A Comparison of Methods for Measuring Implementation Fidelity of a Pregnancy and HIV Prevention Program

Pamela Drake, PhD

Senior Research Associate

ETR Associates, Inc.

RTRworks!

What is implementation fidelity?

- Adherence: delivered as prescribed
- Exposure: dosage/amount
- Quality of delivery
- Participant responsiveness: reaction/engagement
- Program differentiation: distinguish critical components

(Dane & Schneider, 1998; Durlak & DuPre, 2008; Dusenbury et al., 2003; Fagan et al., 2008)

RTRworks!

Why measure implementation fidelity?

- Interpret outcomes

- Assess feasibility of implementation

- Determine which components produce outcomes

- Evaluate training programs

RTRworks!

- NIH-funded SBIR

- Develop an online training to improve teacher fidelity when implementing Reducing the Risk (RTR)

- Focus on role plays: role play management, engagement, sensitive questions, and inclusion of LGBTQ students

Study Design

- Randomized-Control Trial

- National, volunteer sample of 8th-12th grade teachers who were implementing or wanted to implement RTR (44 states; 221 enrolled; 164 post-surveys)

- Control group used only the RTR curriculum guide

- Tx group had access to RTRworks! online training

- Training was interactive, self-directed, accessed throughout implementation

RTRworks!

What do we know about measuring implementation fidelity?

- It's difficult to get teachers to complete fidelity logs
- Teachers tend to inflate fidelity
- Observation is the "gold standard" but is resource intensive

RTR works!

Our Measures

- Online presurvey

- Online postsurvey

- **Online implementation logs – 16 classes

- **Observations (in-person and audio) – 4 classes

- **Lesson-specific interviews – 4 classes

- Follow-up interviews

RTRworks!

Strategies to increase completion of implementation logs

- My Study Page
- Reminders
- $25 Incentive
- Decreased burden
  - 12 "short" logs and 4 "long" logs (4, 5, 10, 11)
  - As little detail as possible

RTRworks!

My Study Page

- Teacher information
  - School; Address; Study group
- Implementation schedule
  - Link to logs; Tracking of incentives
- Other Study Activities
- Link to Training

RTR works!

My Schedule

RTR Class/ Lesson *These four classes are the ones to be audio-taped*	Class Date Planned *Date format: xx/xx/xxxx (no dashes)*	Class Date Actual	Study Data	Date Log or Survey Submitted	Stipend earned
Class 1	(date)	(date)	Log 1 ($25)	(date)	$
Class 2	(date)	(date)	Log 2 ($25)	(date)	$
Class 3	(date)	(date)	Log 3 ($25)	(date)	$
Class 4*	(date)	(date)	Log 4 ($25)	(date)	$
Class 5*	(date)	(date)	Log 5 ($25)	(date)	$

RTRworks!

Our primary fidelity item

1. Did you implement each of the following activities of this class?	Implemented Completely	Implemented Partially or with Changes	Did not Implement
Have students complete My Kid Sister individually then discuss reasons not to have sex, reasons to have sex, and good ways to encourage kids not to have sex	O	O	O
Divide students into small groups and have all students within a group rotate through both role play roles and observer, including small group discussion	O	O	O
Tell students they didn't have to use the scripts and could create their own lines	O	O	O
Discussed role plays	O	O	O

RTRworks!

Observations and Interviews

- Focused on 4 key classes: 4, 5, 10, 11
- Planned 24 observations and 24 interviews
  - 3 Tx & 3 C for each class
  - 8 overlap of observations and interviews (1 for each class by group)
- 4 observers very familiar with RTR curriculum (2 training developers); 2 interviewers
- Interviews took place within 48 hours of class; teacher had to have completed log prior

RTRworks!

Audio Observations

- Small digital recorders

- Straps to hang around neck

- Detailed directions

- Incentives ($25 each)

- Reminders on prior class log

RTRworks!

Data Completion

- Logs:
  - 171 of 221 enrolled completed Class 1 log
  - 162 (95%) completed Class 16 log
  - Average 5 days after implementation; sd=~13 days
- 24 interviews; 25 observations
  - Small variations from what was planned
  - 7 interview/observation overlap (5 also had audio)
- 88 Audio observations (79 recorded all 4 key lessons)

RTRworks!

Observation Challenges

- Getting teachers to commit to and record a schedule of classes

- Determining remote locations where observer could get max number of observations in min days

- Teachers rescheduling

- Weather

- Block classes

Analyses

- Internal consistency of teacher logs
- Created "fidelity" scores for each lesson and overall (percentage)
  - Fully implemented = 1
  - Partially implemented = .5
  - Not implemented = 0
- Calculated agreement between measures on single activities and summed across activities

RTRworks!

What we found – Teacher Logs

- Internal consistency of fidelity scores (for 4 key classes) based on teacher logs was good
  - 24 items (activity)
  - Cronbach's alpha = 0.73
- For N=151 teachers, fidelity scores based on teacher logs were high

	All Teachers	Observation Group	Interview Group
Range	81%-88%	79%-96%	77%-86%
Average	86%	84%	83%

RTR works!

What we found – Observations

- Teachers and observers agreed on activity ratings only about half the time (n=5-8)

- Fidelity scores based on observations were consistently lower than the teacher self report

Class	Matched teacher logs	Observations	Agreement
4	81%	75%	47%
5	79%	64%	59%
10	96%	50%	25%
11	80%	75%	55%
Average	**84%**	**66%**	**50%**

RTR works!

What we found – Interviews

- Fidelity scores based on interviews were very similar to the self report

  - 69% of time without agreement, teachers said they would change rating to match interviewer

- Agreement was over 90% (n=6 for each class)

Class	Matched teacher logs	Interviews	Agreement
4	81.94%	81.94%	92%
5	85.83%	85.83%	95%
10	85.42%	81.25%	92%
11	77.10%	75.00%	98%
Average	**83%**	**81%**	**92%**

RTRworks!

What we found – Activity Types

- Fidelity scores and agreement between teachers and observers varied across activity types

Activity Type	T Mean Fidelity Score (%)	O Mean Fidelity Score (%)	% Agreement
Worksheets	89%	73%	**68%**
Review skills	**93%**	74%	57%
Generate alternatives	92%	**77%**	60%
Demo role play	89%	**52%**	**29%**
Role play	82%	59%	44%
Miscellaneous role play	**65%**	68%	62%
Discuss role play	82%	56%	**28%**

RTR works!

Other Fidelity-Related Items

- Student Involvement
  - Just read their lines from the role play scripts without acting the parts out
  - Acted in the role plays as if they were real life situations
  - Attempted to use eye contact and body language to emphasize their points
  - Came up with good refusal lines
  - Seemed competent at using the refusal skills

RTRworks!

Other Fidelity-Related Items

- Teacher response to role play issues
  - Students off-track in small groups
  - Students complaining about lack of relevance to their lives
  - Students acting or saying they were uncomfortable or intimidated doing the role play
  - Students demonstrating lack of understanding about what to do during the role plays
  - Student resistance to do the role plays
  - Student refusal to participate in the role plays either performing or as observers

RTRworks!

What we found – Other Items

- Agreement varied on other items

Activity Type	Range	Average
Student Involvement	24%-33%	26%
Teacher Response	33%-81%	54%

- Highest: lack of relevance, refusal to participate

- Lowest: off-track, resistance, lack of understanding, uncomfortable

RTRworks!

What we found – Audiotapes

- Some behaviors could easily be coded on the audiotapes, but not all – varies

- Fidelity scores based on audiotapes were more similar to those based on observation than logs

- However, agreement with both was around 70%

Comparison Measure	Matched fidelity score (%)	Audiotape fidelity score (%)	Agreement
Teacher logs	81%	74%	70%
Observations	73%	74%	73%

RTRworks!

Why the differences?

- Social desirability

- Recall
  - Time lag
  - Note taking

- Interpretation

RTRworks!

Conclusions

- Online tracking system, online logs, incentives increase completion of logs but still had a significant lag time

- In-person observation, may be a more accurate measure of fidelity, but for projects like ours they are cost-prohibitive and logistically difficult

- Interviewers may be biased by teacher responses - training

- Audio observations may be an alternative, but only for behaviors that can be picked up via recorder; still resource-intensive to code

RTR works!

Lingering Questions

- Does observation really yield the most accurate fidelity score?
  - Who is the best observer/interviewer? How do we judge?
- How do we balance the need for quality results against available resources?
- Can we provide any type of training on self-report logs that doesn't interfere with our outcomes?
- Are errors random across TX and C groups?
- Why do we have these differences?

RTR works!