PII-TTAP

PERMANENCY INNOVATIONS INITIATIVE

Training & Technical Assistance Project

GUIDE TO DEVELOPING, IMPLEMENTING, AND ASSESSING AN INNOVATION

Volume 4: Initial Implementation

Children's Bureau

Acknowledgements

To support the Permanency Innovations Initiative (PII) Grantees in better meeting the needs of children and families, the PII Training and Technical Assistance Project (PII-TTAP) team created the Development, Implementation, and Assessment Approach (the Approach). The Approach helps organizations develop new innovations or adapt existing ones and effectively implement them to ultimately improve outcomes for children and families. The PII-TTAP team created this Guide which operationalizes the Approach and aids in the transfer of learning by providing detailed information, tools, and instructions for innovation development and adaptation, data and fidelity monitoring, and sustainability planning within child welfare systems. It is based on lessons learned working with the tools, guidance, and resources provided to PII Grantees. This Guide was created by:

JBS International, Inc.

Tori Russell

Roseana Bess

Mary Jane Peck

Emily Fisher

Jacquelyn Spangler

Deborah Lurie

Hung Pho

Brian Jones

Clay Finck

Center for the Support of Families

Will Hornsby

Elizabeth Black

Cathy Welsh

Eliza Van Byrne

Glenda McMillan

NIRN

Karen Blase

Renée Boothroyd

Will Aldridge, III

Allison Metz

Dean Fixsen

Leah Bartley

Suggested citation:

Permanency Innovations Initiative Training and Technical Assistance Project. (2016). *Guide to developing, implementing, and assessing an innovation*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau.

This document was created under JBS International, Inc. Contract No. HHSP23320095638WC funded by the Children's Bureau, Administration for Children and Families, U.S. Department of Health and Human Services. Subcontractors under this contract include: the Center for the Support of Families and the National Implementation Research Network (NIRN) at the University of North Carolina Frank Porter Graham Child Development Institute[1].

[1] NIRN was actively involved in PII-TTAP from October 2010 through June 2015.

TABLE OF CONTENTS

Initial Implementation 1

Section 8: Initial Implementation and Testing 3

- **Chapter 8.1:** What Is Usability Testing? 4
- **Chapter 8.2:** How to Conduct Usability Testing 7
- Usability Testing Tool 10
- **Section 8:** Test Your Understanding 28

Volume 4: Test Your Understanding Answer Key 30

Appendix 31

Initial Implementation

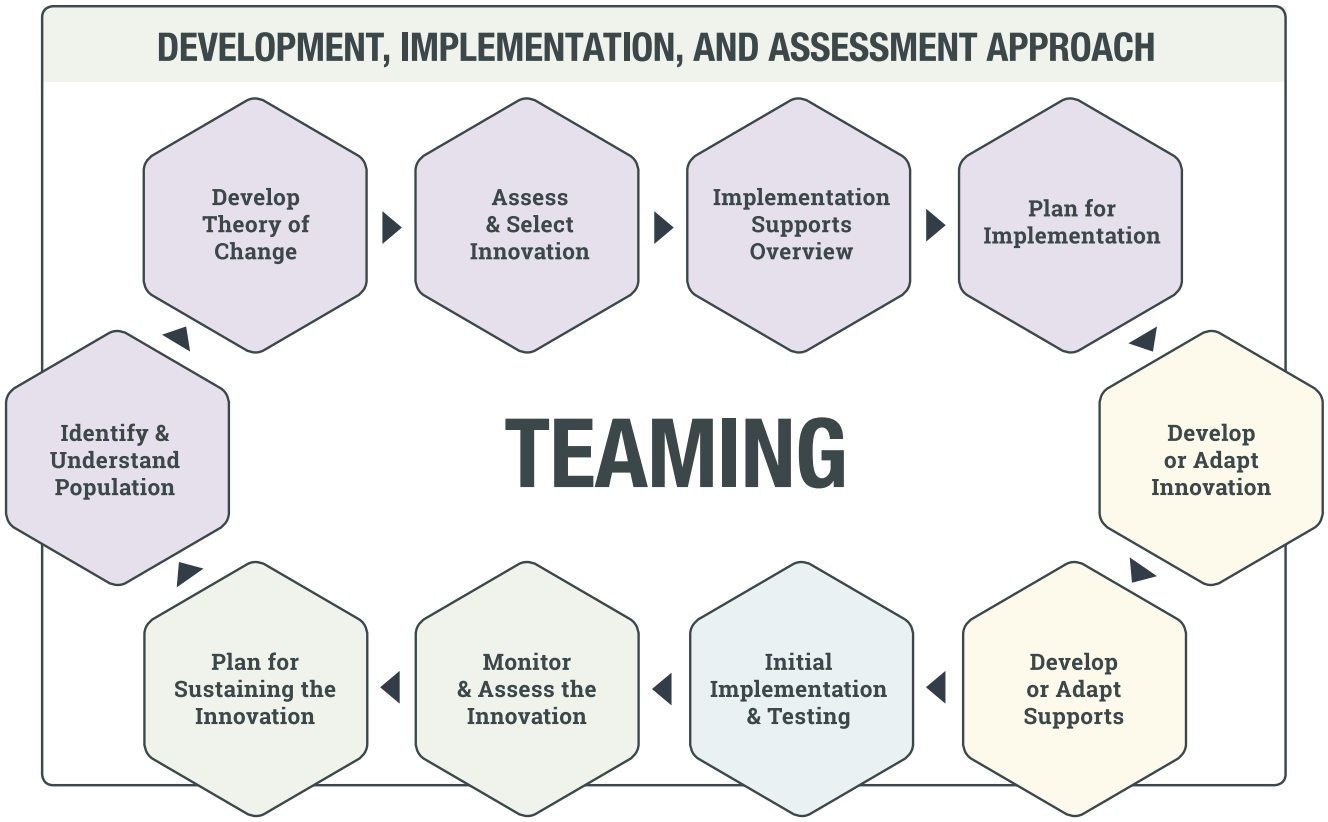
The purpose of the Initial Implementation stage is to: (1) test critical elements, such as key processes and data collection activities, and (2) modify components so that innovation processes are improved and implementation supports are supporting the right processes. Testing is also a valuable strategy to further identify and operationalize the essential functions of an innovation. During initial implementation, children and families begin to receive the intervention, all components of the innovation are at least partially in place, and the implementation supports begin to function.

The *Guide to Developing, Implementing, and Assessing an Innovation* (the Guide) focuses on initial implementation and testing during the Initial Implementation stage.

Initial Implementation & Testing: Through a process called "usability testing," teams test critical elements such as key processes, data collection activities, or essential supports for implementation.

Getting Started With Initial Implementation

The questions included in the Introduction section of the Guide help determine which implementation stage an innovation is in and which



volume of the Guide addresses that stage. If an innovation is currently in the Initial Implementation stage, answering the questions below will help determine which chapter of the Guide is the most appropriate starting place. An answer of “no” to any of the questions most likely indicates the stage that the initiative is in and the section of the Guide associated with that stage. Implementation is not a linear process, however, so before beginning Volume 4, it may be wise to review previous steps included in the Installation stage (Volume 3 of the Guide).

- Have you begun implementing your innovation? (Chapter 8.1)
- Is your team using rapid-cycle problem solving to test critical elements and make rapid adjustments to your innovation? (Chapter 8.1)
- Have you developed a plan for the usability testing process? (Chapter 8.2)

SECTION 8

Initial Implementation and Testing

This section consists of two chapters that help to identify the parts of an innovation that may need to be tested during initial implementation of an innovation. Also known as “usability testing,” this process is used to improve the innovation, implementation supports, and data collection processes. Chapter 8.1 provides a thorough description of the usability testing process. Chapter 8.2 explains how to conduct usability testing and provides scenarios for developing and initiating a usability testing plan. This section includes a knowledge check to test understanding of the materials and a Usability Testing Tool to support the development of a usability testing plan.

What is Usability Testing?

Overview

After an implementation team has developed a theory of change, developed or adapted an innovation, and learned about the supports necessary for implementation, it can proceed with initial implementation and testing. This chapter describes the process of establishing the innovation within the organization and learning whether procedures, processes, or innovation components need to be adapted for implementation to move forward. This process is often referred to as "usability testing."

Learning Objectives: This chapter

- Improves understanding of the purpose of usability testing
- Helps determine how to decide what to test during usability testing

Competencies: Meeting the learning objectives will develop foundational knowledge to conduct usability testing.

Key Terms

Plan, Do, Study, Act (PDSA)[1] - A learning and improvement cycle that combines management thinking with statistical analysis in the following four stages:

- Plan: identify what can be improved and what change is needed
- Do: implement the design change
- Study: measure and analyze the process or outcome
- Act: take action if the results are not as hoped for

This cycle is used to make changes that lead to improvement in a manner of continuous quality improvement. This is a never ending process.

Usability testing - A process of trying out the critical components of the implementation, the innovation, and data collection to see how they fit within the organization. Adjustments can be made to improve the innovation and implementation supports and the fit.

Purpose of Usability Testing

Implementing a new innovation in an organization is often accompanied by uncertainty. Components of an innovation may represent a new approach, and some may be more difficult than others to implement. Aspects of the innovation or data collection may be challenging to staff. For example, a fidelity assessment may be a new concept to the organization and therefore take more time than expected to conduct. This knowledge allows the organization to make refinements to the process.

The purpose of usability testing is to help further operationalize the essential functions of the innovation, implementation supports (e.g., training, coaching, recruitment, selection, and fidelity assessment), and data collection. Testing only a few components a few times allows for efficient feedback to make the necessary adjustments.

What Can Be Learned

During usability testing, teams may learn that:

- Implementation supports should be adjusted to improve practitioners' abilities to implement the innovation as intended.

[1]Best, M., & Neuhauser, D. (2006). Walter A. Shewhart, 1924, and the Hawthorne factory. *Quality and Safety in Health Care, 15(2)*, 142-143.

- The essential functions of the innovation must be more clearly defined.
- Data collection processes need more time and/or uses of collected data need more clarification.

Timing and Functions

Usability testing begins during initial implementation and can be used to test critical components throughout the life of an innovation. Some of its functions are to:

- Maximize learning from available examples (i.e., testing results)
- Quickly detect challenges related to key implementation and innovation processes
- Revise and re-test the processes to see if they have improved
- Stabilize essential functions of an innovation, implementation supports, and data collection processes

Usability testing may be most beneficial in situations that involve:

- New processes that are perceived to be challenging
- Transitions for children or families
- Shared responsibility for an activity

Plan, Do, Study, Act

Usability testing is a type of rapid-cycle problem solving known as Plan, Do, Study, Act (PDSA). PDSA is the quick and iterative process of:

- Planning or identifying the component to be tested (Plan)
- Putting that specific component into place (Do)
- Gathering feedback about the process and examining the results (Study)
- Deciding whether and where to make improvements (Act)

Detect Strengths and Gaps Through Multiple Testing Cycles

Usability testing applies PDSA to detect strengths and gaps via quick tests. Testing teams should "test" the way the innovation is going to be implemented across different implementation locations, if applicable. This approach allows improvements to be made quickly from one cycle to the next.

For example, if 50 practitioners are available to test training, five cohorts of ten practitioners could be used for the usability test instead of one large group. This approach allows the testing team to learn more and obtain better training outcomes. It also enables the team to make improvements as the testing moves from one cohort to the next, until all cohorts have been tested.

Testing cycles should include a sufficient number of workers from a variety of locations and/or offices to help differentiate individual, regional, county, and systemic issues. Each cycle within a usability test must include a plan to deliver the components of the innovation or implementation supports that are being tested and a means for assessing findings and then revising the innovation as needed.

Conducting more than one round of usability testing allows a greater number of problems to be revealed

and addressed. As the team detects challenges, they make revisions then test the "improved" processes again. Usability testing cycles continue, as needed, to move beyond identifying surface issues and detect deeper challenges related to operationalizing the innovation. The goal is to solve both simple and challenging problems before further implementation.

Example Questions to Ask

Team A is in charge of implementing an innovation. Carried out by front line practitioners, the innovation's goal is to maintain children in their homes. The innovation represents a new way for practitioners to interact with children and families. The innovation requires the practitioner to administer an assessment to the child in an initial meeting. It also requires practitioners to collect more data than they have in the past. Coaching and fidelity assessments have never been used in the organization and will be implemented to support this innovation.

Before implementing the innovation agency wide, Team A conducts usability testing with a small group of practitioners to test the new processes and identify components that need modification prior to full-scale implementation. While it is impossible to outline all the critical components for usability testing, sample questions to help identify priority areas include:

- Will families engage?
- Will individual children engage?
- Are practitioners using the essential functions of the innovation when interacting with children and families?
- Are case transfers occurring as intended?

Given that both coaching and fidelity assessments are new concepts for the organization, sample questions to address implementation supports include:

- Is coaching occurring as intended?
- Have practical and reliable fidelity assessments been developed and used?
- Can fidelity assessments be used as intended?

Sample questions related to new data-collection responsibilities include:

- Will staff complete the youth assessments in a timely manner?
- Does data entry and reporting occur as intended?
- Is the expected amount of data collection feasible, given the other demands placed on practitioners?

The examples in this section highlight different types of questions related to innovation, implementation supports, and data collection. The actual questions for the testing will vary based on the innovation being implemented.

What Usability Testing Is Not

When planning an approach, teams should keep in mind that usability testing:

- Is not an evaluation
- Does not present a set of benchmarks
- Does not lead to the final word on any component of the practice, implementation supports, or data collection processes

Usability testing IS the process of trying out the critical components of the innovation, implementation supports, and data collection processes to see how they fit within an organization.

Connection to Evaluation

Usability testing helps pave the way for evaluation efforts by ensuring that the innovation's key components and implementation supports are working. Usability testing identifies and improves upon data collection protocols essential for tracking implementation supports and evaluation outcomes. Testing these components early helps identify and resolve issues, which promotes an easier transition to evaluation. The innovation must be stable prior to evaluation; that is, the innovation's core components should not change.

How to Conduct Usability Testing

Overview

The previous chapter (Chapter 8.1: What is Usability Testing?) helps develop understanding of what usability testing is and offers a list of possible questions related to an innovation, implementation supports, and data collection. This chapter outlines a plan for usability testing. It provides additional guidance on key steps that should be taken when creating a usability testing plan and a Usability Testing Tool to help outline a plan.

Learning Objectives: This chapter explains how to

- Develop a usability testing plan
- Determine when to engage in usability testing

Competencies: Meeting the learning objectives will develop foundational knowledge to conduct usability testing.

Multiple Quick Cycles of Testing

Usability testing is the process of engaging in multiple, quick cycles of testing, learning, and improvement. Each cycle within a usability test must include:

- Clearly defined essential function(s) of the innovation, implementation process, or data collection process that is being tested
- A method for assessing the findings
- A procedure for revising the innovation, implementation supports, or data collection processes, as needed

End of Cycle Decisions

When a cycle of usability testing is completed, the implementation team reviews the results and decides whether additional usability tests are needed. At this point, the team decides whether to conduct more usability testing on that component or test a different component, or determine whether the innovation is ready for full scale implementation.

Create a Usability Testing Plan

Before beginning usability testing, it is helpful to have a plan in place. There are nine key steps for creating a plan for the usability testing process.

Step 1. Clarify who is responsible for coordinating the test (e.g., implementation team).

Step 2. Clearly outline the scope of the test, such as testing the training curriculum or fidelity assessment process.

Step 3. Outline a process for identifying cases or practitioners that will be included in the test. For example, the test will include only practitioners who have been completely trained in the innovation and have received one or more new cases after completing training.

Step 4. Clarify who is responsible for reporting the results to the coordinator of the test—are they practitioners, supervisors, and/or coaches?

Step 5. Identify the type of metrics or key outputs that can be discovered quickly. These can be qualitative, such as the opinion of practitioners and family members about the new engagement approach. Or they can be quantitative, such as the number of parent-worker contacts that occurred.

Step 6. Set a clear timeline for data reporting of these metrics. For example, they will be reported to the implementation team every week or every other week on Friday for 6 weeks.

Step 7. Identify who is responsible for collecting the data, where the data will be stored, and who will be responsible for pulling together the reports. For example, one member of the implementation team could be responsible for entering the data they received into a spreadsheet. They can then use this data to create reports that summarize the results of the usability testing metrics.

Step 8. Identify the decision-makers responsible for acting on the results of the testing, for example, the implementation team in conjunction with agency leadership.

Step 9. Identify the criteria for revision of the process or for declaring it "good enough" when the test is completed. For example, "If a certain percentage of the parents meet twice with a practitioner, we are satisfied with the engagement process."

The Appendix includes an example of a usability testing plan from a child welfare agency.

Example: Planning Overview (Diamond County)

Diamond County has already developed a practice profile and a fidelity assessment for their innovation. They are ready to begin implementation. Before moving forward, they would like to test some key implementation components, such as whether the practitioners will voluntarily participate in the new coaching process and whether families will engage in the innovation in the allotted amount of time.

Initiating the Plan

Diamond County leadership has decided that the implementation team will be responsible for these tests. The team decides that every week the supervisors of the first three units that completed training will report to the implementation team the number of workers who signed up for coaching. In addition, every 2 weeks, they will report to the implementation team the number of families that have agreed to participate in the new innovation. The usability testing period will last for 6 weeks.

The implementation team designates a data analyst to be responsible for the data collection process. The data analyst will enter the data on coaching and family participation into a spreadsheet. At the end of the 6 weeks, the data analyst will run the reports regarding participation and send them to both the implementation team and the agency leadership.

Together, the implementation team and agency leadership decide that, if in the first 6 weeks after training,

65% of their workers sign up for coaching, they will be satisfied with the current communication about the purpose of coaching. In addition, if at least 50% of the families engage in the new innovation, they will be satisfied with the engagement process as currently outlined in the practice profile.

Note: The desired result is not perfection. The desired result is a determination that there are no systemic problems with the critical components, and that these critical components are working well enough.

Ready to Test?

To undergo usability testing, an innovation must be defined. That is, essential function(s) of the innovation, implementation supports, or data collection process must be clearly articulated. In addition, testing cannot occur until the staff responsible for implementing the innovation, the implementation supports, or data collection processes have been trained or prepared to fulfill these roles. For example, the initial practitioners who participate in usability testing should have been selected and trained to deliver the innovation as intended. Similarly, if data collection protocols or implementation supports are to be tested, the protocols for collecting data and providing implementation supports, such as the selection of staff and the training or coaching process, should be clearly outlined before testing. As with usability tests for the innovation, individuals charged with data collection or with providing implementation supports must be prepared for those roles and to carry out the protocols. The testing may help to further operationalize the training, the data collection processes, or other implementation supports.

Select Components for Testing

When the staff and innovation are reading for testing, the implementation team can select a limited number of critical components for usability testing. The previous chapter provided ideas for brainstorming these components. These critical components should be relevant to the theory of change, challenging, and critical to the success of the innovation. However, as noted in Chapter 1: What Is Usability Testing? usability testing should be conducted throughout the life of the innovation.

A Reminder

Conducting more than one round of usability testing allows a greater number of problems to be revealed and addressed. As the team detects challenges, it makes revisions then tests the “improved” processes again. Usability testing cycles continue, as needed, to move beyond identifying surface issues to detecting deeper challenges related to operationalizing the innovation. The goal is to solve both simple and challenging problems before further implementation.

The Appendix includes an example of a usability testing report, which can be a useful method for documenting findings and next steps.

Usability Testing Tool

The Usability Testing Tool on the following pages helps outline a plan for usability testing, including measures, persons responsible, and timelines for completion. The tool can also be useful for tracking test results and assessing whether another round of testing is needed.

USABILITY TESTING TOOL

FROM THE *GUIDE TO DEVELOPING, IMPLEMENTING, AND ASSESSING AN INNOVATION, VOLUME 4*

Instructions

The tool below will help you outline a plan for usability testing. It is divided into two parts. The first part allows you to outline the various measures to test during the usability testing period, who should be responsible, and the timeline. The second part is designed to be used after you have conducted the various tests. It will help you keep track of the results of each of the tests and whether another round of testing will be needed. If another round of testing will be needed, you can describe the next course of action and who will be responsible.

This tool may be used in two ways:

1. ***Print the following pages and use them as a discussion guide with your team. Write your answers in the space provided.***
2. ***Type your information into the space provided and save to your computer. This will allow you to print the completed document or e-mail it to your team members.***

PART I-USABILITY TESTING MEASURES

Use the following boxes to outline the measures you want to test related to the innovation. Remember that testing only a few components a couple of times will allow for quick and efficient feedback and adjustments.

INNOVATION MEASURES

Measures in this category pertain to items that help the team to further refine and operationalize the essential function of the innovation.

Step 1. Who (e.g., the implementation team) is responsible for coordinating the test?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

INNOVATION MEASURES

Steps 2 & 3. Clearly outline the scope of the test. This includes identifying what is being tested (e.g., the training curriculum or fidelity assessment process) and who will be included (e.g., practitioners who have been completely trained in the innovation and have received one or more new cases after completing training). **What is being tested?**
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Who will be included?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

INNOVATION MEASURES

Step 4. Who is responsible for reporting the data gathered to the coordinator of the test? The coordinator was identified in Step 1. (The practitioners, supervisors, etc. report data to the implementation team.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 5. What type of metrics or key outputs can be tested quickly? (These can be qualitative, such as the opinion of practitioners and family members about the new engagement approach. Or they can be quantitative, such as the number of parent-worker contacts that occurred.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

INNOVATION MEASURES

Step 6. What is the timeline for reporting data about this measure? (For example, they will be reported to the implementation team every week or every other week on Friday for 6 weeks, or they will be reported once to the implementation team at the end of the 8-week testing period.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 7. Identify who is responsible for collecting the data, where it will be stored, and who will be responsible for pulling together the reports. (For example, one member of the implementation team could be responsible for entering the data they received into a spreadsheet, which is used to create reports that summarize the results of the usability testing metrics.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

INNOVATION MEASURES

Step 8. Who determines if the results are good enough (e.g., the implementation team in conjunction with agency leadership)?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 9. What are the criteria by which the results will be assessed (e.g., 75% of parent-worker contacts are completed)? (Identify the criteria for revising the process of declaring it "good enough," when the test is completed. For example, if a certain percentage of the parents meet twice with a practitioner, we are satisfied with the engagement process.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

IMPLEMENTATION SUPPORTS MEASURES

Measures in this category help the team to further refine and operationalize their implementation supports (e.g., training, coaching, recruitment, selection, and fidelity assessment).

Step 1. Who (e.g., the implementation team) is responsible for coordinating the test?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

IMPLEMENTATION SUPPORTS MEASURES

Steps 2 & 3. Clearly outline the scope of the test. This includes identifying what is being tested (e.g., the training curriculum or fidelity assessment process) and who will be included (e.g., practitioners who have been completely trained in the innovation and have received one or more new cases after completing training). **What is being tested?**
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Who will be included?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

IMPLEMENTATION SUPPORTS MEASURES

Step 4. Who is responsible for reporting the data gathered to the coordinator of the test? The coordinator was identified in Step 1. (The practitioners, supervisors, etc. report data to the implementation team.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 5. What type of metrics or key outputs can be tested quickly? (These can be qualitative, such as the opinion of practitioners and family members about the new engagement approach. Or they can be quantitative, such as the number of parent-worker contacts that occurred.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

IMPLEMENTATION SUPPORTS MEASURES

Step 6. What is the timeline for reporting data about this measure? (For example, they will be reported to the implementation team every week or every other week on Friday for 6 weeks, or they will be reported once to the implementation team at the end of the 8-week testing period.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 7. Identify who is responsible for collecting the data, where it will be stored, and who will be responsible for pulling together the reports. (For example, one member of the implementation team could be responsible for entering the data they received into a spreadsheet, which is used to create reports that summarize the results of the usability testing metrics.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

IMPLEMENTATION SUPPORTS MEASURES

Step 8. Who determines if the results are good enough (e.g., the implementation team in conjunction with agency leadership)?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 9. What are the criteria by which the results will be assessed (e.g., 75% of parent-worker contacts are completed)? (Identify the criteria for revising the process of declaring it "good enough," when the test is completed. For example, if a certain percentage of the parents meet twice with a practitioner, we are satisfied with the engagement process.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

DATA COLLECTION MEASURES

Measures in this category help the team to further refine and operationalize their data collection procedures.

Step 1. Who (e.g., the implementation team) is responsible for coordinating the test?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

DATA COLLECTION MEASURES

Steps 2 & 3. Clearly outline the scope of the test. This includes identifying what is being tested (e.g., the training curriculum or fidelity assessment process) and who will be included (e.g., practitioners who have been completely trained in the innovation and have received one or more new cases after completing training). **What is being tested?**
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Who will be included?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

DATA COLLECTION MEASURES

Step 4. Who is responsible for reporting the data gathered to the coordinator of the test? The coordinator was identified in Step 1. (The practitioners, supervisors, etc. report data to the implementation team.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 5. What type of metrics or key outputs can be tested quickly? (These can be qualitative, such as the opinion of practitioners and family members about the new engagement approach. Or they can be quantitative, such as the number of parent-worker contacts that occurred.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

DATA COLLECTION MEASURES

Step 6. What is the timeline for reporting data about this measure? (For example, they will be reported to the implementation team every week or every other week on Friday for 6 weeks, or they will be reported once to the implementation team at the end of the 8-week testing period.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 7. Identify who is responsible for collecting the data, where it will be stored, and who will be responsible for pulling together the reports. (For example, one member of the implementation team could be responsible for entering the data they received into a spreadsheet, which is used to create reports that summarize the results of the usability testing metrics.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

DATA COLLECTION MEASURES

Step 8. Who determines if the results are good enough (e.g., the implementation team in conjunction with agency leadership)?
Measure #1:
Measure #2:
Measure #3:
Measure #4:

Step 9. What are the criteria by which the results will be assessed (e.g., 75% of parent-worker contacts are completed)? (Identify the criteria for revising the process of declaring it "good enough," when the test is completed. For example, if a certain percentage of the parents meet twice with a practitioner, we are satisfied with the engagement process.)
Measure #1:
Measure #2:
Measure #3:
Measure #4:

PART II-MEASURE RESULTS

The boxes below are meant to be completed after the usability testing outlined above has been carried out. These boxes are designed to help you track your results and decide whether an additional round of usability testing will be necessary. Remember that conducting more than one round of usability testing allows a greater number of problems to be revealed and addressed.

INNOVATION MEASURE RESULTS

What were the results? (See criteria set in Step 9 and who determines whether the results are good enough in Step 8).		**Will another round of testing be needed?**	**What action will be taken in light of these results?**	**Who will be responsible for taking this action?**
Measure #1:		**Yes** ☐ **No** ☐		
Measure #2:		**Yes** ☐ **No** ☐		
Measure #3:		**Yes** ☐ **No** ☐		
Measure #4:		**Yes** ☐ **No** ☐		

IMPLEMENTATION SUPPORTS MEASURE RESULTS

What were the results? (See criteria set in Step 9 and who determines whether the results are good enough in Step 8)		**Will another round of testing be needed?**	**What action will be taken in light of these results?**	**Who will be responsible for taking this action?**
Measure #1:		**Yes** ☐ **No** ☐		
Measure #2:		**Yes** ☐ **No** ☐		
Measure #3:		**Yes** ☐ **No** ☐		
Measure #4:		**Yes** ☐ **No** ☐		

DATA COLLECTION MEASURE RESULTS

What were the results? (See criteria set in Step 9 and who determines whether the results are good enough in Step 8)		**Will another round of testing be needed?**	**What action will be taken in light of these results?**	**Who will be responsible for taking this action?**
Measure #1:		**Yes** ☐ **No** ☐		
Measure #2:		**Yes** ☐ **No** ☐		
Measure #3:		**Yes** ☐ **No** ☐		
Measure #4:		**Yes** ☐ **No** ☐		

Test Your Understanding

The following questions will help test understanding of the concepts in Section 8. An answer key is provided at the end of this volume.

1. Usability testing serves to further operationalize:

- ☐ a. The innovation
- ☐ b. Implementation supports
- ☐ c. Data collection processes
- ☐ d. All of the above

2. Usability testing is only used during initial implementation:

- ☐ a. False
- ☐ b. True

A child welfare organization is in the process of implementing a home-based parenting program. Before implementing organization wide, they would like to engage in usability testing with a small group of practitioners. Based on this information, please answer questions 3 and 4 below.

3. Which of the following are examples of the usability testing they may engage in? (Choose all that apply)

- ☐ a. Testing whether there needs to be one engagement session or two before beginning the program
- ☐ b. Testing whether the data collection protocol can be done when therapists are at the client homes
- ☐ c. Testing whether the sustainability plan was well received by the community stakeholders
- ☐ d. Testing whether better fidelity scores lead to better client outcomes

4. As a result of usability testing, implementation teams and the larger organization might be able to learn whether (choose all that apply):

- ☐ a. Practitioners need more training to effectively engage families.
- ☐ b. Parents understand why there is a need for video recording to assess for fidelity.
- ☐ c. The innovation proves to be very effective with children aged 4-7 but not with children aged 14-17.
- ☐ d. Child outcomes improve faster when services are provided in their home.

5. The point of engaging in PDSA is to:

- ☐ a. Detect strengths and gaps via quick tests and make improvements where needed
- ☐ b. Completely revise the practice profile
- ☐ c. Take the place of a formal evaluation
- ☐ d. Test the theory of change

6. Usability testing works best when:

- ☐ a. There is time for only one round of testing.
- ☐ b. There is no outlined plan before beginning the testing.
- ☐ c. There is time for more than one round of testing.
- ☐ d. The same single practitioner is used from one cycle to the next.

7. Why is it important to clearly define who (e.g., practitioners and target population) will be included in each of the usability tests?

- ☐ a. So everyone can have a role and feel included
- ☐ b. Because it will be useful for the formal evaluation
- ☐ c. Because these people will not have to participate in coaching
- ☐ d. To ensure that the test is based on a real-world performance of the component being tested

8. The final step in creating a usability testing plan is identifying the criteria for determining whether the results are "good enough." This is important because:

- ☐ a. It will be used to help predict whether the innovation will address the desired outcomes.
- ☐ b. The goal is not perfection; the goal is a determination that the innovation is ready for further implementation.
- ☐ c. After the results are deemed good enough, that particular element will never need to be tested again.
- ☐ d. After determining the results are good enough, they must be further analyzed.

9. Before beginning usability testing there must be:

- ☐ a. A clearly defined innovation
- ☐ b. Clients willing to participate in the testing
- ☐ c. Organization leadership willing to carry out each test
- ☐ d. Only one location for testing

10. When engaging in usability testing, it is most helpful to:

- ☐ a. Try to test as much as possible.
- ☐ b. Only test one of the implementation supports.
- ☐ c. Focus solely on practitioner and client interactions.
- ☐ d. Define a limited number of components to test.

11. To move forward with implementation after usability testing, the results should indicate:

- ☐ a. Everything is working perfectly and there is nothing left to test.
- ☐ b. The innovation needs more clearly defined essential functions.
- ☐ c. There are multiple systemic barriers.
- ☐ d. Critical components are working well enough to proceed.

VOLUME 4: TEST YOUR UNDERSTANDING ANSWER KEY

Section 8

1. d

2. a

3. a,b

4. a,b

5. a

6. c

7. d

8. b

9. a

10. d

11. d

Volume 4: Appendix

Arizona Department of Economic Security
Division of Children, Youth and Families

Fostering Readiness & Permanency Project

Usability Testing Plan
2012

This document presents the Usability Testing Plan for the FRP Project. The document identifies the components of the FRP Project that are most critical to ensuring that the interventions can be implemented as intended. The plan and selection of components was completed by members of the Usability Testing Sub-committee of the FRP Implementation Team by reviewing and discussing each of the practice profiles to identify critical components for testing, and eventually determining the seven components identified in the plan.

Following identification of the five components, members then began developing the criteria for capturing and measuring each component. Criteria for the successful implementation of each component are identified, as is the process for obtaining the data. Details on how the data will be used to determine the need for changes to the interventions and who will use the data to make those decisions is also included in this Usability Testing plan. Finally, each component includes a section requiring the assessment of barriers and identification of ways to address the barriers.

Usability Testing will begin on August 1, 2012 and will include two groups of randomly selected young people in the treatment group and their assigned CPS Specialists, Youth Advocates and CARE Coordinators. Group 1 will start on August 1, 2012 and include 2 CARE Coordinators, 4 Youth Advocates, 12 Young Persons and (at most) 12 CPS Specialists (and their respective CPS Unit Supervisors). If the Usability Testing indicates that changes are needed, the approximate starting date for Group 2 would be on September 1, 2012, and would include 1 CARE Coordinator, 2 Youth Advocates, 6 Young Persons and (at most) 6 CPS Specialists (and their respective CPS Unit Supervisors).

Unless specified otherwise, Usability Testing data will be collected by FRP Project Quality Assurance (QA) teams. The QA Teams are comprised of FRP Project staff members from DES' Division of Children, Youth and Families and consultants from LeCroy & Milligan Associates. Quality assurance team members will be conducted by QA Team members working together in pairs so that each QA Team will be able to address model fidelity and usability testing measures during the following data collection. A team approach increases inter-rater reliability and reduces investigator bias and subjectivity because it incorporates the views of more than one individual. A team approach also helps to ensure that QA Team members have a consistent approach in collecting data and providing feedback to DES and the FRP Project contract provider, Arizona's Children Association (AzCA).

During the Usability Testing phase, QA reviews will be conducted with many measures being reviewed weekly in order to provide rapid feedback to DES FRP staff, teams and subcommittees, and AzCA. A copy of the Usability Testing timeline is provided at the end of this document.

Successful completion of Usability Testing is anticipated to occur by October 31, 2012 and will lead into the next stage of the project, formative evaluation. It is anticipated that the use of this Usability Testing plan will provide the guidance and direction for the FRP Project to successfully begin and complete its Usability Testing process.

Usability Test 1 – Assess the viability of the collaborative aspect of the CARE Team

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria for usability
1	Score on a scale of items assessing the perceived effectiveness of the CARE Team's ability to work together collaboratively to develop the work plan-- e.g., not at all effective (1), rarely effective (2), somewhat effective (3), usually effective (4), always effective (5)	Structured interviews with CPS Specialists, Youth Advocates, and Care Coordinators 2nd week Sep and last week Oct	n/a	QA Team	The required average score across items will be **3.5** for each interviewee.
2	Score on the CARE Team meeting quality of interactions checklist, sections 2 and 3 (A minimum of two meetings will be observed.)	QA Instrument #14 *CARE Team Meeting Observation Rating.* Observations of CARE Team meetings with the CPS Specialists, Youth Advocates, and Care Coordinators	QA Plan #21, 38, 39 and 41	QA Team	The requirement will be that **75** percent of the relevant items checked by the time of the last meeting observed will receive a rating of (3) indicating expected level of implementation.
3	Number of CARE team meetings held in accordance with the young person's individualized work plan.	QA Instrument #8 *Case Record Review.* Reviews of case record and CARE Team meeting minutes and cross-walk with each young person's work plan.	QA Plan #29, 30 and 31	QA Team	The requirement will be that the number of meetings held by the end of the usability testing phase was consistent with the young person's work plan **90** percent of the time. The requirement will be that the number of meetings that all participants attended by the end of the usability testing phase is **90** percent.

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria for usability
4	Assessment of barriers and facilitators to effective collaboration in development of the work plan.	Interviews with Youth Advocate, Care Coordinator, and CPS Specialist. Information gathered from case record reviews on CARE Team attendance		QA Team	No Criteria: Information will be used to inform improvement strategies.

Review of data

1. The QA Team will gather the data collected from the interviews and quality assurance instruments and share the results with the CORE Team.

2. The CORE Team will review the findings and determine whether the measurements were met and if any changes are needed.

3. If it is determined that no changes are needed, the findings will be shared with the Implementation Team.

4. If changes are needed, the data will be shared with the Intervention Subcommittee who will be responsible for developing and identifying potential changes to the coordination process. The PII T/TAP and ET technical assistance providers, Darla Henry (3-5-7 Model), and Bob Friend (Family Finding Model) will be consulted for technical assistance and advice as needed. All recommendations will be shared with the Implementation Team which will determine the recommendations to be implemented.

Usability Test 2 – Determine whether the quality assurance tools will provide the FRP Project with the information needed to sufficiently measure fidelity to the interventions as specified in the Practice Profiles and compliance with contractual requirements. QA tools are case review protocols, observation checklists, survey instruments, etc. (Note that the assessment of the viability of the QA tools also will yield information about fidelity to the various interventions and compliance with contractual requirements.)

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria
1	Score on the effectiveness of each applicable QA tool in gathering information on fidelity to the 3-5-7 intervention.	Cross-walk to assess the match between information from the QA tools that assess 3-5-7 Model fidelity and practice profile specifications for the 3-5-7 Model and Youth Advocate. QA Tools #3 *Pre-Post Assessment*, #6 *Training Observation Checklist*, #8 *Case Record* Review, #9 *Clinical* Supervision, #11 *Youth Advocate Observation Rating Checklist*, #17 *Coaching Feedback*, #20 *Clinical Supervision Observation Checklist*	QA Plan #6, 6a – 6f, 9, 10, 14, 25, 26, 27, and 28 related to 3-5-7 Model fidelity.	QA Team	The requirement will be that **90** percent of the practice profile specifications regarding the 3-5-7 Model intervention are matched by the information obtained from the QA tools
2	Score on the effectiveness of each applicable QA tool in gathering information on fidelity to the Family Finding intervention.	Cross-walk to assess the match between information collected by the QA tools that assess Family Finding fidelity and the practice profile specifications. QA tools #4 and #15 *Pre-Post Assessments*, #7 *Training Observation Checklist*, #8 *Case Record* Review, #9 *Clinical Supervision Record* Review, #12 *CPS Specialist Observation Rating Checklist*, #18 *Coaching Feedback Survey*	QA Plan #7, 7a – 7h, 8, 20, 33, 34, 35, 36, 37 and 38 related to FF model fidelity.	QA Team	The requirement will be that **90** percent of the practice profile specifications regarding the Family Finding intervention are matched by the information from the relevant QA tools.

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria
3	Score on the effectiveness of each applicable QA tool in gathering information on fidelity to the CARE Team model.	Cross-walk to assess the match between information collected by the QA tools that assess CARE Team fidelity and the practice profile specifications. QA tools #2 *Pre-Post Assessment*, #5 *Training Observation Checklist*, #8 *Case Record Review*, #9 *Clinical Supervision Record Review*, #14 *CARE Team Meeting Observation Rating Checklist*, #16 *Coaching Feedback Survey*, #21 *Clinical Supervision Record Review*.	QA Plan #5, 5a – 5d, 11, 12, 13, 16, 21, 23, 29, 30, 31, 32, 37, 38, 39, 41 related to CT fidelity.	QA Team	The requirement will be that **90** percent of the practice profile specifications regarding the CARE Team intervention are matched by the information from the relevant QA tools.
4	Score on the effectiveness of each applicable QA tool in gathering information on compliance with contractual requirements	Cross-walk to assess the match between information collected from the QA tools and the requirements of the procedures manual and the DES Contract. QA tools #8, *Case Record Review*, #9 *Clinical Supervision Record Review*, #10 *Program Administrative Review*	QA Plan and Scope of Work/ Contract	QA Team	The requirement will be that **100** percent of the requirements specified in the DES contract and the Procedures Manual are captured by information from the relevant QA tools.
5	Assessment of barriers and facilitators to effectiveness of QA tools	Interviews with QA team members and Youth Advocates		Core Team	No Criteria: Information will be used to inform improvement strategies if necessary

Review of data

1. The QA Team will gather the data collected from the interviews and quality assurance instruments and share with the CORE Team.

2. The CORE Team will review the findings and determine whether any changes are needed.

3. If it is determined that no changes are needed, the findings will be shared with the Implementation Team.

4. If changes are needed, the data will be shared with the Evaluation/Quality Assurance Subcommittee who will be responsible for developing and identifying potential changes to the quality assurance process. The PII T/TAP and ET technical assistance providers, Darla Henry (3-5-7 Model), and Bob Friend (Family Finding Model) will be consulted for technical assistance and advice as needed. All recommendations will be shared with the Implementation Team which will determine the recommendations to be implemented.

Usability Test 3 – Determine if the young person is willing to initially meet with the Youth Advocate and if the young person will continue to participate after initial meetings with the Youth Advocate

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria
1	Percentage of young people randomly selected to be project participants who attend an initial appointment with the Youth Advocate within **10** business days of assignment to the project.	Log of young people randomly assigned to the project identifying young people who were assigned and began the project. Log of case assignment from AzCA	N/A	QA Team	The requirement will be that **75** percent of the young people selected to participate in the FRP will agree and attend at least one session with the Youth Advocate within 15 days of assignment
2	Percentage of young people who attend at least 1 session with the Youth Advocate who also attend 2 or more subsequent sessions during the period under review	QA Tool #8 *Case Record Review*	QA Plan #24	QA Team	The requirement will be that **88** percent of the young people who participate in at least 1 session with the Youth Advocate will attend 2 or more subsequent sessions during the period under review
3	Assessment of barriers and facilitators to young people participating in the project, including reasons for not participating initially or stopping participation in the project	Unstructured interview with the young person and qualitative analyses of results		CPS Specialist	No criteria – Information will be used to inform improvement strategies

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria
4	Assessment of resistance from other individuals in the young person's life to the young person participating in the project	Interviews with CPS Specialist and qualitative analysis of interview information		Core Team	Requirement will be that by the end of the period under review (e.g., usability testing period), strategies will be in place for addressing any identified resistance encountered to the young person participating in the project
5	Assessment of external conditions (i.e. transportation, scheduling, etc.) that may be a barrier to the young person participating in the intervention	Interviews with CPS Specialist and qualitative analysis of interview information		QA Team	Requirement will be that by the end of the period under review, strategies will be in place for addressing any identified external barriers to young person participation

Review of data

1. The QA Team will gather the data collected from the interviews and quality assurance instruments and share the results with the CORE Team.

2. The CORE Team will review the findings and determine whether the measurements were met and if any changes are needed.

3. If it is determined that no changes are needed, the findings will be shared with the Implementation Team, Darla Henry (3-5-7 Model) and Bob Friend (Family Finding Model).

4. If changes are needed, the data will be shared with the Intervention Subcommittee and/or the Evaluation/Quality Assurance Subcommittee who will be responsible for developing and identifying potential changes to the initial and ongoing engagement process. Darla Henry and Bob Friend will both be contacted for technical assistance and advice as needed. All recommendations will be shared with the Implementation Team which will determine the recommendations to be implemented.

Usability Testing 4 – Determine the viability of the clinical supervision processes for the Youth Advocate

	Measure	Individuals Participating in the Measure	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria
1	Number of supervisory sessions between the Youth Advocate and the Care Coordinator	QA Tools #8 *Case Record Review*, #20 *Observation Checklist of Clinical Supervision*	QA Plan #13 and 14	QA Team	The requirement will be that a formal supervisory session (i.e., a formal, scheduled session) occurred at least 1 time a week during the period under review
2	Score on a scale (Likert-Type) assessing the effectiveness of the supervisory sessions in assisting the Youth Advocate in doing his/her work i.e., not at all effective (1), rarely effective (2), somewhat effective (3), Usually effective (4), Always effective (5)	Structured interviews with Youth Advocate and Care Coordinator	n/a	QA Team	The requirement will be that the score will be **3.5** or higher on the scale for each Youth Advocate and Care Coordinator
3	Score on the Care Coordinator and Youth Advocate supervisory checklist	Observations of supervisory sessions (at least 4) or Case Record reviews QA Tools *#9 Clinical Supervision Record Review and #20 Observations Checklist of Clinical Supervision*	QA Plan #14	QA Team	The requirement will be that at least **75** percent of the items will receive a check by the end of the 3rd/4th session
4	Assessment of barriers and facilitators to the effectiveness of the supervisory process	Interviews with Care Coordinator and Youth Advocate		QA Team	No criteria – Information will be used to inform improvement strategies

Review of data

1. The QA Team will gather the data collected from the interviews and quality assurance instruments and share the results with the Usability Testing Subcommittee and CORE Team.

2. The Usability Testing Subcommittee and CORE Team will review the findings and determine whether the measurements were met and if any changes are needed.

3. If it is determined that no changes are needed, the findings will be shared with the Implementation Team and Darla Henry (3-5-7 Model Purveyor).

4. If changes are needed, the data will be shared with the Intervention Subcommittee who will be responsible for developing and identifying potential changes to the Youth Advocate supervisory process. Darla Henry will also be contacted for technical assistance and advice. All recommendations will be shared with the Implementation Team which will determine the recommendations to be implemented.

Usability Testing 5 – Determine the viability of the case mining and other processes for identifying and obtaining potential permanent people.

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria
1	Number of people identified who could possibly be a permanent family or connection for the young person through case mining within 30 days of the young person's enrollment in project	CARE Coordinator logs (completed by the assigned CARE Coordinator and another CARE Coordinator as a peer reviewer) QA Tool #8 - *Case Record Review* Another CARE Coordinator will be identified to review the case and a comparison of the findings of the two reviewers will be completed.	QA Plan #34	QA Team and Peer Case Review	The requirement will be that the extent of agreement between the assigned CARE Coordinator and other CARE Coordinator of the potential permanent family people or connections identified in the case mining process will be at least **80%**. (Based on taking the list of names of the total number of people identified from the assigned CC and the peer reviewer and comparing the highest list with the total number of the same people.)

	Measure	Methodology	Quality Assurance Cross-walk reference	Individuals collecting information	Criteria
2	Number of people identified in the first 30 days who could possibly be a permanent family or connection for the young person through other strategies – e.g., parent locator, calls to foster parents, information from young person, information from CPS Specialist, etc.	QA Tool #8 - *Case Record Review* and Care Coordinator logs	QA Plan #34	QA Team	The requirement will be that at least **10** potential permanent family options or permanent connections will be identified through processes other than case mining.
3	Assessment of barriers and facilitators to identifying people who could possibly be a permanent family or connection for the young person.	Interviews with Care Coordinator and CPS Specialist		QA Team	No Criteria: Information will be used to inform improvement strategies, if necessary

Review of data

1. The QA Team will gather the data collected from the interviews and quality assurance instruments and share the results with the Usability Testing Subcommittee and CORE Team.

2. The Usability Testing Subcommittee and CORE Team will review the findings and determine whether the measurements were met and if any changes are needed.

3. If it is determined that no changes are needed, the findings will be shared with the Implementation Team and Bob Friend (Family Finding Model).

4. If changes are needed, the data will be shared with the Intervention Subcommittee who will be responsible for developing and identifying potential changes to the initial family finding process. Bob Friend will also be contacted for technical assistance and advice. All recommendations will be shared with the Implementation Team which will determine the recommendations to be implemented.

KIPP

KANSAS INTENSIVE PERMANENCY PROJECT
A SERVICE OF THE CHILDREN'S BUREAU

KU THE UNIVERSITY OF KANSAS School of Social Welfare

March 2012

Usability Testing Report for the
Permanency Innovations Initiative in Kansas

KIPP Usability Testing

Introduction

In November 2011, KIPP began usability testing of its services, implementation, and evaluation/data collection procedures. Substantial planning and installation activities occurred prior to the kick-off of KIPP. The highlights of these installation activities are briefly summarized below to provide important background to the usability testing information that follows.

In July 2011, the KIPP Steering Committee began preparing to hire new KIPP staff and formed subcommittees for each staff position. These subcommittees developed written job descriptions and staff selection protocols to be used statewide by four different private foster care agencies. Once hired, KIPP supervisors and KIPP therapists undertook an extensive training and coaching process to work toward certification in the Oregon Model of Parent Management Training (PMTO). The training and certification process includes: 5 four-day workshops (approximately 160 hours); 2 full-days of in-person coaching sessions; 2 booster sessions; approximately 16 coaching sessions via video-conference, phone, or written feedback; and, a minimum of 10 fidelity ratings per therapist. All KIPP treatment sessions are video-recorded and used as potential data for training, coaching, and self-reflection/self-evaluation.

In addition to initiating a rigorous training regimen, the KIPP Steering Committee formed a subcommittee to plan staff and stakeholder orientations. The purpose of these orientations was to promote an understanding of the project and develop buy-in among staff and stakeholders to the service approach. KIPP's first official public event occurred in late September 2011, when the Steering Committee hosted a leadership summit. The first day of the leadership summit was attended by 49 stakeholders who represented a variety of community services, including mental health, substance abuse, and state level administrators. The latter three days of the summit convened individuals in key leadership positions of the KIPP agencies; it focused on the PMTO theoretical model, training curriculum, training and coaching program, adaptations for KIPP, planning, and problem-solving.

Finally, installation activities also included the formation of subcommittees that wrote a KIPP policies and procedures manual, hosted an all-staff orientation, created a KIPP brochure, designed a private and secure KIPP website, and developed a data collection system in REDCap. Additionally, KIPP leadership within each private agency engaged in several important outreach and orientation activities to create a hospitable environment for KIPP's initial implementation. Leaders held internal meetings with their executive managers and foster care staff to inform them of KIPP and make agency procedural modifications as needed. They also met with judges and other court personnel to introduce them to KIPP. By early November 2011, the implementation drivers of selection, training, coaching, performance assessment, decision support data systems, and facilitative administrative changes by KIPP agencies, had been installed. KIPP was ready to enter its initial implementation phase.

Initial Implementation and Usability Testing

Nine usability testing (UT) metrics were identified by the KIPP Steering Committee in consultation with our site leads for the PII-Training/Technical Assistance Project (TTAP) and the PII Evaluation Team (ET). Each metric was selected because it represented a fundamental task or milestone related to the initial implementation of KIPP. Gathering data on these metrics allowed us to: 1) test whether critical activities of KIPP operated as planned and expected; 2) detect obstacles to implementing essential activities; and, 3) determine whether obstacles were region-specific or statewide, and 4) engage in problem-solving related to identified challenges.

Table 1 shows KIPP's nine UT metrics and their associated targets. The metrics comprise three domains of information related to KIPP's installation and initial implementation: 1) eligibility processes (i.e., metric 1); 2) evaluation/data collection processes (i.e., metrics 2 through 5); and, 3) initial service delivery processes (i.e., metrics 6 through 9). All targets were set at 70%. The 70% target was set as the metric for eligibility and evaluation/data collection processes because it was determined that a 70% level would be sufficient to engage enough families and youth for evaluation purposes and to have a service seen as useful and viable by the community. The 70% target for initial service delivery processes was deemed adequate because families and youth were asked to engage in intensive interaction with the therapist (e.g. videotaping, frequent home visits) shortly after having their child placed in foster care. In light of these circumstances, a 30% non-engagement rate seemed reasonable.

Table 1. KIPP's Usability Testing Metrics and Targets

Domain	Metric	Target
Eligibility	1. Percent of CAFAS/PECFAS that were completed by day 14 of the child's foster care episode	70%
Evaluation/Data Collection	2. Percent of cases randomly assigned within 2 working days of KU receiving request from the foster care agency	70%
Evaluation/Data Collection	3. Percent of parents in the treatment group that agree to participate in KIPP treatment group	70%
Evaluation/Data Collection	4. Percent of parents in the comparison group that agree to participate in KIPP comparison group	70%

Domain	Metric	Target
Evaluation/Data Collection	5. Percent of Time 1 assessments that were completed within 7 working days of group assignment (treatment group only)	70%
Initial Service Delivery	6. Percent of parents in treatment group that were contacted by KIPP therapist within 3 working days of referral from the KIPP supervisor	70%
Initial Service Delivery	7. Percent of parents in the treatment group that participated in video-recordings of the KIPP/PMTO intervention	70%
Initial Service Delivery	8. Percent of parents in the treatment group that continued to participate in KIPP beyond session 2	70%
Initial Service Delivery	9. Percent of treatment cases who had a session with KIPP therapist, parent, and child at least once a week after session module B3	70%

Study Period and Sample

Study Period

KIPP's usability testing (UT) metrics were measured during a study period that spanned three months, October 17, 2011 to January 18, 2012. The UT study period started on the date that foster care agencies began administering the Child and Adolescent Functional Assessment Scale (CAFAS) and the Preschool and Early Childhood Functional Assessment Scale (PECFAS).[1] The first UT cases were assigned on November 7, 2011. The final UT case was assigned on January 18, 2012. This report summarizes findings of UT cases as of February 17, 2012.

Sample

The UT sample consisted of 118 cases assigned between November 7, 2011 and January 18, 2011. Cases were selected for random assignment if they met the KIPP eligibility criteria.

Eleven (11) comparison cases and 98 treatment cases were assigned across five regions. Eight of the 98 treatment cases were identified as "dual reintegration"[2] cases. That is, these cases involved divorced or separated parents who were working separately on dual, or parallel, paths toward reunification. Thus, the total number of treatment cases was 98 + 8 = 106.

NUMBER OF KIPP UT CASES

Treatment cases, N 106
Comparison cases, N 11
Total cases, N 117

[1] CAFAS is administered with children aged 6-18 and the PECFAS is administered with children aged 3-5.
[2] "Dual reintegration" is a term used by Kansas child welfare agencies. *Reintegration* and *reunification* are generally used as interchangeable terms.

Ineligible Cases

An additional 19 cases were assigned but were excluded because they did not meet eligibility rules. The primary reasons that cases did not meet eligibility rules were: 1) no case plan goal of reunification; 2) the parent(s) were not available for services because they had moved from the state of Kansas, were incarcerated, or could not be located; and, 3) the child had been discharged from foster care to the Juvenile Justice Authority.

Findings

Table 2, below, summarizes the results of KIPP's nine UT metrics. Color-coded shading is used to indicate the level of performance on each metric. Green signals metrics for which the target was met; yellow indicates that the performance was below the target of 70%, but at or above 50%; and red shows metrics for which performance was substantially low (i.e., below 50%).

At the statewide level, seven of nine metrics met the pre-determined target of 70%. Six of these were met at 91% or higher. Metric 3—the metric that measured the percent of parents that agreed to participate in the KIPP treatment group—fell just shy of the target at 68%. Metric 5, on the other hand, missed the mark considerably with a performance rate of 49%. Below, each metric's performance is described in detail.

Table 2. Summary of Findings on Usability Testing Metrics

Metric Definition	Reg 1		Reg 2		Reg 3		Reg 4		Reg 5		Statewide	
	Percent	N	Percent	N	Percent	N	Percent	N	Percent	N	Percent	N
1 - % of CAFAS that are completed by day 14	90%	71	77%	71	88%	91	93%	85	77%	57	**86%**	375
2 - % of cases randomly assigned within 2 working days of receiving request from agency	100%	20	100%	27	100%	35	100%	19	100%	16	**100%**	117
3 - % of parents in treatment group that sign consent form	78%	18	64%	25	52%	33	71%	17	100%	13	**68%**	106
4 - % of parents in comparison group that sign consent form	100%	2	100%	2	100%	2	100%	2	67%	3	**91%**	11
5 - % of Time 1 assessments that are completed within 7 working days of group assignment (tx cases only)	14%	14	63%	16	53%	17	75%	12	38%	13	**49%**	72
6 - % of parents contacted by KIPP Therapist within 3 working days of referral from KIPP Supervisor	92%	13	94%	16	100%	13	92%	12	92%	12	**94%**	66
7 - % of parents who participate in video recordings of KIPP/PMTO intervention	100%	12	100%	16	92%	12	91%	11	100%	12	**97%**	63
8 - % of parents who continue to participate in KIPP beyond session 2	100%	9	100%	14	100%	8	100%	8	90%	10	**98%**	49
9 - % of parents who had a session with the KIPP Therapist, parent and child at least once a week	100%	7	92%	12	100%	5	100%	2	100%	6	**97%**	32

Metric 1

% of CAFAS completed by day 14 (target = 70%)

Metric 1 was measured as of December 27, 2011. This date was selected because the final UT case was assigned on January 18, 2012, and to accurately measure this metric, CAFAS data must be observed 14 working days prior to the assignment date—December 27, 2011.

From October 17, 2011 to December 27, 2011, 449 children between 3-16 years old entered foster care. At the time of this report, data were missing or were being reconciled on 74 of the 449 cases. The

major reason for missing data was that the child had been discharged from foster care prior to day 14. Among the 375 cases for which we had complete data, the CAFAS had been administered by day 14 on 86% of them. Even if all 74 of the missing data cases were assumed to have failed this metric, then the statewide rate still would have been within the target range at 72%.

Metric 2

% of cases randomly assigned within 2 working days of request (target = 70%)

The total number of cases assigned to treatment and comparison groups was 117. KU was able to assign 100% of the cases to the agency within 2 working days of the agency requesting a case assignment.

Metric 3

% of treatment group parents that consent to participate (target = 70%)

The sample for Metric 3 included 98 families assigned to the treatment group, plus 8 families identified as "dual reintegration" families. Thus, the total number of families assigned to the treatment group was 106. Of the 106 eligible treatment families, 72 agreed to participate and 35 declined to participate, a successful consent rate of 68%.

Metric 4

% of comparison group parents that consent to participate (target = 70%)

Metric 4 comprised 11 families assigned to the comparison group. Ten (10) of 11 families assigned to the comparison group agreed to participate, a successful consent rate of 91%.

Metric 5

% of Time 1 assessments completed within 7 working days of group assignment (target = 70%)

Seventy-two (72) families agreed to participate in the KIPP treatment group. Among these 72 families, 35 families (49%) completed the evaluation assessments within 7 working days of case assignment. Metric 5 was marked as red because it was substantially lower than the 70% target we aimed to achieve.

Table 3, below, shows the average number of working days between case assignment and completed assessments by region. Statewide, it took 9.1 working days for KIPP Data Liaisons to get evaluation assessments completed with families. It ranged from a low of 6.0 working days in region 4 to a high of 12.3 working days in region 1.

Table 3. Average Number of Days Between Case Assignment and Completed Assessments

	Region 1	Region 2	Region 3	Region 4	Region 5	Statewide
Average # of days to complete evaluation assessments	12.3	8.0	8.4	6.0	11.2	9.1

Metric 6

% of cases contacted by KIPP therapists within 3 working days of referral

The numerator for Metric 6 was 66. That is, of the 72 cases that agreed to participate in KIPP treatment, 6 were not yet referred to a KIPP therapist and 66 cases were referred to a KIPP therapist. Sixty-two (94%) of the families referred to a KIPP therapist were contacted within 3 working days of referral. The average number of days it took the KIPP therapist to contact families was 1.1 working days; the mode was 0 days; and the range was 0 to 12 working days.

Metric 7

% of cases who participate in video recordings of KIPP/PMTO services

Seventy-two (72) families agreed to participate in KIPP treatment. Nine (9) families were excluded from the numerator of Metric 7 for the following reasons:

- 4 families were waiting to complete evaluation assessments
- 3 families were scheduled to begin but had not yet started KIPP services
- 1 family was court-ordered to stop KIPP services
- 1 family was unexpectedly discharged by the court before KIPP services could begin

Thus, the numerator for Metric 7 was 63 families. Among these 63 families, 61 (97%) participated in video recordings of KIPP/PMTO sessions. The reasons that the two families did not participate were as follows:

1. These parents were reluctant to begin KIPP services because their teen child was missing (runaway).
2. This parent dropped out before services began; the parent changed his mind and stated that he was not comfortable with video-recording.

Metric 8

% of cases who continued to participate in KIPP beyond session 2

Among the 72 families that agreed to participate in KIPP treatment, 49 were considered valid for Metric 8. Twenty-three (23) families were excluded from this metric for the reasons listed below.

- 9 families had not begun KIPP treatment as described under Metric 7
- 14 additional cases were excluded due to the following reasons:
  - 9 were scheduled for sessions
  - 3 discharged by the court
  - 1 dropped out, accounted for under metric 7
  - 1 not sure about KIPP, accounted for under metric 7

Metric 8 was met for 98% of the eligible families. Forty-eight (48) of the 49 families participated in KIPP services beyond the second KIPP/PMTO session.

Metric 9

% of cases who had a session with KIPP therapist, parent, and child at least once per week (sessions beyond B3)

Among the 72 families that agreed to participate in KIPP treatment, 32 were considered valid for Metric 9. Forty (40) families were excluded from this metric for the reasons listed below.

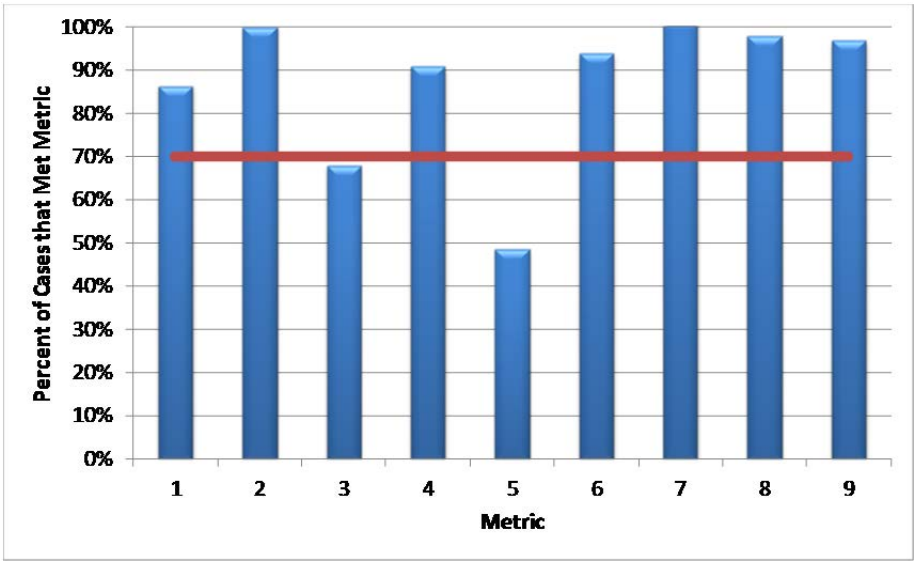
- 24 families were excluded for the reasons described under Metrics 7 and 8.
- 16 families were progressing but had not completed session module B3 as of the date of this report. Among these 16 families, case status notes indicated possible concerns about case progression on 4 cases. Three families were noted for no-showing on their scheduled appointments. One family delayed sessions with the parent, child, and KIPP therapist because the child was in the juvenile detention center.

Thirty-one (31) of 32 families (97%) had begun KIPP sessions that included the parent, child, and the KIPP therapist, and were progressing as planned.

Lessons Learned

Usability testing on KIPP's nine metrics demonstrated that the implementation of KIPP was generally stable across all five regions of Kansas. **Figure 1**, below, shows that KIPP achieved its target of 70% for all metrics with only two exceptions. While Metric 3 was only slightly below the 70% target (68%), it was considered a metric of high value and weight to the KIPP implementation. On the other hand, Metric 5 was well below the 70% target at 49%, but was seen as less essential to the central and core functions of the KIPP implementation. In both cases, the KIPP Steering Committee took seriously these results that showed lower-than-desired performances and addressed each of them as described below. In addition to drawing attention to consent procedures (Metric 3) and the timing of data collection (Metric 5), usability testing revealed several other implementation and evaluation processes that required further refinement in the early months of KIPP's implementation. For each of these, the lessons learned and the adjustments made to KIPP are explained below.

Figure 1. KIPP's Performance on 9 Usability Testing Metrics with Reference to a 70% Target



Metric 3 – Obtaining Parental Consent for Participation in KIPP Treatment Group

The target for Metric 3 was that *70% of parents assigned to the treatment group would agree to participate in KIPP and sign consent forms.* During usability testing, we attained a successful consent rate

of 68% with parents of the treatment group. This slightly low consent rate was not surprising or unexpected. Since KIPP is an intensive, in-home treatment that requires a substantial investment of time from parents, we expected some parents to be reluctant to participate in KIPP treatment. Beyond asking parents to voluntarily invite a KIPP therapist into their home three times a week, KIPP also requires parents to agree to a rather unconventional request—that is, parents must agree to video-record every KIPP session. Given these expectations for video-recording, some refusals were anticipated. In the end, the KIPP Steering Committee concluded that a 68% successful consent rate was adequate and acceptable. Still, we closely examined regional and statewide performance on Metric 3 and discussed strategies for improving the consent rate among parents assigned to the treatment group.

KIPP's relatively lower performance on Metric 3 can be primarily attributed to two areas—regions 2 and 3. Region 2's consent rate was 64%, while region 3's was 52%. Both of these regions struggled with persuading parents to participate in KIPP services, particularly parents of older youth. The average age of focal children randomly assigned to regions 2 and 3 was higher than the average age of focal children randomly assigned to other regions. Specifically, regions 2 and 3 had an average age of 13, while regions 1 and 4 had an average age of 12 and region 5—the region with the highest consent rate—had an average age of 10. Among the families who declined to participate in KIPP treatment, 74% had a focal child who was 12 or older. In contrast, among the families who agreed to participate in KIPP treatment, 61% had a focal child who was 12 or older. The KIPP Steering Committee discussed this issue and unanimously agreed that parents of older youth are generally harder to engage in child welfare services.

Region 3's challenges with Metric 3 were especially noticeable because they consistently had the lowest consent rate throughout usability testing. After reviewing usability data, region 3's leadership took several key steps to attempt to improve their consent rate. In mid-February, region 3's directors decided to replace the KIPP supervisor. They have also initiated several other strategies, such as: 1) pairing the region 3 supervisor with the region 1 supervisor, who has experienced more success with obtaining parental consent; 2) gaining support of the foster care case management team; 3) working closely with the foster care case management team to reach parents (e.g., attending case planning meetings held by the case manager to meet parents in person and invite them to participate in KIPP); and, 4) requiring the supervisor to make multiple attempts and in-person contact with parents.

In addition to regions working locally to bolster consent rates, the KIPP Steering Committee will continue to closely monitor consent rates in all five regions. The KIPP Implementation Team, made up of agency directors that represent the five regions, will also regularly discuss consent and other engagement issues at its bi-monthly meetings. Finally, the KIPP evaluation liaison hosts quarterly meetings with KIPP supervisors during which the agenda regularly covers strategies for parent and youth engagement.

Metric 5 – Completing Time 1 Assessments Within 7 Working Days

Metric 5 established a timeline for completing time 1 evaluation assessments based on KIPP's assumption that KIPP therapist should engage parents *early* in the life of a case. We guesstimated that 70% of time 1 evaluation assessments could be completed within 7 working days of assignment to the treatment group. Usability testing data showed that we underestimated the amount of time it would take to complete time 1 evaluation assessments. Data liaisons were able to complete time 1 evaluation assessments within 7 working days about one-half of the time. In contrast to the findings on Metric 3,

which showed that the problem was largely isolated to two regions, the findings on Metric 5 showed that these glitches were experienced statewide. That is, only one of the five regions – region 4 – was able to meet the target of 70%.

Upon reviewing the data on Metric 5, the KIPP Steering Committee explored the reasons that time 1 evaluation assessments did not occur on a quicker timeline. We learned that the biggest obstacle was scheduling the Family Interaction Task (FIT) assessment, which requires a video-recording of parents and focal child together. Most challenging is the logistics of scheduling multiple parties – the child, the parents, the foster parents, the case manager, the transportation provider, the data liaison, a person to monitor for safety, etc. FITs are typically scheduled for a late afternoon, evening, or weekend when children are not in school and caregivers are not at work. Frequently, agencies must transport children from several hours away, and transportation providers require appointments to be scheduled with several days' notice. The time it took to complete evaluation assessments was drawn out further in those situations that required waiting for a child to exit a psychiatric facility or a parent to be released from jail. One court jurisdiction routinely delays the FIT assessment by requiring all parents to provide a clean urine analysis prior to any contact with their children. In many cases, the FIT assessment is scheduled to coincide with a regular parent/child visit[3] because this is efficient and makes the most sense for transporting children. For example, it does not make sense to transport a child for three to six hours solely for a 30-minute FIT assessment. Data liaisons also report experiencing a number of no-shows, which require the logistics planning to start all over. In sum, time 1 evaluation assessments can be delayed for a variety of complex factors that are difficult to change.

KIPP usability testing demonstrated that the timeline for time 1 evaluation assessments needed to be extended beyond 7 working days. Accordingly, the KIPP Steering Committee adjusted this metric for further monitoring and set the timeline at 14 working days.

Eligibility Clarification

During usability testing, several eligibility criteria were modified or clarified. The following is a summary of the eligibility topics that were addressed.

SED Determination

The timeframe in which a foster care agency determines the presence of serious emotional disturbance (SED) was expanded from 14 days to 45 days. This modification was made to ensure adequate time for an accurate assessment of children's mental health functioning. The issue initially emerged by the second week of KIPP implementation when a foster parent reported behaviors of a child that would indicate the presence of SED, but had not been observed by day 14 of the child's foster care episode and, thereby, included in the data used to complete the CAFAS/PECFAS.

Foster care agencies continue to screen all children between the ages of 3 and 16 years old by day 14 of foster care entry. Expanding the window to 45 days allows sufficient time for rescreening if information emerges after the initial CAFAS/PECFAS score is given.

[3] In the state of Kansas, private foster care agencies are required by contract to facilitate a weekly parent/child visit. The KIPP evaluation assessments may not be used to replace a regular parent/child visit.

Children Re-Entering Foster Care

During usability testing, we clarified that while children re-entering foster were generally eligible for KIPP if they met other eligibility criteria, those who entered care after a failed trial home visit would not be considered eligible.

Families That Move Outside the Service Area and/or Change Venues

Families that move outside the KIPP service area, or outside an agency's region, may become ineligible for KIPP services. The foster care agency may consider proximity of the family to the assigned KIPP therapist. If the family's residence is outside the agency's service area and is too far to deliver in-home services three times per week, then the family will not be eligible for KIPP services.

Incarcerated Parents

If a parent is incarcerated at the time of random assignment to the treatment group, the foster care agency may make efforts to obtain consent. An incarcerated parent may participate in KIPP services if they meet other eligibility criteria with regard to engagement of services. That is, if a parent can begin KIPP services within 3 months of assignment, they may participate.

Changes in Case Plan Goal

Cases that present with imminent changes in case plan goal will be rendered ineligible for KIPP services (i.e., case plan goal will no longer be reunification). If a case is likely to have a change in case plan goal within 45 days of initiating KIPP services, that case may be deemed ineligible.

Engagement Protocol – Parents

Obtaining Consent. Usability testing on Metric 3—consent from parents randomly assigned to the treatment group—clarified efforts that KIPP supervisors should make to obtain consent. Supervisors will make every effort to engage parents on the phone, enlist case managers to help engage parents, and engage parents in person as soon as possible. If consent has not been obtained after two weeks of repeated attempts to engage parents by phone or in person, the KIPP supervisor may cease efforts and report the case as "declined participation."

Inactive Case Protocol. Usability testing on metrics that required observation of case progression (e.g., metrics 7, 8, and 9) provided an opportunity to examine protocols for cases in which parents do not ultimately engage in KIPP services. The following is the *KIPP Inactive Case Protocol* that was developed during usability testing.

> Once consent has been gained, attempts should be made to engage parents in KIPP sessions. If barriers to engagement can be identified and eliminated, effort should be made to do so (e.g., providing concrete assistance). Diligent efforts should continue for two weeks. Diligent efforts include making phone calls, visiting the parent's current residence, coordinating a meeting with the foster care case manager, and scheduling in-person meetings. If, after two weeks, the parent does not respond to repeated attempts to schedule or complete an initial appointment, the therapist will notify the supervisor, who may choose to attempt contact with the parent. When the supervisor is satisfied that the case is inactive, he or she will notify the agency's KIPP Implementation Team (KIT) member. At the discretion of the KIT member, the agency may choose to discharge the case or continue making efforts up until 3 months have passed. The case may not be discharged until both

supervisor and KIT member have been notified. The supervisor or KIT member will notify the KIPP evaluation liaison. The therapist responsible for the case will ensure that a completed Discharge Note is submitted to KIPP's REDCap website.

Engagement Protocol – Youth

Early on in usability testing, we experienced older youth who objected to participating in KIPP services and/or video-recordings of KIPP services. These early challenges provided the catalyst for the KIPP team to develop the following *Youth Engagement Protocol*.

- To qualify as a training family[4], the family must include: 1) a parent willing to work toward reunification, and 2) the randomly assigned focal child who is in out-of-home care at the start of treatment.
- Parental consent must be obtained prior to the start of treatment. Additionally, the focal child should receive information about KIPP services and evaluation, and assent documentation should be completed.
- All evaluation assessments should be completed, including the FIT. The FIT should be completed with the reunifying parent(s) and the focal child.[5] However, if the child refuses, the case will be staffed on a case-by-case basis by the KIPP supervisor, the KIT member, and the KIPP evaluation liaison. A decision will be made regarding the case's ability to move forward into KIPP services.
- <u>If the child refuses to be video recorded</u>: KIPP/PMTO sessions may begin with the child off camera. It is the therapist's job to engage the child in video-recording, and they should continue to make attempts to do so.
- <u>If a child refuses to participate in parent-child-therapist sessions</u>: Repeated efforts should be made to engage the child. These may include individual preparation and engagement sessions by the KIPP therapist assigned to the family or by another KIPP therapist, if available.
    - If, after extraordinary efforts to engage the focal child—and after completing PMTO session module C3—the focal child still refuses to participate, the KIPP supervisor will consult with ISII to determine whether the case can count as a training case. Additionally, the KIPP supervisor will inform the KIT member and the KIPP evaluation liaison of the case disposition.
- CAVEAT: Given the unpredictability of engagement and the likelihood of attrition, it is possible that KIPP therapists will work with KIPP treatment families that will count in the evaluation but will not count as PMTO training families. This is acceptable to KU and ISII; however, it will likely lengthen the KIPP therapists' certification processes. Clinical judgment and agency discretion should be used to determine a case's disposition.

---

[4] "Training families" are families that the KIPP therapist uses for her or his training in PMTO. Therapists must video-record every session with these families. The sessions are observed by PMTO coaches in order to rate the therapists fidelity and to coach on the PMTO model.

[5] FIT procedures have not yet been developed for children with significant developmental disabilities or for children under five. FITs should not be completed with children in either category until these procedures have been finalized.

Discharge Protocols

During usability testing a handful of cases discharged to reunification unexpectedly, mainly due to court decisions that proceeded without the recommendation of the foster care agency. In response, we developed the following discharge protocol.

> If a treatment family reunifies unexpectedly, then the therapist will work with the family for a minimum of one-month to transition them to case closure and, if applicable, to aftercare services. At the agency's discretion, and with the voluntary participation of the family, the KIPP therapist may continue working with the family until they have completed the PMTO curriculum.

Areas Identified for Additional Tailoring

The need for additional tailoring of the intervention and its measurement were identified during usability testing. Because PMTO is largely a verbal, behavioral intervention, it does not align well with the needs of two subgroups of youth who will likely comprise a significant proportion of our intended target population (children 3-16 with SED): 1) very young children (aged 3-5); 2) and children with co-occurring SED and developmental disabilities. The FIT, for example, is not yet tailored for the needs of these groups; observations of FITS conducted with nonverbal children brought this issue to our attention. ISII is working with a child development specialist to modify the FIT and the intervention so that it better meets the needs of our youngest, most vulnerable focal children.