

Evaluating Fidelity: Predictive Validity for a Measure of Competent Adherence to the Oregon Model of Parent Management Training

Marion S. Forgatch, Gerald R. Patterson, David S. DeGarmo

Oregon Social Learning Center

When efficacious interventions are implemented in real-world conditions, it is important to evaluate whether or not the programs are practiced as intended. This article presents the Fidelity of Implementation Rating System (FIMP), an observation-based measure assessing competent adherence to the Oregon model of Parent Management Training (PMTO). FIMP evaluates 5 dimensions of competent adherence to PMTO (i.e., knowledge, structure, teaching skill, clinical skill, and overall effectiveness) specified in the intervention model. Predictive validity for FIMP was evaluated with a subsample of stepfamilies participating in a preventive PMTO intervention. As hypothesized, high FIMP ratings predicted change in observed parenting practices from baseline to 12 months. The rigor and scope of adherence measures are discussed.

THE LAST TWO DECADES stand witness to what has essentially been a tectonic shift in studies of intervention processes. The change has involved establishing an empirical base for therapy processes and outcomes. For example, Pinsof (1981) carefully reviewed existing coding systems that sampled psychodynamic, psychoanalytic, and client-centered therapies to study therapy process. Analogous empirical developments within Parent Management Training (PMT) came somewhat later with an observational system describing therapists' behaviors and parent reactions to the therapists, including cooperative and resistant behaviors (Chamberlain & Ray, 1988). Better understanding

Support for this project was provided by Grants R01 MH38318 and R01 MH54703 from the Child and Adolescent Treatment and Preventive Intervention Research Branch, DSIR, NIMH, U.S. PHS; Grant R01 DA16097 from the Prevention Research Branch, NIDA, U.S. PHS; and Grant P30 MH 46690 from the Prevention and Behavioral Medicine Research Branch, Division of Epidemiology and Services Research, NIMH and ORMH, U.S., PHS.

Address correspondence to Marion S. Forgatch, Senior Scientist, Oregon Social Learning Center, 160 E. 4th Avenue, Eugene, OR 97401-2426; e-mail: marionf@oslc.org.

BEHAVIOR THERAPY 36, 3-13, 2005

005-7894/05/0003-0013\$1.00/0

Copyright 2005 by Association for Advancement of Behavior Therapy

All rights for reproduction in any form reserved.

of therapist actions that enhance and impede behavior change has led to increased attention to the clinical and teaching skills necessary for effective interventions.

Advances in the study of intervention processes have been accompanied by an explosion of outcome studies that employ randomized trials, particularly given the call for replication across sites (Chambless & Hollon, 1998). The need to replicate has led to a proliferation of manuals that specify intervention components and procedures. Using manuals, however, does not guarantee competent application of a method. Intervention delivery must be evaluated for fidelity to the program content and processes or one cannot explain whether failure to replicate is a problem with the program or with its application.

This article introduces a measure of fidelity to the Oregon Model of Parent Management Training (PMTO), a theory-based intervention developed by the group at Oregon Social Learning Center (OSLC; e.g., Forgatch, Bullock, & Patterson, 2004). We use the word fidelity to incorporate two concepts: *adherence* to the intervention's core content components and *competent* execution using accomplished clinical and teaching practices. We see adherence as limited to following procedures while covering specific topics, which can be readily assessed with a checklist. Competent adherence to PMTO requires that the procedures be carried out with sophisticated clinical and teaching skills that promote behavior change. The measure, Fidelity of Implementation (FIMP; Knutson, Forgatch, & Rains, 2003), employs direct observations from videotapes of intervention sessions in which core components were covered.

The PMTO approach is based on social interaction learning (SIL), a developmental model that defines mechanisms of etiology, maintenance, and change as a set of core parenting practices that directly affect child outcomes (Reid, Patterson, & Snyder, 2002). The SIL model specifies that harsh contextual factors (e.g., family structure transitions, parental maladjustment) have indirect effects on child outcomes and are mediated by coercive processes and ineffective parenting skills. Coercive and

ineffective parenting practices are assumed to flourish in stressful contexts. The hallmark of PMTO interventions, thus, is a focus on enhancing effective parenting and diminishing coercive practices while making relevant adaptations for contextual factors. The main features of coercion are negative reciprocity, escalation, and negative reinforcement. These coercive components are controlled in turn with five core effective parenting skills. *Skill encouragement* promotes prosocial development through scaffolding techniques (e.g., breaking behavior into small steps, prompting appropriate behavior) and contingent positive reinforcement (e.g., praise and incentives). *Discipline* (i.e., limit setting) decreases deviant behavior through the appropriate and contingent use of mild sanctions (e.g., time-out, privilege removal). *Monitoring* (i.e., supervision) involves tracking youngsters' activities, associates, whereabouts, and transportation. *Problem-solving* skills help families negotiate disagreements, establish rules, and specify consequences for following or violating rules. *Positive involvement* reflects the many ways parents provide their youngsters with loving attention.

OSLC interventions have been tailored for clinical problems and prevention designs associated with children's antisocial behavior. Several manuals detail related procedures for various contexts or focal populations. These include treatment foster care (Chamberlain, 1994); substance abuse (Dishion & Kavanagh, 2003); single-mother families (Forgatch, 1994); stepfamilies (Forgatch & Rains, 1997); and risk for conduct disorder (Ramsey, Antoine, Kavanagh, & Reid, 1992a, 1992b). The common thread in these programs is replacing coercive family process with the five effective parenting practices described earlier. Several studies have demonstrated that intervention effects on these parenting practices produce benefits to child behavior problems, including aggression, delinquency, deviant peer association, internalizing problems, school problems, and growth in substance abuse (DeGarmo, Eddy, Reid, & Fetrow, submitted; DeGarmo & Forgatch, in press; DeGarmo & Forgatch, 2004; DeGarmo, Patterson, & Forgatch, 2004; Eddy & Chamberlain, 2000; Forgatch & DeGarmo, 1999; Martinez & Forgatch, 2001). Because PMTO is strongly theory based, rigid adherence to manuals is not required as long as the core components are taught in a skillful manner that brings about change. Assessing PMTO fidelity, thus, requires evaluating both content and process.

A state-of-the-art design for evaluating intervention adherence is suggested by Hogue, Liddle, and Rowe (1996). They offer this view of the field:

A primary strategy within the category of observational adherence methods entails training non-participants to recognize a roster of intervention techniques and then to code videotapes of randomly selected sessions. Raters are typically kept blind to therapist identity, session number. . . . This strategy introduces the highest levels of rigor and reward in adherence research." (p. 334)

The large-scale prevention study now under way by Dumas and colleagues (2001) could be an exemplar of this approach. Their prevention trials took place in three different settings. Audio- or videotapes were collected in each setting and coded for adherence. Data are not yet available but are designed to test the hypothesized relation between adherence and effective outcome. Another modern program that emphasizes the importance of collecting the more expensive observation data is the Winning New Jobs (WNJ) program. In that study, observers sampled 40% of the training sessions and included evaluation of process as well as content (Price & Vinokur, 1995).

DeRubeis and Feeley (1990) suggested that adherence should be significantly related to treatment outcome. Henggeler and colleagues (1997) used a measure of adherence to MultiSystemic Therapy (MST) to explain failure to produce expected outcomes in a randomized treatment study with juvenile offenders. They constructed 15 rating scales from 26 items that assessed general events (e.g., "The therapist tried to change some ways that family members interact with each other," "There were awkward silences and pauses during the session"). Therapists, adolescents, and parents made separate ratings and factor analyses were carried out for each set of raters. The factor scores by agent were correlated separately with 7 outcome measures (e.g., self-reported delinquency, emotional adjustment, incarceration). Although 8 of the 105 correlations were significant in the expected direction, 5 were significant in the wrong direction, making it questionable that the findings exceeded chance expectations. Thus, that intensive effort failed to demonstrate that treatment adherence produced successful outcomes. In the present study, we expect the fidelity construct to predict improvement in parenting practices for mothers and stepfathers.

The FIMP rating system (Knutson et al., 2003) evaluates content of core components as well as competent clinical and teaching processes observed during intervention. FIMP incorporates objectivity through the use of observation of videotaped sessions that deal with core content material. In the present pilot study, we evaluated the predictive

validity of FIMP and hypothesized that FIMP ratings would be significantly associated with change in parenting practices from baseline to 12 months, after intervention completion. The study drew from families participating in Marriage and Parenting for Stepfamilies (MAPS), a preventive PMTO intervention for recently married stepfamilies (Forgatch & Rains, 1997).

Methods

PARTICIPANTS

A subsample of 20 families was randomly drawn from the experimental group of families participating in a preventive intervention trial for stepfamilies. In the full sample, 110 stepfamilies were recruited from a metropolitan area in the Pacific Northwest through media advertisements, with 67 experimental and 43 control families. From the 53 families that had completed intervention at the time of the study, we randomly selected 20 families, 5 each from the caseload of the four MAPS interventionists. In this subsample, stepfamily couples had been married an average of 17.8 months ($SD = 13.4$) and resided with a participating focal child who was the mother's biological child. Children ranged in age from 5.8 to 9.5 (mean = 7.7) and 75% were boys. The sample was originally restricted to boys because they are more likely than girls to exhibit early problems after divorce and remarriage (Hetherington & Clingempeel, 1992). Girls were later added to the sampling criteria to complete the recruitment schedule.

In terms of socioeconomic characteristics, most mothers and stepfathers had obtained on average 2 years of education or training beyond high school at study entrance. None of the mothers and 5% of the stepfathers had a GED or less than high school education. Thirty-five percent of the mothers and 25% of the stepfathers had completed high school with no further training or education. For mothers and stepfathers respectively, 5% and 5% were not working outside the home, 20% and 15% were in unskilled occupations, 5% and 30% in skilled worker/tenant farmer/small business jobs, 35% and 10% in clerical or sales occupations, 15% and 10% were in semiprofessional occupations, and 15% and 5% were medium business owners or in administrative/professional positions. The combined average gross annual income for each household reported by the mother and stepfather was \$41,500 ($SD = \$23,015$). The average age of the mothers and stepfathers in this sample was 33.9 ($SD = 7.0$) and 35.1 ($SD = 8.0$), respectively. For mothers, stepfathers, and children respectively, 89.1%, 91.8%, 88.2% were of European-American descent; 2.7%, 3.6%, 1.8% were Latina/o, .9%,

.9%, 0% were Native American, 0%, .9%, 0% were African American, 0%, 1.8%, 0% were Asian American, and 7.3% .9%, 10% were multiracial or other.

DESIGN

The study employed a randomized experimental longitudinal design with 61% and 39% of families assigned to experimental and control groups, respectively. The groups were assigned unequally to provide sufficient sample size within the experimental group to examine potential full-implementation effects of the intervention (Vinokur, van Ryn, Gramlich, & Price, 1991). Prior to randomization, parents were told about the study, the intervention, and that intervention participation would be randomly determined. Families in the control condition received no intervention but were given referral support if requested. Families in both conditions were assessed on the same time line: baseline, 6, 12, and 24 months. Participants were paid \$10/hour. Parenting measures are from baseline and 12-month assessments and intervention fidelity measures are from the period following baseline and prior to 12 months.

ATTRITION

For the overall group design study, we retained 91% of the sample at the 12-month follow-up, with no differential attrition for the experimental or control conditions. For the subsample, 100% were retained.

INTERVENTION PARTICIPATION

The mean number of sessions attended was 14.3 ($SD = 2.6$). The average duration until termination was 31.8 weeks ($SD = 3.3$), more than twice as many weeks as sessions.

INTERVENTION

The MAPS intervention, fully described in the manual *Marriage and Parenting in Stepfamilies* (Forgatch & Rains, 1997), was designed to promote healthy family and marital functioning and to prevent child adjustment problems. The MAPS manual contains 13 session agendas with information for interventionists and materials for parents. Each session is detailed with agenda, objectives and rationales, procedures, exercises and role-plays, and process suggestions. Parent materials include summaries of principles, home practice assignments, charts, and other necessary resources. The interventionists were flexible in their timing and application of materials, thus families progressed through the program components at differing rates. General procedures in PMTO and MAPS call for

debriefing home practice assignments, introducing new topics, rehearsing the new strategies through role-play and other exercises until parents are well practiced, applying the procedures at home, troubleshooting problems, and tailoring procedures to specific family needs and values. Program components are introduced in an integrated fashion, with skills gradually building one upon another.

MAPS is a PMTO intervention with emphasis on the five core parenting practices (i.e., discipline, skill encouragement, monitoring, problem-solving, and positive involvement). MAPS has an enhancement component with two sessions specifically designed to enrich the marital relationship in newly constituted stepfamilies (e.g., debunking common stepfamily myths, presenting a united parenting front, the role of the stepparent). Additionally, each of the core parenting components is framed to deal with common problems stepfamilies have in applying the parenting principles. For example, the contingencies (positive and negative) for youngsters are the same in both programs; who applies the consequences and how they do it may differ. A common example has to do with discipline confrontations. As in standard PMTO, couples learn to work together as a team, but in MAPS we recommend that the biological parent take the lead in setting consequences and the stepparent support the biological parent. If couples agree to this strategy, they are given practice and coaching through role-play.

In the present study, we focused on two sessions for each type of contingency because these components are the sine qua non of PMTO. We FIMP scored samples of sessions on skill encouragement and discipline if they contained significant discussion of the topic either through introduction of new material or troubleshooting problems in its application. Skill encouragement was typically introduced in the first third of the intervention and discipline in the middle third. Skill encouragement includes discussions, exercises, role-plays, games, or other activities relevant to teaching and improving behavior by establishing environments that set up children to be successful and providing contingent positive response. Positive response includes social reinforcement, material or instrumental incentives, and incentive programs. Discipline includes discussions, role-play, exercises, games, and other activities to encourage use of contingent penalties to reduce problematic, undesirable, or otherwise negative behavior. Examples of strategies include time-out, work chores, fines, and privilege removal, as well as repetitive or boring events like writing sentences or copying words from the dictionary. These strategies may be referred to as punishment, discipline, limit setting, sanctions, and negative consequences.

FIDELITY OF IMPLEMENTATION

The FIMP rating system was based on the two prior OSLC observational systems designed to evaluate the intervention process: Therapist Performance Observational System (TPOS; Reid et al., 1979) and the Therapy Process Code (TPC; Chamberlain et al., 1986). In FIMP, we incorporated information from studies using the earlier systems (Chamberlain et al., 1984; Chamberlain & Ray, 1988; Patterson, 1988; Patterson & Chamberlain, 1988; Patterson & Forgatch, 1985; Stoolmiller, Duncan, Bank, & Patterson, 1993). The five FIMP dimensions were developed in an iterative process linking evaluation of an interventionist's competent adherence to theoretical PMTO core components and processes. We used a foundation of behavioral categories that proved relevant in previous studies, eliminated or combined categories with little empirical value, and expanded categories that required better definition. Next we attempted to apply the categories using videotapes and transcripts of family intervention sessions, which made it possible to parse categories into appropriate units. Then we subsumed discrete behaviors into global dimensions of competent adherence. Fidelity ratings took into account factors contributed by the family, evaluating the interventionist's PMTO knowledge and delivery in the context of a family's needs, characteristics, and reactions.

Training coders to reliability required approximately 40 hours. Trainees memorized the coding manual, viewed and scored videotapes, and discussed disagreements. At the conclusion of training, the reliability criterion required coders to achieve an intra-class correlation (ICC) of at least 70% on scores for tapes from three different families. During the study, 10-minute segments were sampled from each of two videotaped sessions per family on the topics of skill encouragement and discipline. To identify meaningful segments for rating, an assistant spot checks tapes labeled with the relevant topic, looking for a meaningful segment of approximately 10 minutes with appropriate content (e.g., debriefing home practice on the component, role plays; brainstorming for possible incentives or negative consequences; adjusting size of steps).

The session could be either introduction or troubleshooting. The 40 sessions were independently scored by calibrator coders (two trained PMTO interventionists) and by a reliability coder (the intervention supervisor). The reliability coder scored 100% of the samples; the calibrators each scored approximately 50% of the sample. Coders did not score their own sessions.

The FIMP manual details procedures for scoring videotapes, defines each dimension, and provides

necessary forms. Five dimensions are rated: PMTO knowledge, structure, teaching skills, clinical skills, and overall quality. For each FIMP dimension, the manual provides a general definition, description of key features, rating guidelines, and rating examples. Each dimension uses a 9-point scale, in which 1–3 reflects unacceptable performance (“needs work”), 4–6 is “acceptable,” and 7–9 is “good work.” The rating sheet also includes prompts for relevant dimensions of each category (Figure 1). In addition to familiarity with the manual, FIMP raters are knowledgeable with intervention manuals and training handbooks. The FIMP manual is available upon request. Brief definitions of the five dimensions follow.

Knowledge of PMTO. Knowledge reflects demonstration of proficiency in PMTO and SIL principles. Given their central role in the program, parenting practices are spelled out in detail. Ratings incorporate the interventionists’ apparent understanding of the theoretical model.

Structure. This category refers to skill at balancing several activities during the course of a session. Aspects include following an agenda, maintaining an orderly flow, leading without dominating, being responsive to family issues, and using sensitive timing and pacing.

Teaching. The earlier TPC defined *Teach* as a broad category, which was associated with increased resistance to direction provided by the therapist (Patterson & Forgatch, 1985). We separated the FIMP definition into verbal and active teaching activities. Verbal teach includes standard pedagogical tactics (give information, make suggestions, provide rationale); active teach engages the family in the learning process by brainstorming, prompting, engaging in role-plays, asking questions that elicit solutions, and other interactive approaches. Teaching involves demonstrating proficiency in strategies and tools to promote parents’ mastery and independent use of PMTO practices. Overreliance on verbal teaching can result in lower scores.

Clinical process. The importance of clinical process skills is well known to professionals. Behavioral approaches seldom specify clinical skills as a necessary component of the intervention. We differentiate between teaching and clinical process according to their aims. The goal of teaching is to impart information so that parents master new skills. A key goal of clinical process is to provide support necessary to create a safe and supportive context for learning. The skills can include questioning process that leads to openness, maintaining appropriate balance, encouraging skill development, and joining the family’s storyline.

Overall quality. This category requires the rater

to draw on a gestalt of information. Key components are knowledge of and performance with PMTO components and accomplishment of session goals. Other dimensions include the likelihood that the family can and will use the procedures, the family’s apparent satisfaction, their likelihood of continuing intervention, difficulty of the family, unique aspects of the situation, and sensitivity of the issue.

PARENTING MEASURES

Observational assessments of parent-child interactions were conducted at baseline and 12 months. The laboratory observations included a set of structured interaction tasks totaling 48 minutes in which the child interacted with mother alone, stepfather alone, and mother and stepfather together. The activities included: four 7-minute problem-solving discussions (1 alone with each parent and 2 with both parents) about current hot conflicts (e.g., chores, school problems, behavior), a 10-minute teaching task in which the mother assisted her child with a series of academic problems one grade level above current grade, a 5-minute cooperation/play task with both parents, and a 5-minute period to share refreshments with both parents.

Trained observers scored the family interactions using the Family and Peer Process Code (FPP; Stubbs, Capaldi, Forgatch, & Crosby, 1998). The FPP is a real-time system containing codes for discrete positive, negative, and neutral behaviors, which provides information on initiator and recipient, sequence, content, affect, and duration. Following the microsocial scoring, coders make ratings using a global rating system (Forgatch, Knutson, & Mayne, 1992).

Coder training in the FPP code required approximately 16 to 20 weeks at 20 hours per week. The reliability criterion at the conclusion of training required scoring two tapes in a row with 75% event-by-event agreement and a Cohen’s kappa coefficient of at least .65. Approximately 15% of the interactions were randomly selected for blind reliability checks. The intercoder reliability was consistent with other observational studies. The average kappa coefficient of coder agreement was .74 for content behaviors and .69 for affect.

To evaluate change over time, mean construct scores were created comprising multiple indicators using methods outlined by Stoolmiller and colleagues (Stoolmiller & Bank, 1995; Stoolmiller, Duncan, & Patterson, 1995) for modeling growth of construct scores. Each indicator was rescaled on a scale from 0 to 1 and then averaged. For microsocial observational scores (e.g., frequency counts) that were not rated on a fixed Likert-type scale

	9	8	7	6	5	4	3	2	1
	PMTO Knowledge								
<ul style="list-style-type: none"> - Applies principles and model - Understands core/supporting parenting practices 									1
<ul style="list-style-type: none"> - Follows an agenda - Includes appropriate sections - Manages orderly flow - Gives appropriate attention to relevant dimensions 	<ul style="list-style-type: none"> - Is responsive to family - Maintains leadership - Leads without dominating 								1
	Structure								
									1
									2
									3
									4
									5
									6
									7
									8
									9
	Teaching								
	Uses variety of activities								
<ul style="list-style-type: none"> - Gives information/instructions - Provides rationales - Uses good PMTO raps - Pinpoints 	<ul style="list-style-type: none"> - Balances verbal teach/active teaching - Elicits goal behavior - Engages family - Integrates/weaves information - Reviews material - Assesses skills and fills in gaps - Pre-corrects - Breaks into teachable units - Balances encouragement/correction - Predicts problems - Provides enough information - Adjusts PMTO tools to family - Punctuates 								1
	Use of role play								
									1
									2
									3
									4
									5
									6
									7
									8
									9
	Process Skills								
	Uses variety of tools								
<ul style="list-style-type: none"> - Prevents/manages resistance - Prevents/manages conflict - Maintains balance - Promotes united approach - Focuses on encouragement/support - Connects with storyline 	<ul style="list-style-type: none"> - Normalizes - Interpret/reframes - Metaphors - Mirrors/matches - Strategic warning - Takes responsibility - Paraphrases/summarizes - Humor - Paradox - Reflects - Supportive interrupts - Keeps contact - Movement 								1
	Use of role play								
									1
									2
									3
									4
									5
									6
									7
									8
									9
	Overall Quality								
<ul style="list-style-type: none"> - Demonstrates knowledge of PMTO - Performance on PMTO dimensions 	<ul style="list-style-type: none"> - Accomplishment of goals - Likelihood family will use 								1
									2
									3
									4
									5
									6
									7
									8
									9

FIGURE 1 FIMP Dimensions Rating Form. From "Fidelity of Implementation Rating System: The training manual for PMTO" by N. M. Knutson, M. S. Forgatch, and L. A. Rains, 2003, p. 25. Copyright 2003 by Oregon Social Learning Center. Reprinted with permission of the authors.

with known boundaries, across wave upper-boundaries determined scaling for computing growth scores. Constructs were required to obtain an eigenvalue of greater than 1 for each wave across time and a one-factor solution using principle components analyses.

The parenting measures are theoretically based, have demonstrated sensitivity to change in PMTO, and have shown convergent and predictive validity in at-risk and divorce samples (Forgatch & DeGarmo, 1999). For mothers, the parenting construct consisted of five coder-rated subscales: positive involvement, skill encouragement, problem-solving outcomes, monitoring, and inept discipline. The stepfathers' parenting construct contained the same subscales as mothers with the exception of skill encouragement, which was coded from the mother-child teaching task. No comparable task for skill encouragement was conducted for stepfathers.

Positive involvement was a mean scale score of coder ratings following each task. The seven indicators included 38 Likert-scale items (e.g., showed empathy, warmth, support, genuine concern; was accepting, affectionate). Cronbach's alphas for mothers and stepfathers respectively were .88 and .89 at baseline; .83 and .89 at 12 months.

Skill encouragement was a mean scale score based on global coder ratings following the 10-minute teaching task. Eleven items assessed promotion of skill development through contingent encouragement and scaffolding strategies (e.g., breaks task into manageable steps, reinforces success, prompts appropriate behavior, corrects in a nonaversive way). Cronbach's alphas were .80 for mothers at baseline and .81 at 12 months.

Problem solving was a scale score of coder ratings made following each of the three problem-solving interactions about a parent-identified issue. Nine Likert-scale items were averaged (e.g., solution quality, extent of resolution, likelihood of follow through, apparent satisfaction). Cronbach's alphas for mothers and stepfathers respectively were .91 and .94 at baseline; .94 and .94 at 12 months.

Monitoring was a scale score from two agents' reports. Parent interviewers provided Likert-scale ratings on three items: supervision during assessment, tracking outside of lab, and skillful at obtaining information. Coders rated two items: apparent knowledge of child's activities and tolerance of negative behavior. Cronbach's alphas for mothers and stepfathers respectively were .70 and .62 at baseline; .59 and .73 at 12 months.

Inept discipline was a scale score based on coder impressions. The scale consisted of 10 items (e.g., overly strict, authoritarian, oppressive; erratic, inconsistent, haphazard; threatens unlikely disci-

plines; nags to get compliance). Cronbach's alphas were .83 and .82 at baseline for mothers and stepfathers, respectively; .86 and .87 at 12 months.

Results

The main goal of the analyses was to evaluate the psychometric properties of the FIMP scores and to evaluate their predictive validity. We also assessed whether the subsample of families randomly selected for this study systematically differed from the nonselected families. Further, we needed to assess whether there were any systematic biases associated with the FIMP coders and the FIMP scores. Therefore, the analyses were conducted in three basic stages. First we conducted independent sample *t* tests comparing the FIMP scored and non-FIMP scored families. Second, we conducted exploratory factor analyses of the FIMP items. Third, we employed structural equation modeling (SEM) to conduct a simultaneous confirmatory factor analysis of FIMP and test the predictive validity of FIMP scores in a path analysis.

PRELIMINARY ANALYSES

We compared parenting scores for families within the experimental group between those randomly selected for FIMP scoring and those not FIMP scored. There were no baseline differences between the two groups for mothers' effective parenting or for stepfathers' effective parenting. Significant differences emerged at the 12-month follow-up. Non-FIMP-scored mothers displayed better parenting practices than did those who were FIMP scored ($M = .85, SD = .07$ and $M = .79, SD = .07, p < .05$, respectively). Non-FIMP-scored stepfather parenting scores were also higher at the 12-month follow-up compared to FIMP-scored families ($M = .82, SD = .08$ and $M = .73, SD = .09, p < .01$, respectively). Change scores indicated that over time the non-FIMP-coded families improved on the average in comparison with the FIMP-coded mothers and stepfathers. The mean score for change in effective parenting for non-FIMP-coded mothers was .08 ($SD = .12$) and for FIMP-coded mothers was .00 ($SD = .06, p < .05$); for stepfathers, $M = .08, SD = .10$ and $M = -.00, SD = .11, p < .05$, respectively).

FIMP ANALYSES

Each session produced a 5-item scale including *knowledge of PMTO, structuring of the session, teaching skill, clinical skill, and overall quality*. Using principal components factor analysis, one-factor solutions within encouragement and discipline produced eigenvalues of 3.3 and 4.3, explaining 66% and 85% of the variance for the respective

scales. Across topics, a one-factor solution produced an eigenvalue of 5.3, explaining 53% of the variance; however, two factors were extracted with an eigenvalue greater than 1.0 corresponding to encouragement and discipline subscales. Therefore, separate indicators were retained for the fidelity evaluation. Cronbach's alpha reliability for calibrators was .87 for encouragement and .95 for discipline. Alpha was .89 and .95, respectively, for the reliability coder. The intraclass correlation (ICC) coefficient between the calibrators and the reliability coder was .71 for encouragement, .93 for discipline, and .90 for a mean of the two scores.

The basic fidelity hypothesis was tested using multivariate analyses within SEM path models using the AMOS program (Arbuckle, 1997). An advantage to using SEM was the ability to control for within-couple dependence on the outcome by (a) freeing the covariance among the residuals for mothers and stepfathers, and (b) controlling for cross-lagged baseline paths. The model specified was essentially a seemingly unrelated least squares regression (see Greene, 1993). To test the hypothesis that higher fidelity would be associated with better parenting outcomes within the experimental group, a latent fidelity factor was modeled predicting change in observed parenting from baseline to 12 months within the experimental condition. The model included control variables of child age and child gender. Each change score controlled for initial status. The results are presented in Figure 2 in the form of standardized beta paths. For visual

clarity, paths from the control predictors were not displayed.

The dependent variables were entered as change scores regressed on their initial status. Thus, the effects of baseline predictors do not represent an auto-regressive or stability path, but rather the true effect of initial status on change (Kessler & Greenberg, 1981). Results of the model indicated that the main hypothesis was supported. Controlling for child age and gender, therapist fidelity was associated with increases in mothers' effective parenting ($\beta = .51, p < .05$) and stepfathers' effective parenting ($\beta = .49, p < .05$), independent of their initial levels of parenting. The model obtained good fit to the data, $\chi^2(12) = 9.11, p = .69, CFI = 1.00$. The model also indicated that fidelity alone accounted for 30% of the variance in both mother's parenting and stepfather's parenting change. The number of variables in this SEM exceeds a recommended number of at least 5 cases per variable for basic regression models (e.g., Tabachnik & Fidell, 1983). Thus, we replicated results with a 4-variable model using 2 fidelity indicators and 2 change scores, as well as a 3-variable system using one factor score and two change scores. Those path models are within the recommended number of cases per variable for OLS regressions.

On a related note, the sample size severely limits the number of relevant multivariate tests that can be conducted, such as testing for clustering of families within interventionists. Therefore, tests for individual interventionist effects were excluded from

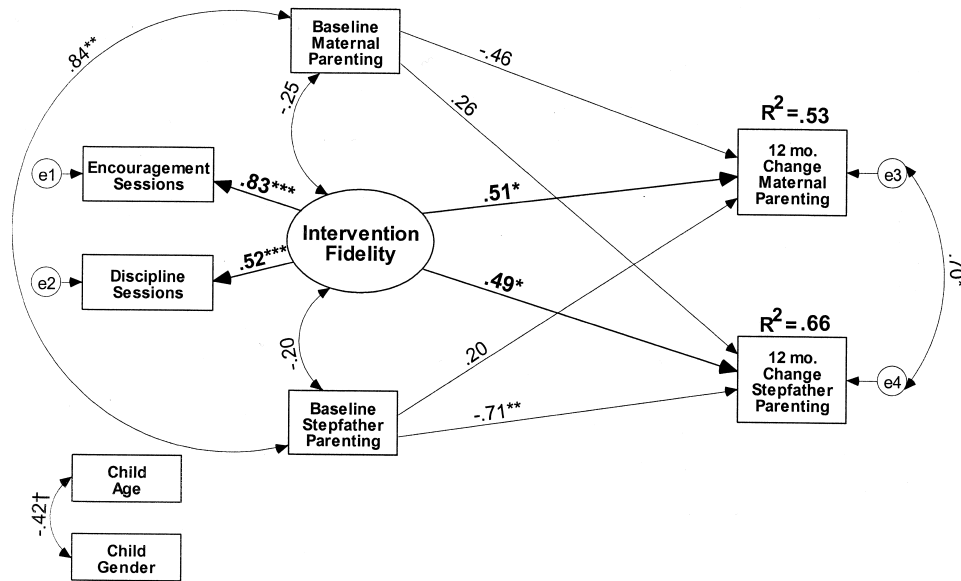


FIGURE 2 Structural equation path model testing effects of interventionist fidelity on change in effective parenting practices. Standardized coefficients presented, $\chi^2(12) = 9.11, p = .69, CFI = 1.00, n = 20, \dagger p < .10; *p < .05, **p < .01, ***p < .001$.

the above model. However, dummy coded variables were entered in an alternative model (not shown) that indicated interventionist contrast effects for the fidelity model presented in Figure 2. The fidelity parameters remained virtually unchanged, indicating that effects associated with family assignment or variance due to interventionist skill did not account for the effects of fidelity on change in parenting.

Discussion

In this study, we evaluated the predictive validity of FIMP, a measure of competent adherence to PMTO, a theoretically based intervention. The SIL model underlying the intervention specifies that five core parenting practices function as mechanisms for child outcomes. Prior studies have shown that a construct comprising observed measures of each of these parenting practices is sensitive to change through PMTO intervention and that this change in parenting produces commensurate benefits to child outcomes (DeGarmo et al., 2004; Forgatch & DeGarmo, 1999; Martinez & Forgatch, 2001). In the present study, we used FIMP to score videotaped sessions from a PMTO preventive intervention with stepfamilies and evaluated FIMP ratings as predictors of change in parenting practices measured separately for mothers and stepfathers. Observational methodology was employed to assess competent adherence to PMTO and change in parenting practices. To score FIMP, we sampled 10-minute segments from videotapes of sessions on the topic of skill encouragement and on discipline. To score parenting practices, videotapes of parent-child interactions in the laboratory were coded at baseline and 12 months later. Independent coders using distinct coding systems scored the intervention sessions and the parent-child interactions. As expected, high ratings on FIMP predicted improvements in parenting practices for mothers and for stepfathers. The analyses emphasize the potential utility of FIMP. The results show that fidelity predicts efficacious treatment outcomes.

Although we did not follow the recommendation set forth by Hogue and colleagues (1996) to randomly select sessions for scoring, this study meets many other aims for evaluating competent adherence to an intervention. Because our theory specifies that it is the training in parenting that produces change in parenting practices, we rated sessions focused on contingency management (i.e., skill encouragement and discipline). In future studies, we will FIMP score sessions on the other core components and test the FIMP scores for predictive validity on change in parenting practices. We expect that sampling portions of sessions covering core

components will provide a valid assessment of competent adherence.

The theoretical bases of given treatment approaches lead to use of specific and disparate process skills in the competent delivery. For example, psychodynamic orientations focus on self-discovery through guided questioning and interpretation. Behavioral approaches such as PMTO build skills through teaching, role-play, and continued coaching to shape behavior change. In their study of PMTO therapy process, Patterson and Forgatch (1985) found that teaching, especially combined with confrontation, was associated with increased levels of resistance to the therapist. The same study showed that therapist support was associated with decreased resistance. For this reason, we train PMTO interventionists to avoid pairing teaching with confrontation. Instead, teaching is done within a rich context of support. We emphasize the use of teaching methods that actively engage the parents in the learning process. And we draw on clinical skills that promote openness to growth and development, techniques that have been developed by a number of therapies (e.g., strategic and structural family therapy). Such skills include reframing, joining, mirroring, and use of paradox, to name a few. We think that good teaching and strong clinical skills go hand in hand in PMTO. Both are necessary but not sufficient without the core content components. Thus, teaching skills and clinical process skills have separate and equal weighting in the FIMP. Presumably a psychodynamic approach would mark a therapist down for teaching during intervention. PMTO raters would give poor marks to one who did little or no teaching and for teaching without good clinical support.

Dumas and colleagues (2001) define intervention fidelity as a demonstration that "each of its components is delivered in a comparable manner to all participants and is true to the theory and goals underlying the research" (p. 38). We agree that the components should be true to the theory and goals but we believe that interventionists must adapt timing and approach to match family characteristics. In PMTO, as in all programs that operate through strong working alliances, interventionists must spend time being sociable and responsive to family issues, especially when issues are pressing, chronic, or recurring. For example, to insist on discussing putting the 5-year-old to bed on time when the teenager in the family is experimenting with heroin is like attending to a hangnail when someone is having a heart attack. Thus, being responsive to family circumstances may require additional sessions or less time on agenda topics during the session. In support of this, Reid found that

experienced PMTO therapists spent about 60% of session time on agenda-driven topics; the equivalent for less experienced therapists was 80%. The less experienced therapists also had higher dropout rates (Reid et al., 1979). Rigidly following protocols may not be necessary for therapists with a deep understanding of the theory and who can make adjustments that do not interfere with mechanisms of change.

Some family characteristics make it difficult for interventionists to produce their best work, and this can detract from fidelity ratings. Other qualities may enhance fidelity. In the present study, we found that parents with lower baseline parenting scores had higher FIMP scores. While we were somewhat surprised at the finding, on reflection we decided that lower parenting effectiveness draws on the need to use PMTO techniques. This relationship between low parenting skill and higher fidelity may not occur in clinical treatment studies in which all participants are selected for experiencing significant problems. In prevention studies in which some families have little need for assistance, distressed families may elicit more PMTO structure. It would be useful to test a hypothesis comparing prevention and treatment samples with levels of FIMP based on baseline levels of parenting skill. We also found a difference in change in parenting practices for families in FIMPed subsample and the larger MAPS sample. This difference would presumably make our findings in the present study more conservative since there is less variance to predict.

In the pilot study reported here, the FIMP raters were the intervention supervisor and two of the interventionists, introducing the potential for bias. To control for this problem, the interventionists never scored their own sessions and the supervisor scored 100% of the tapes as a reliability coder. The FIMP raters were also blind to the outcomes for parenting practices. Nevertheless, it will be important in future studies to have nonparticipant raters. We have recently trained to reliability a new cadre of seven FIMP raters in Norway. Their training required approximately 40 hours of supervisor training time. The Norwegian coders viewed videotapes of Norwegian trainees working with families and rated tapes in groups. The OSLC supervisor viewed and rated tapes with them and discussed differences in ratings. Finally the trainees scored three tapes from the dataset presented in this article. Their mean ICC score was .80 ($SD = .12$). Now these FIMP raters have begun scoring a rather large dataset of families throughout Norway. All FIMP raters thus far are certified PMTO specialists, and therefore expert in the use of the technique. Use of trained therapists does, however, lead to increased expense.

Recording intervention procedures in manuals makes it possible to replicate studies across sites and take efficacious programs into communities. Unfortunately, most of us who write these manuals leave out important aspects of our interventions. Some do not explain how the theory relates to the intervention components. Others do not sufficiently describe the teaching and process skills required to activate the mechanisms to their fullest. Evaluations of competent adherence to intervention protocols must take all relevant dimensions into account. Until that time, we who conduct intervention research have far to go before we sleep.

References

- Arbuckle, J. L. (1997). *Amos users' guide Version 3.6*. Chicago: SmallWaters.
- Chamberlain, P. (1994). *Family connections: A treatment foster care model for adolescents with delinquency* (Vol. 5). Eugene: Castalia.
- Chamberlain, P., Davis, B., Forgatch, M. S., Frey, S., Patterson, G. R., Ray, J. R., Rothschild, A., & Trombly, J. (1986). *The therapy process code: An observational system*. Eugene: Oregon Social Learning Center.
- Chamberlain, P., Patterson, G. R., Reid, J., Kavanagh, K., & Forgatch, M. (1984). Observation of client resistance. *Behavior Therapy, 15*, 144–155.
- Chamberlain, P., & Ray, J. (1988). The therapy process code: A multidimensional system for observing therapist and client interactions in family treatment. In R. J. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 4, pp. 189–217). Greenwich, CT: JAI Press.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*, 7–18.
- DeGarmo, D. S., Eddy, J. M., Reid, J. B., & Fetrow, R. A. (submitted). *Evaluating components of the Linking the Interests of Teachers and Families (LIFT) school-based intervention: Predicting initiation and growth in substance use across adolescence*.
- DeGarmo, D. S., & Forgatch, M. S. (2004). Putting problem solving to the test: Replicating experimental interventions for preventing youngsters' problem behaviors. In R. D. Conger, F. O. Lorenz, & K. A. S. Wickrama (Eds.), *Continuity and change in family relations* (pp. 267–290). Mahwah, NJ: Erlbaum.
- DeGarmo, D. S., & Forgatch, M. S. (in press). Early development of delinquency within divorced families: Evaluating a randomized preventive intervention trial. *Developmental Science*.
- DeGarmo, D. S., Patterson, G. R., & Forgatch, M. S. (2004). How do outcomes in a specified parent training intervention maintain or wane over time? *Prevention Science, 5*, 73–89.
- DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14*, 469–482.
- Dishion, T. J., & Kavanagh, K. (2003). *Intervening on adolescent problem behavior: A family centered approach*. New York: The Guilford Press.
- Dumas, J. E., Lynch, A. M., Laughlin, J. E., Smith, E. P., & Prinz, R. J. (2001). Promoting intervention fidelity: Conceptual issues, methods, and preliminary results from the

- EARLY ALLIANCE prevention trial. *American Journal of Preventive Medicine*, 20 (Suppl. 1), 38–47.
- Eddy, J. M., & Chamberlain, P. (2000). Family management and deviant peer association as mediators of the impact of treatment condition on youth antisocial behavior. *Journal of Consulting and Clinical Psychology*, 68, 857–863.
- Forgatch, M. S. (1994). *Parenting through change: A programmed intervention curriculum for groups of single mothers*. Eugene: Oregon Social Learning Center.
- Forgatch, M. S., Bullock, B. M., & Patterson, G. R. (2004). From theory to practice: Increasing effective parenting through role-play: The Oregon model of parent management training (PMTO). In H. Steiner (Ed.), *Handbook of mental health interventions in children and adolescents: An integrated developmental approach*. San Francisco: Jossey-Bass.
- Forgatch, M. S., & DeGarmo, D. S. (1999). Parenting through change: An effective prevention program for single mothers. *Journal of Consulting and Clinical Psychology*, 67, 711–724.
- Forgatch, M. S., Knutson, N., & Mayne, T. (1992). *Coder impressions of ODS lab tasks*. Eugene: Oregon Social Learning Center.
- Forgatch, M. S., & Rains, L. (1997). *MAPS: Marriage and Parenting in Stepfamilies* (parent training manual). Eugene: Oregon Social Learning Center.
- Greene, W. H. (1993). *Econometric analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology*, 65, 821–833.
- Hetherington, E. M., & Clingempeel, G. (1992). *Coping with marital transitions* (Vol. 57). Chicago: University of Chicago Press.
- Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy*, 33, 332–345.
- Kessler, R. C., & Greenberg, D. F. (1981). *Linear panel analysis: Models of quantitative change*. New York: Academic Press.
- Knutson, N. M., Forgatch, M. S., & Rains, L. A. (2003). Fidelity of Implementation Rating System (FIMP): The training manual for PMTO. Eugene: Oregon Social Learning Center.
- Martinez, C. R., Jr., & Forgatch, M. S. (2001). Preventing problems with boys' noncompliance: Effects of a parent training intervention for divorcing mothers. *Journal of Consulting and Clinical Psychology*, 69, 416–428.
- Patterson, G. R. (1988). Family process: Loops, levels and linkages. In N. Bolger, A. Caspi, G. Downey & M. Moorehouse (Eds.), *Persons in context: Developmental processes* (pp. 114–151). Cambridge: Cambridge University Press.
- Patterson, G. R., & Chamberlain, P. (1988). Treatment process: A problem at three levels. In L. C. Wynne (Ed.), *The state of the art in family therapy research: Controversies and recommendations* (pp. 189–223). New York: Family Process Press.
- Patterson, G. R., & Forgatch, M. S. (1985). Therapist behavior as a determinant for client noncompliance: A paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology*, 53, 846–851.
- Pinsof, W. M. (1981). Family therapy process research. In A. Gurman & D. Kniskern (Eds.), *Handbook of family therapy* (pp. 699–741). New York: Brunner/Mazel.
- Price, R. H., & Vinokur, A. D. (1995). Supporting career transitions in a time of organizational downsizing: The Michigan JOBS program. In M. London (Ed.), *Employees, careers, and job creation* (pp. 191–209). San Francisco: Jossey-Bass.
- Ramsey, E., Antoine, K., Kavanagh, K., & Reid, J. B. (1992a). *LIFT Parent Education Component: Grade 1. OSLC Technical manual*. (Available from Oregon Social Learning Center, 160 E. 4th Ave., Eugene, OR 97401.)
- Ramsey, E., Antoine, K., Kavanagh, K., & Reid, J. B. (1992b). *LIFT Parent Education Component: Grade 5. OSLC Technical manual*. (Available from Oregon Social Learning Center, 160 E. 4th Ave., Eugene, OR 97401.)
- Reid, J. B., Fleischman, M. J., Arthur, J., Toobert, D. J., Stern, S., & Patterson, G. R. (1979, December). *Therapist performance observational system*. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, San Francisco.
- Reid, J. B., Patterson, G. R., & Snyder, J. (Eds.). (2002). *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention*. Washington, DC: American Psychological Association.
- Stoolmiller, M., & Bank, L. (1995). Autoregressive effects in structural equation models: We see some problems. In J. M. Gottman & G. Sackett (Eds.), *The analysis of change* (pp. 263–276). Hillsdale, NJ: Erlbaum.
- Stoolmiller, M., Duncan, T. E., Bank, L., & Patterson, G. R. (1993). Some problems and solutions in the study of change: Significant patterns of client resistance. *Journal of Consulting and Clinical Psychology*, 61, 920–928.
- Stoolmiller, M., Duncan, T. E., & Patterson, G. R. (1995). Predictors of change in antisocial behavior during elementary school for boys. In R. Hoyle (Ed.), *Structural equation modeling: Issues and applications* (pp. 236–253). Newbury Park, CA: Sage.
- Stubbs, J., Crosby, L., Forgatch, M. S., & Capaldi, D. M. (1998). *Family and peer process code: A synthesis of three Oregon Social Learning Center behavior codes (Training manual)*. (Available from Oregon Social Learning Center, 160 East 4th Ave., Eugene, OR 97401.)
- Tabachnick, B. G., & Fidell, L. S. (1983). *Using multivariate statistics*. New York: Harper & Row.
- Vinokur, A. D., van Ryn, M., Gramlich, E. M., & Price, R. H. (1991). Long-term follow-up and benefit-cost analysis of the Jobs Program: A preventive intervention for the unemployed. *Journal of Applied Psychology*, 76, 213–219.

RECEIVED: August 4, 2003

ACCEPTED: December 22, 2003