to which intended outcomes of an innovation, used as intended, are realized in the organization. Fidelity is one of the Implementation Drivers. Here, we are speaking of fidelity as part of the definition of an innovation. If there is no way to assess fidelity (intended use in practice), the innovation is not a Usable Innovation.

Operationalizing Interventions

As noted by Dane and Schneider (1998) and Michie, Fixsen, Grimshaw, and Eccles (2009) there is little empirical evidence to support assertions that the components named by an evidence-based program developer are, in fact, the essential functions necessary for producing the intended effects. There may be other un-named, un-measured components involved in a treatment that actually produce the effects, and the components identified by the program developer may or may not be important to the outcomes achieved. The mention or lack of mention of certain components by a developer should not be confused with their function or lack of function in an innovation-based exchange.

Fixsen, Blase, Duda, Naoom, and Van Dyke (2010) described the methods used to develop well-described programs so they can be implemented on a scale sufficient to solve social problems.

To be useful to consumers and functional across thousands of practitioners and organizations operating in locations across the country, Implementation Teams need to know what to train, what to coach, and what performance to assess to make full and effective use of an evidence-based program or other innovation. Implementation Teams need to know what "it" (essential functions) is so they can efficiently and effectively assure proper use of the intervention now and improve "it" over time.

For the vast majority of evidence-based programs and other innovations, usability testing and PDSA improvement cycles can be used to rapidly define them to meet the four program criteria. Usability testing methods were developed by computer scientists as a way to de-bug and improve complex software programs or websites. PDSA cycles grew out of the work of applied researchers at the Bell Labs in the 1920s as they sought to develop more reliable and error free communication equipment and networks. Both methods are readily useful in human service environments (e.g. Institute for Healthcare Improvement, 2010). Fraser and Galinsky (2010, p. 462) note that "The development of an intervention takes place over a series of studies that are sequenced from less-controlled pilot tests to more-controlled efficacy and effectiveness tests. However, negative findings at any point may be cause for reconceptualization of the intervention design. Thus, the process is not linear. It has a recursive feature in which, though progress may be made over time, an intervention may be revised and retested iteratively until it reaches a benchmark for efficacy (e.g., an effect size comparable to or greater than effects observed with other interventions in the field of practice)."

Usability Testing and Program Development

For programs (evidence-based or otherwise) that are poorly defined, testing must be done to gain the experience and information to meet the four criteria for a Usable Innovation. Usability testing is an efficient and effective method for operationalizing a program and for establishing the necessary implementation supports in practice.

Usability testing (Akin et al., 2013; Nielsen, 2005) employs a small number of participants for the first trial, assesses results immediately, makes corrections based on those results, and plans and executes the next iteration. This process is repeated (say, 5 times with 4 participants in each iteration for a total N = 20) until the intervention is producing credible outcomes. The "trial and learning" embodied in the usability testing approach allows Implementation Teams and others to quickly adjust the "program" with each iteration so each small group is testing a new and, hopefully, improved version. Based on usability testing outcomes in other fields we might expect the first iteration to reveal about 30% of the information we need to know. Each iteration after that results in incremental improvements needed before proceeding to a larger scale.

Marshall et al. (2017, p. 578) describe their attempts "as an experienced team of practitioners, improvers, commissioners and evaluators to design an effective intervention to improve the safety of people living in care homes in England. We highlight how the design of the

intervention, as described in the original grant proposal, changed significantly throughout the initiative. We outline how the changes that were made resulted in a more effective intervention but how our failure to design a better intervention from the start reduced the overall impact of the project." Problem analysis (Atkins et al., 2017) and detailed reviews of ongoing practices and outcomes (Lewis, 2014) are ways of conducting usability testing to design better innovations.

Usability Testing and Implementation Supports

The usability testing approach simultaneously allows for repeated assessments and improvements in implementation supports – the "execution" or "do" part of usability testing. Each iteration allows for adjustments in training, coaching, and performance assessments as well as the program / intervention itself.

The benefits of the PDSA cycle in highly interactive environments have been evaluated across many domains including manufacturing, health, and substance abuse treatment. This "trial-and-learning" approach allows developers of complex human service programs and Implementation Teams to identify the essential functions of an intervention itself as they evaluate "what works" and discard non-essential components.

When applied to human services in a usability testing format, the "plan" can be the intervention as it is intended to be used in practice. To carry out the "do" part of the PDSA cycle, the "plan" needs to be operationalized (what we will do and say, with whom, where, and when, to enact the plan). This compels attention to the essential functions of an innovation and provides an opportunity to begin to develop a training and coaching process (e.g. this is how to do the plan) and to create an assessment of fidelity (e.g. did we "do" the "plan"). As 3 or 4 newly trained practitioners begin working with recipients in an actual service environment, the budding fidelity measure can be used to interpret the outcomes in the "study" part of the PDSA cycle (e.g. did we do what we intended; did doing what we intended result in desired outcomes). Once the results are known (fidelity assessment and recipient outcomes), the Implementation Team can adjust the program and adjust the implementation supports for the next cohort of 3 or 4 practitioners and begin the next iteration.

The similarities between the PDSA cycle and usability testing are obvious. The difference is that usability testing includes 3-5 participants (e.g. practitioners, service units) in each cycle to further maximize the learning from the trial and learning approach. In human services, program developers and Implementation Teams may employ several iterations to arrive at a functional version of a Usable Innovation that is effective in practice and can be used with recipients as intended on a socially significant scale (Fixsen, Blase, Timbers, & Wolf, 2007; Wolf et al., 1995).

For the 70% to 95% of innovations, and nearly 100% of standard practices, that have not been operationalized by the developers, Implementation Teams will need to make use of usability

testing to uncover the essential functions that are teachable, learnable, doable, and assessable in practice before they can proceed with broader scale implementation.

Improvement Cycles in practice

An example of an approach to establishing usable interventions and implementation supports is provided below. Note how PDSAC is used on purpose to develop simultaneously the innovation and the implementation supports for the innovation in an education context (Fixsen, Hassmiller Lich, & Schultes, 2018).

The usability testing process outlined below employed 9 teachers over the course of 4 months. In a usability testing format, the Implementation Team worked intensively with 3 teachers at a time to maximize the learning and to quickly make use of learning in the work with the next group of 3 teachers. This provides three times more learning and improvement opportunities for the Implementation Team compared to one experience with 9 teachers.
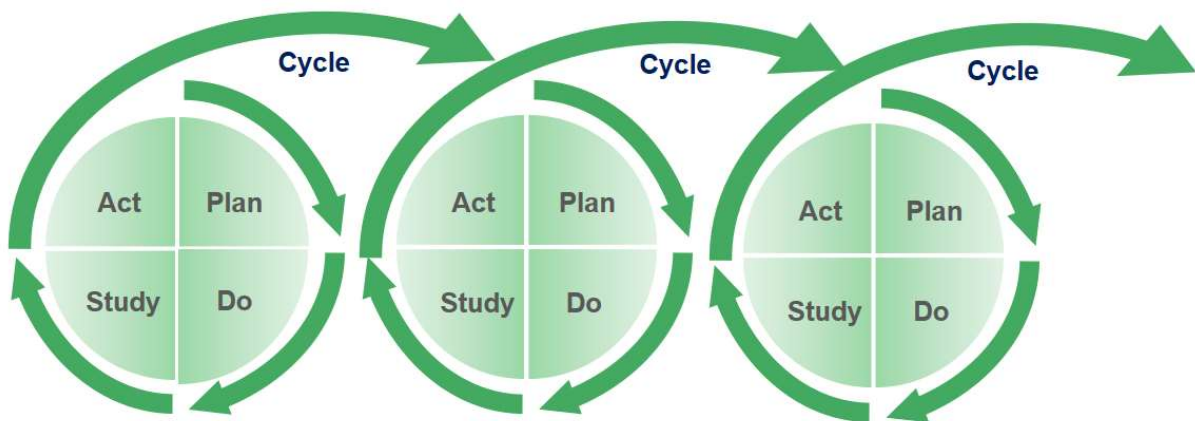
Figure 1. Usability testing with repeated iterations to develop a solution to a problem or achieve an aspiration.

Iteration #1

Plan: The state legislature just passed a law mandating new standards for grade 3 literacy. The state department of education asked faculty of the state university to summarize the research on early literacy instruction with an emphasis on instructional practices that might be useful for children and students from age 3 through grade 3. The research summary specified the following two instruction practices found to be effective in the literature (e.g. Hattie, 2009):

- Effective instructors encourage high levels of student engagement with education content

- Effective instructors provide frequent, prompt, and accurate feedback to students when they respond

A Practice Profile was drafted to identify expected, developmental, and poor examples of instructional behavior related to each of the instruction practices.

Do: To begin the process, the Implementation Team contacted a nearby school district. After some Exploration Stage work with principals and teachers, they secured the cooperation of 9 K-3 teachers and their principals. The teachers agreed to try to use the instruction methods, participate in training, allow two people to observe their classroom every day for two weeks, give students a weekly quiz related to literacy content taught that week, and participate in up to one hour of de-briefing discussion during each week. In a meeting with the teachers and their principals, a schedule was developed so teachers 1-3 would participate during Month 1, teachers 4-6 would begin to participate in Month 2, and teachers 7-9 would begin to participate in Month 3.

Just prior to Month 1, the Implementation Team developed a two-hour training workshop to review and discuss the literature regarding the two key instruction practices, created video tapes to model the two key components, and developed "behavior rehearsal scenes" to provide opportunities for teachers to practice the skills in a mock classroom. At the beginning of Month 1 the Implementation Team provided the training to teachers 1-3 and debriefed with the teachers at the end of training to obtain their opinions of the training methods and content.

Prior to Month 1, the Implementation Team drafted 4 fidelity items to assess the use of the two key instruction practices. During the behavior rehearsal section of training, one member of the Implementation Team used the items to observe teacher instruction in the mock classroom. The items were modified based on those observations. The "fidelity scores" related to teacher instruction at the end of training were analyzed to see how training could be improved next time.

Immediately after training, the 3 teachers began using the instruction practices in their classrooms. Starting on the third day and every other day thereafter, the Implementation Team observed each classroom for 2 hours with two members of the Team simultaneously observing one classroom at a time. The Team members used the Practice Profile outline to note instances of expected, developmental, and poor examples of instruction. At the end of week 1 and again at the end of week 2, two members of the Implementation Team did a teacher instruction fidelity assessment using the 4 items developed prior to training and modified during training. Each teacher provided the Implementation Team with the average scores for the weekly student quiz related to literacy content taught that week.

At the end of each week, 2 Implementation Team members met with the 3 teachers as a group to discuss the instruction practices. Teachers provided their perspectives on what was easy or difficult for them to do in their interactions with students. Implementation Team members offered suggestions for using the instruction practices based on their observations of all 3

teachers. Implementation Team members began drafting a coaching service delivery plan based on teachers' input.

Study: At the end of weeks 2 and 3 the Implementation Team met to consider the information being developed. The information and data being gained from the experience with the first 3 teachers were used to revise the innovation and improve implementation supports as noted in the Act section.

Act: Based on classroom observations and comments from teachers, the Implementation Team re-defined the key instruction components of the innovation. The Implementation Team expanded the component, "Instructors encouraging high levels of student engagement with education content" to include "provides explicit instruction" and "models instruction tasks." The Implementation Team drafted a Practice Profile (including the new components) with detail based on the classroom observations. A draft of the Practice Profile was reviewed with the 3 teachers and their ideas were included regarding how to define expected, developmental, and poor examples of use of each component of the innovation.

The Implementation Team compared notes on the fidelity assessments to see if they agreed or not on scoring each of the 4 items. Agreement was not good so the fidelity items were revised to be more specific and the number of items was increased to include the new components being operationalized in the Practice Profile. A protocol for how a fidelity observer should enter the classroom and conduct the observation was drafted for use in subsequent fidelity observations. The fidelity scores and the scores for the weekly student quizzes were summarized. No discernable relationship between the two was apparent.

As noted above, the Implementation Team began studying training during and after the training session for teachers 1-3. In week 3 the Team began work on how to improve training methods and how to include the new content in training for the next 3 teachers.

Iteration #2

Plan: The Implementation Team met with the principal and teachers to set the time for a two-hour training workshop for teachers 4-6. The Implementation Team discussed the work during Month 1 and invited questions about the classroom observations and the de-brief times.

Do: In Month 2, the Implementation Team provided the revised training to teachers 4-6. The training content was based on the expanded essential components. The revised training methods were based on the experience and feedback from teachers 1-3.

The Implementation Team provided a two-hour training workshop to review and discuss the literature regarding the key instruction practices, model the key components, and provide opportunities for teachers 4-6 to practice the skills in a mock classroom. During training, practice continued until the teachers felt competent and confident. The Implementation Team

debriefed with the teachers at the end of training to obtain their opinions about the training methods and content.

During the behavior rehearsal section of training, one member of the Implementation Team used the revised fidelity items to observe teacher instruction in the mock classroom. The fidelity items were modified further based on those observations.

To collect pre-post training data, a version of the behavior rehearsal (used in training) was conducted individually for each teacher just prior to training. The teacher's behavior was scored using the fidelity criteria. The scores for each fidelity item prior to training and during the last behavior rehearsal at the end of training were analyzed to see the extent to which teachers improved instruction skills during training. The data provided direction on how training could be improved next time.

Immediately after training, teachers 4-6 began using the instruction practices in their classrooms. Starting on the third day and every other day thereafter, the Implementation Team observed each classroom for 2 hours with two members of the Team jointly observing one classroom at a time. For teachers 1-3 one observation per week was conducted. During the observations, the Team members used the Practice Profile outline to note instances of expected, developmental, and poor examples of instruction.

Two members of the Implementation Team did a fidelity assessment. The new fidelity assessment was used for assessments of teachers 1-6 each week to gain more experience with the items and to continue to develop the observation protocol. Each teacher provided the Implementation Team with the average scores for the weekly student quiz related to literacy content taught that week.

At the end of each week, two Implementation Team members met with the 6 teachers to discuss the instruction practices. Teachers provided their perspectives on what was easy or difficult for them to do. Implementation Team members offered suggestions for using the instruction practices based on their observations of all 6 teachers. Implementation Team members revised the coaching service delivery plan based on teachers' input.

Study: The Implementation Team now has two months of information from teachers 1-3 and one month of information from teachers 4-6. In Month 2, Teachers 1-3 were gaining experience and using the innovation with confidence in their interactions with students. The Implementation Team began seeing more nuanced versions of the 4 key components of the innovation.

The pre-post training data were summarized to see where training produced more or less improvement in teachers learning the instruction skills. Those data were compared to the ongoing fidelity assessments to see if the post-training scores for teachers predicted later fidelity scores.

The fidelity scores for the 6 teachers and the scores for the weekly student quizzes were summarized. A pattern emerged indicating a possible relationship between higher fidelity scores and better scores on student quizzes.

Act: Based on observations and teacher comments, the Implementation Team again re-defined the key instruction components of the innovation. The Implementation Team expanded the component, "Effective instructors provide frequent, prompt, and accurate feedback to students when they respond" to include "corrects errors by modeling a correct response" and "limits corrective feedback to the task at hand." These new components were included in the draft Practice Profile. The draft of the Practice Profile was reviewed with the 6 teachers and their ideas were included regarding how to define expected, developmental, and poor examples of use of each component of the innovation.

The Implementation Team compared notes on the fidelity assessments to see if they agreed or not on scoring each of the items. The items were revised to be more specific and the number of items was increased to include the new components being operationalized in the Practice Profile. The protocol for how a fidelity observer should enter the classroom and conduct the observation was revised based on the experiences with all 6 teachers.

The pre-post training data summary made it clear that trainers were more effective when teaching the instruction components related to delivering information to students. However, the trainers were producing mixed outcomes when teaching instruction components related to providing feedback to students after they responded. The Implementation Team developed new behavior rehearsal scenarios to provide more training on those skills.

Iteration #3

Plan: The Implementation Team met with the principal and teachers to set the time for a two-hour training workshop for teachers 7-9. The Implementation Team discussed the work during Months 1 and 2 and invited questions about the classroom observations and the de-brief times.

Do: In Month 3, the Implementation Team provided the revised training to teachers 7-9. The training content was based on the expanded essential components and practice profiles. The revised training methods were based on the experience and feedback from teachers 1-6. The Implementation Team debriefed with the teachers at the end of training to obtain their opinions of the training methods and content.

During the behavior rehearsal section of training, one member of the Implementation Team used the revised fidelity items to observe teacher instruction in the mock classroom. The fidelity items were modified further based on those observations.

Pre-post training data were collected by using a version of the behavior rehearsal (used in training) individually for each teacher just prior to training. The teacher's behavior was scored

using the revised fidelity criteria. The scores for each fidelity item prior to training and during the last behavior rehearsal at the end of training were analyzed to see the extent to which teachers improved instruction skills. The data provided direction on how training could be improved next time.

Immediately after training, teachers 7-9 began using the instruction practices in their classrooms. Starting on the third day and every other day thereafter, the Implementation Team observed each classroom for 2 hours with two members of the Team simultaneously observing one classroom at a time. For teachers 1-6 one observation per week was conducted. During the observations, the Team members used the Practice Profile outline to note instances of expected, developmental, and poor examples of instruction.

For teachers 1-9, at the end of week 1 and again at the end of week 2 two members of the Implementation Team did a fidelity assessment. The revised fidelity assessment was used for assessments of teachers 1-9 each week to gain more experience with the items and to continue to develop the observation protocol.

At the end of each week, two Implementation Team members met with the 9 teachers to discuss the instruction practices. Teachers provided their perspectives on what was easy or difficult for them to do. Implementation Team members offered suggestions for using the instruction practices based on their observations of all 9 teachers. Implementation Team members revised the coaching service delivery plan based on teachers' input.

Study: The Implementation Team now has three months of information from teachers 1-3, two months of information from teachers 4-6, and one month of information from teachers 7-9. With daily use of the new instruction methods in the classroom, teachers 1-6 were using the innovation with confidence in their interactions with students. As each teacher "made the new skills her own," the Implementation Team began seeing nuanced versions of the key components of the innovation.

Fidelity scores for teachers 1-3 and 4-6 seemed to be improving from the first week after training to Month 3. The continued revision and expansion of the fidelity items made these data difficult to interpret, but the impression from observations and teacher reports seemed to confirm the fidelity information. The fidelity scores and the scores for the weekly student quizzes were summarized. Analysis of Month 3 data for all 9 teachers resulted in a positive correlation of 0.50 between fidelity scores and student quiz outcomes.

For two teachers in the 4-6 group, fidelity scores were good and their student outcomes were outstanding! The Implementation Team and teachers met to review the classroom observations and to engage the teachers in discussion of their instruction practices. It turned out that in the previous year these two teachers had been mentored by the same master teacher. During their induction into teaching, they had been taught to stand by the door and greet each student by

name as he/she entered the classroom at the start of the school day and again after lunch period (Embry & Biglan, 2008). They felt this "primed the pump" and helped with student engagement.

The pre-post training data were summarized to see where training produced more or less improvement in teachers learning the instruction skills. Those data were compared to the ongoing fidelity assessments to see if the post-training scores for teachers predicted later fidelity scores.

Act: Based on observations, the Implementation Team again re-defined the key instruction components of the innovation. The Implementation Team expanded the key components to include greeting each student by name at the beginning of the school day. This new component was included in the draft Practice Profile. The draft of the Practice Profile was reviewed with the 9 teachers and their ideas were included regarding how to define expected, developmental, and poor examples of use of each component of the innovation.

The Implementation Team compared notes on the fidelity assessments to see if they agreed or not on scoring each of the items. The items were revised to be more specific and the number of items was increased to include the new "greeting component" being operationalized in the Practice Profile. The protocol for how a fidelity observer should enter the classroom and conduct the observation was revised based on the experiences with all 9 teachers.

The pre-post training data summary showed that trainers produced better outcomes when teaching instruction components related to providing feedback to students after they responded. However, there was need for further improvement. The Implementation Team decided to revise how they were giving feedback to teachers during training (e.g. focus comments on the positive behavior; model expected behavior prior to asking the teacher to practice again) during the behavior rehearsal scenarios.

Cycle: After 4 months, the Implementation Team was refining the fine points of the Practice Profile, assessing pre-post training knowledge and skills of teachers participating in training, using a good set of items to assess instruction practices in the classroom, and collecting information to correlate fidelity scores with student quiz scores. The innovation still needed improvement but met the basic criteria for a usable innovation.

Developing Practice Profiles

The outline in this section originally was developed by the Minnesota Implementation Team for developing a Practice Profile in an education context. The process is the same in other contexts although the participants and roles change (Blase et al., 2011).

Practice Profiles identify the essential functions or core components or critical features (these terms are used interchangeably) for an instructional strategy/practice or Usable Intervention (these terms are used interchangeably). A Practice Profile operationalizes each core component and describes educator behaviors along a continuum from expected or proficient behavior, to developmental, to needs improvement. This document can be used to help guide a team's creation of a Practice Profile for an instructional strategy/practice or Usable Intervention.

Purpose

Practice Profiles are tools that can be used by educators to promote skills, competence and confidence in effective implementation of a strategy/practice or Usable Intervention. Use of a Practice Profile is intended to:

- Increase fidelity to the practice and increase likelihood of implementing the practice as it is intended
- Clearly define expectations and operationalize adult behaviors - to increase the likelihood that everyone knows the expectations when using an instructional strategy/practice
- Develop common understanding of what the practice looks like when adults are "doing it well"
- Guide purposeful coaching
- Be reviewed periodically and be modified over time according to feedback and usability testing

Six Areas to Address in Developing a High Quality Practice Profile

Six major areas have been identified as being critical to the development of a high quality or successful Practice Profile: 1) Prerequisites, 2) Resources, 3) Teaming, 4) Content, 5) Process, and 6) Alignment.

Area 1) Prerequisites
What prerequisites should be addressed prior to developing the Practice Profile? Understanding the purpose and context for developing the Practice Profile must be addressed prior to developing a Practice Profile. Potential coaching questions during this time are listed below.

- What are the essential functions/core components of the instructional practice/strategy or useable intervention – how will they be identified?

- What does the research identify as the critical features of the evidence-based instructional strategy or practice?
- What is the scope of the profile that will be developed (specific group, content, grade level)?
- How will the Practice Profile align with the instructional strategy/practice or framework being implemented?
- How will understanding of the importance of developing a Practice Profile be communicated to the team and school staff?
- When will the Practice Profile be developed – how will it fit with ongoing practice improvement initiatives?
- How will the team ensure that the strategy is teachable, learnable, doable, and assessable/measurable in practice?

Area 2) Resources
What are some resources that might be used to develop content of Practice Profiles?

- Examples of (vetted) Practice Profiles
- Research based instructional strategies that meet the identified need (include the citation of the resources used on the Practice Profile)
  - Hattie, J. (2012) "Visible Learning for Teachers - Maximizing Impact on Learning."
- Video clips of instructional strategies specific to content area – examples of what good instruction looks like.

Area 3) Teaming
Who will do the work of developing Practice Profile? Depending on the size of the Implementation Team, a smaller task group may be assigned to lead the Practice Profile work.

- A task group is typically assigned to lead the Practice Profile work and develop the initial draft, and make any final revisions once feedback has been received. It is important to utilize experts in the identified area of need (e.g. Reading, Special Education, Math). Multiple perspectives are important. The task group can complete the entire Practice Profile for the instructional strategy/practice, or core components or critical features of the instructional strategy/program may be assigned to pairs of individuals who work to create the first draft and then bring it back to the larger task group for feedback and review.
- The Implementation Team is used to provide feedback on the draft Practice Profile and can be used for providing final approval of the Practice Profile once it is completed.
- School advocate may act as a coach during the Practice Profile development and ensure that linked teams are utilized.

Area 4) Content
What does the Practice Profile include?

Practice Profiles describe the essential components or critical features of a specific instructional strategy/practice or Usable Intervention (e.g., close reading). Practice Profiles may also describe research-based pedagogical practices - such as embedded learning targets, formative assessment, or reciprocal teaching, which can apply to all content areas. The format is shown in the table below.

Core Component or Critical Feature	Contribution to the Desired Outcome	Expected/Proficient: Description of educator behavior	Developmental: Description of educator behavior	Needs Improvement: Description of educator behavior
	For each component, include an explanation of how it contributes to the desired outcome	List indicators. Includes activities and describes behaviors that exemplify educators who are able to generalize required skills and abilities to wide range of settings and contexts; skills are used consistently and independently – skills are sustained over time while continuing to grow.	List indicators. Includes activities and describes behaviors that exemplify educators who are able to implement required skills and abilities but in a more limited range of contexts and setting – skills are used inconsistently or need coaching to complete or successfully apply particular skills for improvement in order to move into expected. This column helps to define the focus of coaching.	List indicators. Includes activities and describes behaviors that exemplify educators who are not yet able to implement required skills or abilities in context.

As noted in the table, the content of a Practice Profile includes:

a) Developmental Levels of Proficiency - clear description of headings (across the top of the Practice Profile).

- Expected/Proficient: Instruction that exemplifies educators who are able to use the instructional strategy/practice to meet the needs of all learners through differentiation. Uses strategy consistently and in all settings and in a broad range of contexts - whole group, small group, individual.
- Developmental, Developing, Making Progress. Instruction exemplifies educators who are able to use the required instructional strategy/practice but with a limited range. Use of strategy is inconsistent or needs coaching to move to expected/proficient/fidelity. Words used to describe strategy may include "some of the time," "somewhat inconsistent". This column helps define or identify the focus of additional coaching.
- Needs Improvement, Not in Place. Includes teachers who are not yet able to use the required instructional strategy/practices. Words used to describe this behavior may include "rarely" or "none of the time." Performance in this area may indicate a need to address the Implementation Drivers (e.g., may need more training, coaching, usable data systems, etc.). May include or describe unacceptable practices - i.e. round robin reading (the opposite of what is described in the Expected category)

b) Purpose and scope of the Practice Profile

Delineate a clear understanding of the desired outcome and purpose for using the instructional strategy or practice

- Use of an evidence-based instructional strategy/practice to guide the development of the Practice Profile
- Clearly define the target student group (age, grade, student group – e.g., English Language Learners)
- Narrow the scope of the profile to behaviorally based, measurable, observable indicators for each core component

c) Clear description of the instructional strategy/practice

- Develop clear descriptions and operational definitions of educator behavior for each of the core components/critical features that are in the instructional strategy
- Ensure the strategy is teachable, learnable, doable, and assessable/observable

d) Fidelity – ensuring consistent use of the Practice Profile across the system or school where it is being used

- How are the essential functions and core activities that are specified in the Practice Profile identified during observation?
- How can the observations and assessments of educator behavior be practical and replicated?
- Who will keep track of the instructional strategy/practice observation data using the Practice Profile?
- How will the instructional strategy/practice data be reported to the leadership team and used?

Area 5) Process

What is the recommended process for developing a Practice Profile?

a) Form a task group

- Approximately six members from different aspects of the system - should include experts in the identified area of need – should include member(s) of the leadership and the Implementation Team
- A leader should be identified to serve as point person and organizer of the group
- Ensure strong facilitation and leadership from the beginning of the process
- Enroll "resistant" staff as key players in the process (all voices)

b) Develop consensus on key instructional need/s before beginning profile work

- Identify and gain a thorough understanding of the root cause of instructional need
- Identify intended purpose for the Practice Profile
- Determine who will use the Practice Profile (e.g. grade level, subject area, teachers, principal, instructional coaches, parents)

c) Provide examples, intensive consultation, and protocols

- Create a clear understanding of the purpose from the beginning
- Research the instructional strategy/practice to be used
- Consult with specialists/experts
- Develop consensus on acceptable practice

d) Develop initial draft

- Articulate levels of proficiency for each column - i.e. Expected/Proficient; Developmental, Needs Improvement/Not in Place
- Determine essential functions or core components; "What is essential for the instructional strategy to be effective?"
- Determine indicators of each essential function or core component (indicators are the description of proficiency, developing, and unacceptable for each core component)

- Break into smaller groups/pairs to draft the practice profile for each essential function

e) Sharing between linked teams

- To build consensus for use at all levels, drafts will be shared between task group, leadership team, and other groups such as improvement leaders, for revision and clarification
- Key questions to ask for group work:
  1. Are the indicators critical for the instructional strategy to be effective?
  2. If yes, is this indicator measurable? If not, can we state it so that it is measurable?
  3. Do the indicators make sense?
  4. Does the indicator belong somewhere else? Should it be removed?
  5. Is anything missing?
  6. Is there anything that does not belong?
  7. Are there particular indicators that need further clarification?

f) Allow enough time

- Each essential function profile takes approximately 2 days to complete. This includes 6-8 hours for initial draft, 2 hour group work, 5-6 hours of revisions. This does not include time to put together initial resources for team members.

Area 6) Alignment
How is alignment ensured? Addressing this area helps to ensure that philosophical principles are reflected in the Practice Profile, linking State Academic Standards, benchmarks, and classroom instruction and practices.

- Teachers may request coaching on use of instructional strategy/practice
- Observer/s may refer teacher to an instructional coach
- Peer coaching may be initiated by teacher(s)

References

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health, 38*(1), 4.

Akin, B. A., Bryson, S. A., Testa, M. F., Blase, K. A., McDonald, T., & Melz, H. (2013). Usability testing, initial implementation, and formative evaluation of an evidence-based intervention: Lessons from a demonstration project to reduce long-term foster care. *Evaluation and Program Planning, 41*(0), 19-30. doi:http://dx.doi.org/10.1016/j.evalprogplan.2013.06.003

Atkins, L., Francis, J., Islam, R., O'Connor, D., Patey, A., Ivers, N., . . . Michie, S. (2017). A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implementation Science, 12*(1), 77. doi:10.1186/s13012-017-0605-9

Blase, K. A., & Fixsen, D. L. (2013). *Core intervention components: Identifying and operationalizing what makes programs work*. Retrieved from Washington, DC: http://aspe.hhs.gov/hsp/13/KeyIssuesforChildrenYouth/CoreIntervention/rb_CoreIntervention.cfm

Blase, K. A., Metz, A., Bartley, L., & Fixsen, D. L. (2011). *Developing practice profiles*. Retrieved from University of North Carolina at Chapel Hill: National Implementation Research Network: https://www.activeimplementation.org/resources/developing-practice-profiles/

Braukmann, P. D., Kirigin Ramp, K. A., Braukmann, C. J., Willner, A. G., & Wolf, M. M. (1983). The analysis and training of rationales for child care workers. *Children and Youth Services Review, 5*(2), 177-194.

Chinman, M., Imm, P., & Wandersman, A. (2004). *Getting To Outcomes: Promoting accountability through methods and tools for planning, implementation, and evaluation*. Retrieved from Santa Monica, CA: RAND Corporation:

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*(50). doi:10.1186/1748-5908-4-50

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.

Elwyn, G., Taubert, M., & Kowalczuk, J. (2007). Sticky knowledge: A possible model for investigating implementation in healthcare contexts. *Implementation Science, 2*(44). doi:10.1186/1748-5908-2-44

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clin Child Fam Psychol Rev., 11*(3), 75-113. doi:10.1007/s10567-008-0036-x

Feldstein, A. C., & Glasgow, R. E. (2008). A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. *Joint Commission Journal on Quality and Patient Safety, 34*(4), 228-243.

Fixsen, D. L., Blase, K. A., Duda, M., Naoom, S., & Van Dyke, M. (2010). Implementation of evidence-based treatments for children and adolescents: Research findings and their implications for the future. In J. Weisz & A. Kazdin (Eds.), *Implementation and*

dissemination: Extending treatments to new populations and new settings (2nd ed., pp. 435-450). New York: Guilford Press.

Fixsen, D. L., Blase, K. A., Timbers, G. D., & Wolf, M. M. (2007). In search of program implementation: 792 replications of the Teaching-Family Model. *The Behavior Analyst Today, 8*(1), 96-110. doi:10.1037/h0100104

Fixsen, D. L., Hassmiller Lich, K., & Schultes, M. T. (2018). Shifting systems of care to support school-based services. In A. Leschied, Saklofske, D., Flett, G. (Ed.), *Handbook of school-based mental health promotion: An evidence informed framework for implementation* (pp. 51-63). Toronto: Springer: Toronto: Springer.

Forgatch, M. S., & DeGarmo, D. S. (2011). Sustaining Fidelity Following the Nationwide PMTO Implementation in Norway *Prevention Science, 12*(3). doi:10.1007/s11121-011-0225-6

Fraser, M. W., & Galinsky, M. J. (2010). Steps in Intervention Research: Designing and Developing Social Programs. *Research on Social Work Practice, 20*(5), 459-466. doi:10.1177/1049731509358424

Graham, I., Logan, J., Harrison, M., Straus, S., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in Translation: Time for a Map? *The Journal of Continuing Education in the Health Professions, 26*(1), 13 - 24.

Greenhalgh, T., Robert, G., MacFarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly, 82*(4), 581-629.

Hall, G., & Hord, S. M. (2011). *Implementing change: Patterns, principles and potholes* (4th ed.). Boston: Allyn and Bacon.

Hord, S. M., Rutherford, W. L., Huling-Austin, L., & Hall, G. E. (1987). *Taking charge of change*. Alexandria, VA: Association for Supervision and Curriculum Development.

Institute for Healthcare Improvement. (2010). The sound of two hands washing: Improving hand hygiene. Retrieved from http://www.ihi.org/IHI/Topics/CriticalCare/IntensiveCare/ImprovementStories/FSSoundofTwoHandsWashing.htm

Kilbourne, A. M., Neumann, M. S., Pincus, H. A., Bauer, M. S., & Stall, R. (2007). Implementing evidence-based interventions in health care: Application of the replicating effective programs framework. *Implementation Science, 2*(42). doi:10.1186/1748-5908-2-42

Klein, K. J., & Sorra, J. S. (1996). The challenge of innovation implementation. *Academy of Management Review, 21*(4), 1055-1080.

Lewis, G. (2014). Emerging lessons from the FIGO LOGIC initiative on maternal death and near-miss reviews. *International Journal of Gynecology & Obstetrics, 127*(S1), S17-S20. doi:doi:10.1016/j.ijgo.2014.07.007

Marshall, M., de Silva, D., Cruickshank, L., Shand, J., Wei, L., & Anderson, J. (2017). What we know about designing an effective improvement intervention (but too often fail to put into practice). *BMJ Quality and Safety, 26*(7), 578-582. doi:10.1136/bmjqs-2016-006143

McHugo, G. J., Drake, R. E., Teague, G. B., & Xie, H. (1999). Fidelity to assertive community treatment and client outcomes in the New Hampshire Dual Disorders Study. *Psychiatric Services, 50*(6), 818-824.

Michie, S., Fixsen, D., Grimshaw, J., & Eccles, M. (2009). Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implementation Science, 4*(1), 40.

Nielsen, J. (2005). Usability for the masses. *Journal of Usability Studies, 1*(1), 2-3.

Nord, W. R., & Tucker, S. (1987). *Implementing routine and radical innovations*. Lexington, MA: D. C. Heath and Company.

Phillips, E. L., Phillips, E. A., Fixsen, D. L., & Wolf, M. M. (1974). *The Teaching-Family handbook* (Second ed.). Lawrence. KS: University Press of Kansas.

Rogers, E. M. (1995). *Diffusion of Innovations* (4 ed.). New York: The Free Press.

Sandler, I., Ostrom, A., Bitner, M. J., Ayers, T. S., Wolchik, S., & Daniels, V. S. (2005). Developing effective prevention services for the real world: A prevention service development model. *American Journal of Community Psychology, 35*(3-4), 127-142.

Schoenwald, S. K., Brown, T. L., & Henggeler, S. W. (2000). Inside multisystemic therapy: Therapist, supervisory, and program practices. *Journal of Emotional and Behavioral Disorders, 8*(2), 113-127.

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal, 17*(Special Issue), 27-43.

Teague, G. B., Bond, G. R., & Drake, R. E. (1998). Program fidelity in assertive community treatment: Development and use of a measure. *American Journal of Orthopsychiatry, 68*(2), 216-232.

Tilly III, W. D. (2008). *How To Pin the Pendulum to the Wall: Practical Strategies to Create Lasting Change in Schools.* Paper presented at the ABAI Evidence-Based Education Conference, Reston, VA.

Wolf, M. M., Kirigin, K. A., Fixsen, D. L., Blase, K. A., & Braukmann, C. J. (1995). The Teaching-Family Model: A case study in data-based program development and refinement (and dragon wrestling). *Journal of Organizational Behavior Management, 15*, 11-68.