

Published in final edited form as:

*Child Youth Serv Rev.* 2013 January 1; 35(1): 188–193. doi:10.1016/j.chilyouth.2012.10.002.

## Examining the equivalence of fidelity over two generations of KEEP implementation: A preliminary analysis<sup>☆</sup>

Rohanna Buchanan<sup>a,\*</sup>, Patricia Chamberlain<sup>a</sup>, Joseph M. Price<sup>b</sup>, and Peter Sprenghelmeyer<sup>c</sup>

<sup>a</sup>Oregon Social Learning Center, 10 Shelton McMurfhey Blvd., Eugene, OR 97401, United States

<sup>b</sup>Department of Psychology, San Diego State University, 6363 Alvarado Ct., Suite 103, San Diego, CA 92120, United States

<sup>c</sup>OSLC Community Programs, 1170 Pearl St., Eugene, OR 97401, United States

### Abstract

**Problem**—In order to obtain and maintain positive outcomes garnered from evidence-based practice (EBP) models, it is necessary to implement them effectively in “real world” settings, to continually monitor intervention fidelity to prevent drift, and to train new staff due to turnover. The fidelity monitoring processes that are commonly employed in research settings are labor intensive and probably unrealistic to employ in community agencies given the additional burden and cost that they represent over and above the cost of implementing the EBP. Efficient strategies for implementing fidelity monitoring and staff training procedures within the inner context of agency settings are needed to promote agency self-sufficiency and program sustainability.

**Method**—A cascading implementation model was used whereby agencies who achieved proficiency in KEEP, an EBP designed to prevent placement disruptions in foster and kinship child welfare homes, were trained to take on fidelity management roles to improve the likelihood of program sustainability. Agency staff were trained to self-monitor fidelity and to train internal staff to achieve model fidelity. A web-based system for conducting fidelity assessments and for onsite/internal and remote program quality monitoring was utilized.

**Results**—Scores on fidelity ratings from streamed observations of intervention sessions showed no differences for foster parents treated by first generation interventionists trained by model developers compared to a second generation of interventionists trained by the first generation.

**Conclusion and relevance to child welfare**—Development of the local intra-agency capacity to manage quality intervention delivery is an important feature of successful EBP implementation. Use of the cascading implementation model appears to support the development of methods for effective monitoring of fidelity of the KEEP intervention, for training new staff, and ultimately for the development of internal methods for maintaining program sustainability and effectiveness.

### Keywords

Implementation fidelity; Equivalence design; Cascading dissemination; KEEP intervention

---

<sup>☆</sup>Funding Source: Support for this research was provided by Grant Nos. R01 MH 60195 from the Child and Adolescent Treatment and Preventive Intervention Research Branch, DSIR, NIMH, U.S. PHS; R01 DA 15208 and R01 DA 021272 from the Prevention Research Branch, NIDA, U.S. PHS; R01 MH 054257 from the Early Intervention and Epidemiology Branch, NIMH, U.S. PHS; and P20 DA 17592 from the Division of Epidemiology, Services and Prevention Branch, NIDA, U.S. PHS.

© 2012 Published by Elsevier Ltd.

<sup>\*</sup>Corresponding author: Tel.: +1 541 485 271; fax: +1 815 550 1789. rohannab@oslc.org (R. Buchanan).

## 1. Introduction

The implementation of evidence-based treatments in child mental health systems has become a national priority (Hoagwood, Burns, Kiser, Ringeisen, & Schoenwald, 2001). Hyde, Falls, Morris, and Schoenwald (2010) highlighted two important issues that need to be addressed in the process of implementing evidence-based interventions: (1) whether the intervention works when it is implemented in usual care settings, and (2) ongoing monitoring of how well the intervention is being implemented. The second of these issues is the focus of this paper. We build on previous research that showed that second generation interventionists (trained in a “cascaded” train the trainer condition) achieved comparable child-level outcomes to those obtained by first generation interventionists who were trained by the evidence-based practice (EBP) developers in the context of a randomized clinical trial (Chamberlain, Price, Reid, & Landsverk, 2008). In this paper, a method for monitoring fidelity is described and preliminary findings on fidelity levels achieved by first and second generation interventionists implementing KEEP in real world (non-research) settings are examined. KEEP is a group-based 16 session intervention aimed at strengthening skills of foster and kinship parents serving children ages 4–16 in regular child welfare systems and has been shown to reduce child behavior problems and placement disruptions (Price et al., 2008).

Fidelity has been referred to as, “the demonstration that an experimental manipulation is conducted as planned” (Dumas, Lynch, Laughlin, Smith, & Prinz, 2001), and as incorporating the concepts of *adherence* to the intervention’s core content components and *competent* execution using proven clinical teaching practices (Forgatch, Patterson, & DeGarmo, 2005). Provision of in-depth training in the intervention model and goals, curriculum content, and training procedures is necessary for intervention fidelity, as is a properly supervised staff. However, neither training nor supervision alone is sufficient in ensuring that the intervention is conducted as planned (Dumas et al., 2001). In the current study, we include a dual focus on facilitator adherence to the KEEP core content components and to the competent process oriented delivery of the intervention. Videotaped recordings of the KEEP group sessions were rated for both coverage of key session components, and effective communication processes. Specifically, potential differences in levels of fidelity are examined between Generation 1 (G1) facilitators who were trained and supervised by KEEP developers, and Generation 2 (G2) facilitators who were trained and supervised by certified KEEP G1 facilitators. To examine potential differences in levels of fidelity by generation we used an equivalence testing design strategy.

### 1.1. Equivalence designs

Equivalence designs in medical and mental health have been gaining popularity in recent years as researchers seek to implement research-based interventions into real world practice settings and to examine less costly versions or means of conducting interventions that are practical to administer and that meet the same standards as existing treatments (see Eranti et al., 2007; Greene et al., 2010; Hermens et al., 2007; Lovell et al., 2006; Morland et al., 2010; O’Reilly et al., 2007). Equivalence designs have been used when the researchers seek to demonstrate that two interventions are equivalent on an outcome of interest (D’Agostino, Massaro, & Sullivan, 2003; Greene, Morland, Durkalski, & Frueh, 2008). Given equivalence, one intervention might be more appealing than another if it is more efficient.

### 1.2. The KEEP intervention

KEEP is a group-based parent management training (PMT) intervention for foster or kinship families. Parents receive 16 weeks of foster/kinship family support and training and supervision in behavior management methods. Intervention groups consist of 3 to 10 foster

parents and are conducted by a trained facilitator and co-facilitator team. The 90-minute sessions are structured so that the curriculum content is integrated into group discussions. The overall objective is to give parents effective tools for dealing with child externalizing and other behavioral and emotional problems and to support them in implementing those tools. Curriculum topics included framing the foster/kin parents' role as being key agents of change with opportunities to alter the life course trajectories of the children placed with them, and methods for encouraging child cooperation, for using behavioral contingencies, for using effective limit setting, and for balancing encouragement and limits. Sessions focus on dealing with difficult problem behaviors (including covert behaviors), promoting school success, encouraging positive peer relationships, and strategies for managing stress brought on by providing foster care. There is an emphasis on active learning methods; illustrations of primary concepts are presented via role-plays and videotapes. At the end of each meeting, a home practice assignment is given that relates to the topics covered during the session. The purpose of these assignments is to assist parents in specific ways to implement the behavioral procedures reviewed at the group meetings. The facilitator or co-facilitator telephone foster parents each week to trouble shoot any problems they have in implementing the assignment and to collect data on the child's problem behaviors during the past day. If foster parents miss a parent-training session, material from the missed session is delivered during a home visit at a time convenient for the foster/kinship parents.

### 1.3. Prior studies

The KEEP intervention is an outgrowth of the social learning-based parent management training (PMT) approach that has been shown to produce positive outcomes for the treatment and prevention of child and adolescent behavior problems in numerous randomized controlled trials conducted in Oregon and elsewhere (e.g., Chamberlain, Price, Leve, et al., 2008; Chamberlain, Price, Reid, & Landsverk, 2008; Kazdin, 1997; Kazdin & Weisz, 1998; Leve & Chamberlain, 2007; Leve, Chamberlain, & Reid, 2005). In addition, prior studies have demonstrated that fidelity of PMT interventions can be sustained at a high rate following scale-up (Forgatch & DeGarmo, 2007). In an initial efficacy trial conducted in Oregon foster families were randomly assigned to one of three groups: (a) enhanced services plus a monthly stipend, (b) a monthly stipend only, and (c) a foster-care-as-usual control group. Treatment for the enhanced groups was conducted by an experienced foster parent who was well versed in the OSLC PMT model and supervised by the KEEP developer (Chamberlain). Results showed decreased child behavior problems and increased placement stability in intervention homes (Chamberlain, Moreland, & Reid, 1992). Next, a second larger effectiveness study was conducted in the San Diego child welfare system in partnership with researchers at OSLC and the Child and Adolescent Services Research Center (CASRC, PI: Price). Seven hundred foster and kin parents caring for a 5- to 12-year-old child were randomly assigned to intervention (KEEP) or control (case work services as usual) conditions. In that study in addition to examining outcomes, a cascading dissemination model was tested. In that model, paraprofessional facilitators hired by CASRC were trained by OSLC developers, supervised weekly by the on-site supervisor and during weekly telephone calls by OSLC. These facilitators and the facilitators of the initial study in Oregon are considered generation 1 facilitators, or G1 (they were trained and supervised by developers). Next, the G1 San Diego facilitators trained and supervised a second cohort; facilitators trained by the G1s are considered generation 2, or G2. An OSLC clinical consultant supervised the G1's supervision of G2 interventionists, but had no direct contact with G2. The G1 and G2 group sessions were videotaped and the tapes were reviewed during supervision sessions. The results showed superior outcomes for children and parents in the KEEP condition and of relevance here, there were no differences in treatment effectiveness for participants receiving the intervention from G1 and G2 (Chamberlain, Price, Leve, et al., 2008; Chamberlain, Price, Reid, & Landsverk, 2008).

## 1.4. The current study

The purpose of the current report is to examine whether the fidelity observed in the context of the research trials could be generalized to new samples through the implementation of KEEP in community agency settings where the G1 and G2 group facilitators were not part of research teams, but rather conducting the KEEP intervention as part of routine agency care delivered to foster and kin parents. The appealing characteristics of the KEEP G2 condition include the potential for the agency to increase local capacity to train and supervise their own facilitators. This has obvious cost implications and potentially translates into the delivery of more services to children and foster/kin families. Training agency staff to self-monitor intervention fidelity is also appealing and might relate to program sustainment over time (Ory, Jordan, & Bazzairre, 2002). We view the results from the current study as preliminary because the analysis is a post hoc examination of real-world implementations of KEEP and therefore is subject to a number of limitations discussed later.

## 2. Materials and methods

### 2.1. Study participants and setting

Participants for the current study were 10 KEEP facilitators who were delivering the model in community-based non-research sites including generation 1 (G1) facilitators who were trained and supervised by the KEEP developers ( $n=6$ ) and generation 2 facilitators (G2) who were trained and supervised by G1 facilitators ( $n=4$ ) without any direct contact from the KEEP developers. Two of the six G1s went on to train and supervise G2 facilitators (Fig. 1). Participants video recorded 291 sessions from six sites all providing support services to foster/kin parents in child welfare. Sites agreed to have coded de-identified data included in this analysis. Groups run by G1 and G2 facilitators were comparable with respect to facilities (i.e., groups were run in community settings such as community centers, churches, and child welfare meeting rooms), size of groups (approximately 5–10 foster parents per group), and foster parent incentives (e.g., free childcare, snacks, and parking were provided).

### 2.2. Measure

The facilitator adherence rating (FAR, scale shown in Fig. 2) was used to code video recordings of sessions from G1 ( $n=155$ ) and G2 ( $n=136$ ) groups. The FAR is a 14-item scale shown in Fig. 2 that is used to rate the content, process and structure of the group. A Likert-type scale from 1 (not at all) to 5 (very much) is used for the ratings. The FAR is completed by the supervising consultant after each KEEP session. A score of 56 out of a possible 70 (or an average rating of 4 on all items) indicates that the facilitator has met the certification standard for fidelity for the session. Prior analysis of the FAR (Buchanan, Saldana, & Chamberlain, 2010) indicates that G1 facilitators typically earn a score closer to an average rating of 4.5 on all items,  $M=63.36$ ,  $SD=4.89$ . Data for the present analysis include 291 FAR scores (G1  $n=155$ , G2  $n=136$ ). The number of scores included for individual facilitators varies and is based on the number of cohorts of foster/kin parents they have served (Range= 1–4 cohorts) and sessions they have facilitated ( $M=15.44$ ,  $SD= 3.04$  sessions). Of the 27 cohorts included in this analysis, 67%, or approximately 11 of the 16 sessions, had associated FAR data. Missing data were due to technical difficulties including the site not being able to upload video or no sound associated with the video. Therefore, there was insufficient audible information to rate the session. FAR data appeared to be missing at random and occurred equally for any given session number.

### 2.3. Data sources

Sessions were video recorded using a laptop camera with AVCR software, an application that runs on the Adobe AIR framework. Using AVCR, sessions are uploaded over a secure

channel to the HIPPA-compliant KEEP website that utilizes ColdFusion 9 Standard software. Data points are stored in an SQL server and all entry points are encrypted. Consultants and G2 supervisors access video only for their sites, review the weekly session video, and complete the FAR scoring on the KEEP website. FAR scores are then used in the weekly consultation and supervision with the facilitator. All consultants (G1 and G2) are required to meet a standard of 80% reliability with trained senior G1 consultants.

## 2.4. Analysis

This study tests the hypothesis that the G2 facilitators working in community agencies, who were trained by G1 staff, would have the same fidelity as those trained and supervised by the KEEP developers (G1). We hypothesized that there would be no difference, clinically or statistically, between the G1 and G2 means, or,  $H_A: M_{G1} = M_{G2}$ ;  $H_0: M_{G1} \neq M_{G2}$ . Rejection of the null hypothesis ( $H_0$ ) is necessary to conclude that the two conditions are equivalent. Equivalence designs present the challenge of erroneously concluding that two treatments are equivalent when null results may be due to poor study design and execution (Greene et al., 2008; Jones, Jarvis, Lewis, & Ebbutt, 1996). Researchers can strengthen equivalence designs by (a) a prior specification of the equivalence margin, and (b) constructing a 90% confidence interval (Rogers, Howard, & Vessey, 1993; Westlake, 1981).

**2.4.1. A priori selection of the equivalence margin**—Determination of the equivalence margin is based both on prior studies indicating effectiveness as well as clinical judgment of the degree to which the new treatment can differ from the active control and still be considered equivalent (Greene et al., 2008; Jones et al., 1996; Piaggio, Elbourne, Altman, Pocock, & Evans, 2006). For KEEP, fidelity of implementation is considered to be established when FAR scores reach at least an average of 4 per item, or a total score of at least 56. To determine that there is a clinically non-significant difference (equivalence) between G1 and G2 conditions, an interval margin was set at  $\pm$  one standard deviation of the G1 mean score ( $SD$  adjusted for clustering effects=4.69). This requires the obtained confidence interval around the mean difference of the G1 and G2 scores fall within the range of  $-4.69$  to  $+4.69$ . A true difference in average fidelity for G1 and G2 smaller than 4.69 would likely not be a clinically meaningful difference.

**2.4.2. Selection of the appropriate statistical analysis**—Guidelines for equivalence designs suggest using a 90% confidence interval (90% C.I.) of the difference analysis approach to evaluate whether the new intervention is statistically significantly different from the active control (Rogers et al., 1993; Westlake, 1981). To determine equivalence, the obtained 90% C.I. for the new intervention must fall within the equivalence margin range of  $-4.69$  to  $+4.69$ . Multilevel modeling in Mplus was used to adjust for clustering of FAR scores, where Level 1 = facilitators clustered within generations (G1 and G2), and Level 2 = KEEP groups clustered within facilitators ( $N=10$ ). Multi-level models statistically account for clustering within groups by partitioning within- and between-group effects, which is relevant for designs with a small number of cases per level (Clarke, 2008). The present analysis is a post hoc analysis; therefore, a power analysis was not conducted prior to data collection. With traditional difference testing designs, the purpose is to accurately detect small differences between groups which typically require larger sample sizes. However, with equivalence testing designs, the purpose is to accurately detect larger differences so a smaller sample size is acceptable. A wider range of scores around the difference will be produced by small sample size (e.g.,  $N=10$  facilitators) for the 90% C. I. than would be produced by a larger sample size (Jackson, 2005). A wider range of scores is more likely to fall outside of the equivalence margin.



### 3. Results

To assess equivalence fidelity between G1 and G2, we conducted a two-sided 90% C.I. of the difference between groups using multi-level modeling in Mplus. G1 and G2 mean FAR ratings were adjusted to account for clustering ( $M_{G1}=62.84$ ,  $n_{G1}=155$  and  $M_{G2}=63.20$ ,  $n_{G2}=136$ ) with a mean difference of 0.36 between G1 and G2. Results indicate that the obtained 90% C. I. of 0.36 (-2.62-3.33) falls wholly within the a priori specified equivalence margin of 0+/-4.69. See Fig. 3 for an illustration of the equivalence margin and obtained 90% C. I. These results strongly suggest that the G1 and G2 means are clinically equivalent to one another.

On average, there were 25.8 group sessions per facilitator in G1 ( $ICC=0.30$ ) and 34.00 in G2 ( $ICC=0.17$ ). These results mean that G2 facilitators were rated on more sessions and there were only minimal differences between G1 and G2 ( $ICC$  difference=0.13).

### 4. Discussion

The mean difference between G1 and G2 scores is quite small (0.36) and the 90% C. I. was narrow (-2.62 to -3.33). An obtained 90% C. I. of this magnitude means that we are 90% confident that the true difference between G1 and G2 scores is between 0-2.97 points on the FAR. The narrow 90% C. I. is notable given the small sample size ( $N=10$  facilitators).

Because differences between the G1 and G2 mean scores are not clinically or statistically significantly different we can be reasonably confident that the G2 fidelity is at the same level as the G1 fidelity. This result is consistent with prior findings that clinical outcomes are comparable between G1 and G2 groups (see Chamberlain, Price, Reid, & Landsverk, 2008). It is reassuring that these results seem to indicate that the cascade strategy is producing facilitators who achieve the same acceptable level of fidelity as those who were trained and supervised by the KEEP developers.

#### 4.1. Limitations and next steps

The primary study limitations result from the practical realities of post hoc analyses including a modest amount of missing data, FAR ratings completed by different individuals at G1 than G2, a low number of consultants completing FAR ratings, and lack of within-study reliability coding of the FAR. Differences in FAR ratings for G1 and G2 were small. These differences might be due to a variety of facilitator factors including inconsistencies in quality or competence, slippage in model adherence, or difficulty of adherence for a particular group. Differences between G1 and G2 might also be due to rater effects including training and experience with the FAR or proximity of the rater to the facilitator (G1 ratings were completed by OSLC trained consultants and G2 ratings were completed by the facilitator's primary supervisor). Increasing the sample size of facilitators and the sample size of rated sessions would reduce the impact of a specific facilitator or KEEP group on the overall outcome. Theoretically, relevant covariates could be measured to further explain variation in consistency of ratings from G1 to G2, such as facilitator and consultant ratings of difficulty of the groups and facilitator stress related to the group.

#### 4.2. Conclusions

Given the limitations, we view this as a preliminary study. Nonetheless, the results show that the cascading dissemination model appears to be a feasible method for scaling up the KEEP intervention. If future, more rigorous tests of this model continue to support the maintenance of effects over generations of interventionists, the cascading dissemination model may have applicability to other evidence-based practices as they are implemented in routine care settings.

## Acknowledgments

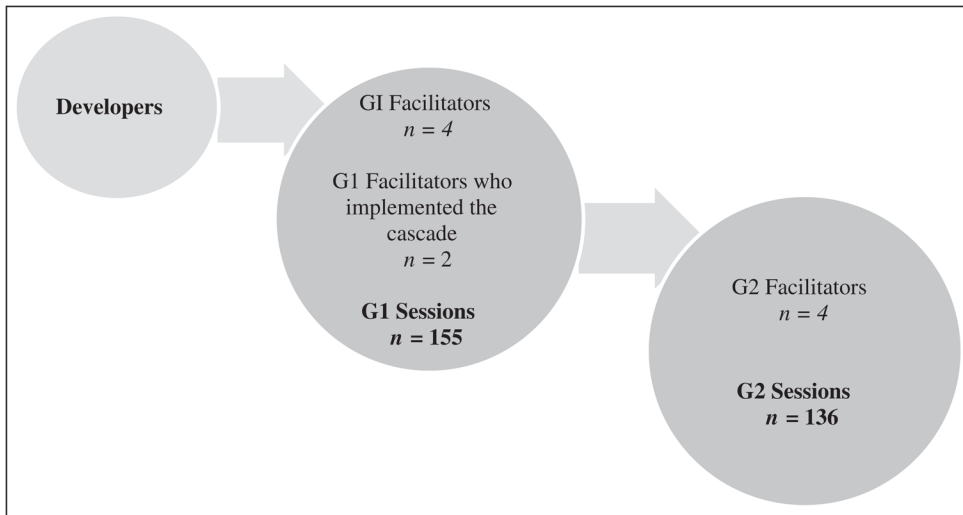
Support for this research was provided by Grant Nos. R01 MH 60195 from the Child and Adolescent Treatment and Preventive Intervention Research Branch, DSIR, NIMH, U.S. PHS; R01 DA 15208 and R01 DA 021272 from the Prevention Research Branch, NIDA, U.S. PHS; R01 MH 054257 from the Early Intervention and Epidemiology Branch, NIMH, U.S. PHS; and P20 DA 17592 from the Division of Epidemiology, Services and Prevention Branch, NIDA, U.S. PHS. We gratefully acknowledge the support of David DeGarmo and Michael Hurlburt for the analysis in this study.

## References

- Buchanan, R.; Saldana, L.; Chamberlain, P. KEEP scoring technical report: Facilitator adherence rating. Eugene, OR: Oregon Social Learning Center; 2010.
- Chamberlain P, Moreland S, Reid K. Enhanced services and stipends for foster parents: Effects on retention rates and outcomes for children. *Child Welfare*. 1992; 5:387–401. [PubMed: 1521497]
- Chamberlain P, Price J, Leve LD, Laurent H, Landsverk JA, Reid JB. Prevention of behavior problems for children in foster care: Outcomes and mediation effects. *Prevention Science*. 2008; 9(1):17–27. [PubMed: 18185995]
- Chamberlain P, Price J, Reid J, Landsverk J. Cascading implementation of a foster and kinship parent intervention. *Child Welfare*. 2008; 87(5):27–48. [PubMed: 19402358]
- Clarke P. When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. *Journal of Epidemiology and Community Health*. 2008; 62:752–758. [PubMed: 18621963]
- D'Agostino RB, Massaro JM, Sullivan LM. Non-inferiority trials: Design concepts and issues — The encounters of academic consultants in statistics. *Statistics in Medicine*. 2003; 22:169–186. [PubMed: 12520555]
- Dumas JE, Lynch AM, Laughlin JE, Smith EP, Prinz RJ. Promoting intervention fidelity: Conceptual issues, methods, and preliminary results from the EARLY ALLIANCE prevention trial. *American Journal of Preventive Medicine*. 2001; 20:38–47. [PubMed: 11146259]
- Eranti S, Mogg A, Pluck G, Landau S, Purvis R, Brown RG, et al. A randomized, controlled trial with 6-month follow-up of repetitive transcranial magnetic stimulation and electroconvulsive therapy for severe depression. *The American Journal of Psychiatry*. 2007; 164:73–81. [PubMed: 17202547]
- Forgatch MS, DeGarmo DS. Sustaining fidelity following the nationwide PMTO™ implementation in Norway. *Prevention Science*. 2007; 8(1):235–246.
- Forgatch MS, Patterson GR, DeGarmo DS. Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of parent management training. *Behavior Therapy*. 2005; 36:3–13. [PubMed: 16718302]
- Greene CJ, Morland LA, Durkalski VL, Frueh BC. Noninferiority and equivalence designs: Issues and implications for mental health research. *Journal of Traumatic Stress*. 2008; 21(5):433–439. [PubMed: 18956449]
- Greene CJ, Morland LA, Macdonald A, Frueh BC, Grubbs KM, Rosen CS. How does tele-mental health affect group therapy process? Secondary analysis of a noninferiority trial. *Journal of Counseling and Clinical Psychology*. 2010; 78(5):746–750.
- Hermens MLM, van Hout HPJ, Terluin B, Ader HJ, Pennix BWJH, van Marwijk HWJ, et al. Clinical effectiveness of usual care with or without antidepressant medication for primary care patients with minor or mild-major depression: A randomized equivalence trial. *Biomed Central Medicine*. 2007; 5(36)
- Hoagwood H, Burns BJ, Kiser L, Ringeisen H, Schoenwald SK. Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*. 2001; 52(9):1179–1189. [PubMed: 11533391]
- Hyde, PS.; Falls, K.; Morris, JA.; Schoenwald, SK. Turning knowledge into practice: A manual for human service administrators and practitioners about understanding implementing evidence-based practices. 2. Boston, MA: Technical Assistance Collaborative, Inc; 2010. (revised)
- Jackson, SL. *Statistics plain and simple*. Belmont, CA: Thomson Wadsworth; 2005.

- Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: The importance of rigorous methods. *British Medical Journal*. 1996;313.
- Kazdin AE. Parent management training: Evidence, outcomes, and issues. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1997; 36:10–18.
- Kazdin AE, Weisz JR. Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*. 1998; 66:19–36. [PubMed: 9489260]
- Leve LD, Chamberlain P. A randomized evaluation of multidimensional treatment foster care: Effects on school attendance and homework completion in juvenile justice girls. *Research on Social Work Practice*. 2007; 17:657–663. [PubMed: 18159224]
- Leve LD, Chamberlain P, Reid JB. Intervention outcomes for girls referred from juvenile justice: Effects on delinquency. *Journal of Consulting and Clinical Psychology*. 2005; 73:1181–1185. [PubMed: 16392991]
- Lovell, K.; Cox, D.; Haddock, G.; Jones, C.; Raines, D.; Garvey, R., et al. Telephone administered cognitive behavior therapy for treatment of obsessive compulsive disorder: Randomised controlled non-inferiority trial. *British Medical Journal*. 2006. <http://dx.doi.org/10.1136/bmj.38940.355602.80>
- Morland LA, Greene CJ, Rosen CS, Foy D, Reilly P, Shore J, et al. Tele-medicine for anger management therapy in a rural population of combat veterans with posttraumatic stress disorder: A randomized noninferiority trial. *The Journal of Clinical Psychiatry*. 2010; 71(7):855–863. [PubMed: 20122374]
- O'Reilly R, Bishop J, Maddox K, Hutchinson L, Fisman M, Takhar J. Is telepsychiatry equivalent to face-to-face psychiatry? Results from a randomized controlled equivalence trial. *Psychiatric Services*. 2007; 58(6):836–843. [PubMed: 17535945]
- Ory MG, Jordan PJ, Bazzairre T. The Behavior Change Consortium: Setting the stage for a new century of health behavior-change research. *Health Education Research*. 2002; 17(5):500–551. [PubMed: 12408195]
- Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJW. Reporting of noninferiority and equivalence randomized trials: An extension of the CONSORT statement. *Journal of the American Medical Association*. 2006; 295(10):1152–1161. [PubMed: 16522836]
- Price JM, Chamberlain P, Landsverk J, Reid JB, Leve LD, Laurent H. Effects of a foster parent training intervention on placement changes of children in foster care. *Child Maltreatment*. 2008; 13:64–75. [PubMed: 18174349]
- Rogers JL, Howard KI, Vessey JT. Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin*. 1993; 113(3):553–565. [PubMed: 8316613]
- Westlake WJ. Bioequivalence testing — A need to rethink (Reader response). *Biometrics*. 1981; 37:591–593.





**Fig. 1.** Cascaded dissemination model implemented as part of routine agency care.

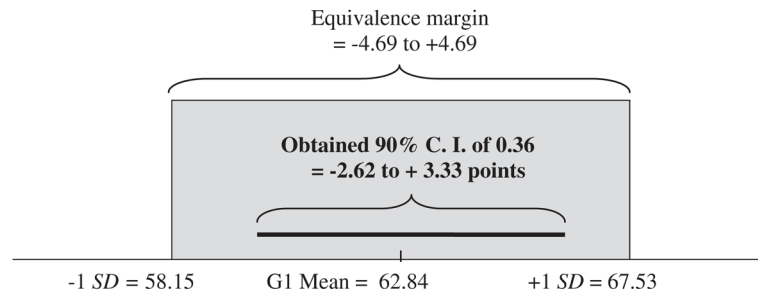
Program: \_\_\_\_\_  
 Date: \_\_\_\_/\_\_\_\_/\_\_\_\_  
 Group #/Name: \_\_\_\_\_  
 Week #: \_\_\_\_\_  
 Facilitator: \_\_\_\_\_  
 Consultant: \_\_\_\_\_  
 Session Date: \_\_\_\_/\_\_\_\_/\_\_\_\_  
 Review Date: \_\_\_\_/\_\_\_\_/\_\_\_\_

**KEEP: Facilitator Adherence Rating**

Please rate from 1-5 your impression of the following items.

		Not at All	Hardly at All	Some what	Pretty Much	Very Much
1.	Discussion included information on positive behaviors and youth progress	1	2	3	4	5
2.	Discussion included short and/or long term goals	1	2	3	4	5
3.	Facilitator taught material in the manual	1	2	3	4	5
4.	Foster parents were encouraged to/or used a behavioral management system (sticker charts, gem jars, etc.)	1	2	3	4	5
5.	Facilitator effectively used individual situations as teaching opportunities for the rest of the group	1	2	3	4	5
6.	Facilitator explained solutions in behavioral terms	1	2	3	4	5
<i>CONTENT Score</i> _____						
7.	The facilitator did not make critical comments	1	2	3	4	5
8.	Foster parents were engaged in discussions	1	2	3	4	5
9.	The atmosphere of the group was friendly and supportive	1	2	3	4	5
10.	Facilitator redirected conversation when necessary	1	2	3	4	5
<i>PROCESS Score</i> _____						
11.	Logistics were well managed (incentives, snacks, child care, etc.)	1	2	3	4	5
12.	Facilitator managed the meeting time well	1	2	3	4	5
13.	Facilitator greeted the foster parents warmly	1	2	3	4	5
14.	Facilitator ended group well (on time, encouraging statements, home practice)	1	2	3	4	5
<i>STRUCTURE Score</i> _____						
<b>TOTAL Score</b> _____						

**Fig. 2.**  
Facilitator adherence rating form.



**Fig. 3.** The obtained 90% C. I. of the difference between G1 and G2 falls within the equivalence margin of +/- 1.0 SD of the G1 mean.