Running Head: PAW

Development and Evaluation of a Parenting Intervention Program: Integration of Scientific and Practical Approaches

David Segal

The Department of Behavioral Science, The College of Management, Israel

Peter Y. Chen

Department of Psychology, Colorado State University, USA

Donald A. Gordon

Christoph D. Kacir

And

Julius Gylys

Department of Psychology, Ohio University, USA

Correspondence regarding the article should be address to:

Peter Y. Chen or Donald A. Gordon. Peter Y. Chen, Department of Psychology, Colorado State University, Fort Collins, CO 80523. Electronic mail can be sent via Internet to chenp@colostate.edu. Donald A. Gordon, Department of Psychology, Ohio University, Athens, OH 45701. Electronic mail can be sent via Internet to gordon@oak.cats.ohiou.edu.

Abstract

We presented a case study to describe practical as well scientific strategies that were applied to assess human-computer interaction, with the foci on software content development (i.e., content validation and interface comparisons), evaluation (i.e., different types of criteria, effect sizes, effect ratios, targeted goals met), and user acceptability. A parenting intervention program entitled Parenting Adolescents Wisely (PAW) was delivered by two formats: non-interactive videotape and interactive multimedia. Based on a content validation model developed in the current study, both formats consisted of critical skills identified from past empirical studies. Results of applying both formats to at-risk families showed improvements on three types of evaluative criteria: reaction, learning, and behavior. Finally, the PAW program showed a substantial cost-benefit based on effect ratios, compared to other parenting interventions.

Development and Evaluation of a Parenting Intervention Program: Integration of Scientific and Practical Approaches

Over the past few years, mental health delivery structures have been drastically changed. These changes are mainly attributed to the rising cost of mental health treatment. Consequently, the pressure to design and conduct cost-effective interventions drives how psychological services are financed, organized, and delivered (Austad & Hoyt, 1992). One plausible solution to this pressing practical challenge is to integrate computer and clinical interventions, and develop an intervention that is efficient and relatively short with a low number of face-to-face contact hours with therapists. With the aid of the current multimedia development technology, it is much easier than ever to develop a multimedia training program. However, identification of valid contents for training programs tended to be processed behind the scene without being thoroughly documented. In the current case study, we first described the process of developing contents for a parenting intervention on the basis of a content validation model described below. In addition, we showed how the program contents were validated by using multiple quantitative approaches (i.e., statistical significance test, targeted goals met, effect size, and effect ratio).

Intervention evaluation is critical but has received less attention in typical clinical practice. A comprehensive evaluation is particularly challenging to researchers and practitioners in the discipline of human-computer interaction. Given its interdisciplinary nature, a comprehensive evaluation of human-computer interaction requires relevant (i.e., valid) and reliable criteria assessing multiple aspects such as mental abilities (e.g., effects on learning speed, recall accuracy, or reasoning), physiological capabilities (e.g., effects on visual acuity, visual field, auditory sensation), interface designs (e.g., ease of use), psychological reactions (i.e., anxiety, or resistance to technology), behavioral changes, and so on. The second purpose of

the study is to demonstrate how the intervention is evaluated by three types of criteria: users' reaction to the program (i.e., reaction criteria), knowledge acquisition from the program (i.e., learning criteria), and behavior transfer to outside of the intervention (i.e., behavior criteria).

Uses of a large sample size and a control group are often restricted due to the practical nature in the field experiment. These constraints render interpretation of field research results quite difficult. We described two supplementary approaches (i.e., targeted-goals met and effect size), in addition to the traditional statistical analysis, that offer insights about the effectiveness of the intervention.

We presented a content validation model depicted in Figure 1, modified from Goldstein's (1993) and Ford, Quinones, Sego, and Sorra's (1992) training evaluation model. The model, served as the guideline to develop the parenting intervention program, consists of two dimensions: parenting skills identified from the empirical literature and parenting skills represented in the intervention program. Valid content of the program is affirmed if the intervention only contains critical and relevant skills identified from the research (i.e., intervention content relevance shown in quadrants A and D). An intervention program would be considered content deficient or contamination if the intervention omitted crucial skills (shown in quadrant B) or included irrelevant skills (shown in quadrant C).

Insert Figure 1 about here

Identification of relevant parenting skills and problems

Based on the empirical literature on child behavior problems including delinquency and substance abuse (e.g., Dumas, 1989; Kazdin, 1988, 1991; McMahon & Wells, 1989; Miller &

Prinz, 1990; Patterson & Forgatch, 1987), we first identified effective parenting skills which include monitoring child behavior, communicating effectively (e.g., clear command, undiluted praise), learning to anticipate and solve new problems and conflicts, using alternatives to physical punishments, administering tangible and social reinforcement, collaborating with the child on setting up contracts and contingency management, and facilitating prosocial behavior. The effectiveness of the above parenting skills has demonstrated improvements in a shift of dysfunctional behavior to the normative behavior range (Webster-Stratton, 1992). It is also evident that improvements often remain one year after treatment with continued benefits up to ten years (Forehand & Long, 1988).

After identifying the relevant parenting skills, we chose nine representative problems based on modified live clinical cases in which the aforementioned skills are critical to the formation of well-adjusted adolescents, and to the amelioration of the common problems experienced by adolescents. These problems include children not doing chores, not completing chores satisfactorily, not complying with parental requests to get off the telephone or turn down the volume of music, not completing homework, fighting with siblings, associating with peers of which parents didn't approve, and speaking disrespectfully.

Choices of Delivery Media

Early parenting interventions were generally conducted by a therapist in an individual, face-to-face format, which has time-constraints, cost, and inflexibility disadvantages in meeting rapid epidemiological and economic demands. As a result, standardized non-interactive videotape parenting programs, which contain pre-recorded parenting-skills, were created to meet the challenge. Webster-Stratton and her colleagues (Webster-Stratton, Kolpacoff, & Hollinsworth, 1988; Webster-Stratton, 1990, 1992) have demonstrated that the effectiveness of

self-administered videotape parenting programs is comparable to more elaborate programs that combined the videotapes with therapist-led discussion groups or therapist consultation.

Facing rapid advancement in software development, it is reasonable and logical for program developers to go beyond the non-interactive videotape parenting program by utilizing interactive multimedia that allows parents to determine the pacing, sequence and selection of learning content. These interactive features accommodate different parents' responses and permits parents to choose different pathways throughout their learning process (e.g., exit any part of the program whenever they desire or choose to review other contents in the program). Since this type of learning is deductively structured with enriched and contrived experience, it presumably allows learners to take responsibility for their learning (Cohen, 1984), which in turn facilitates the learning of new material and the exploration of new ideas (Hapeshi & Jones, 1992). Furthermore, it has been argued that clients prefer to interact with a machine rather than a human therapist when instructional knowledge and its practice were a primary requirement for behavioral change (Burda, Starky, & Dominguez, 1991). This preference may be attributed to the non-judgmental nature of computers which encourages clients to learn tasks thoroughly or repeat difficult segments of the content at their own pace. When the content is parenting, a topic which elicits defensiveness in virtually every parent, this advantage is substantial. Given the above favorable characteristics associated with interactive multimedia programs, we predicted that interactive multimedia format would outperform non-interactive videotape format.

Following the above content validation model, we developed a parenting program entitled Parenting Adolescents Wisely (PAW) with two different formats: non-interactive videotape (NV) and interactive multimedia (IM). The interactive multimedia program, originally written with Hypercard™, and with Authorware™ afterwards, was developed by a team,

directed by the third author, composed of faculty, staff, and graduate students in psychology, multimedia services, film, computer sciences, and telecommunications. As described earlier, nine problems typical of most families, as well as effective and ineffective parenting skills which parents use to deal with these problems, were presented via the multimedia hardware.

Each problem was first presented with video scenes which depicted a typical negative interaction between a parent and child. The running time of each scene is approximately two minutes. After each scene, users are prompted to choose one of 2-3 possible solutions of which only one is an effective approach to problem resolution. There are 26 scenes depicting possible solutions to the nine problems listed above. The characters in each scene were enacted by parents, children, and adolescents with previous acting experience.

Each possible solution is followed by a critique presented on-screen in a written question-answer format. The critiques are a critical part of the program, for they point out the errors parents make, explain why ineffective solutions lead to problems and why effective parenting solutions prevent and solve problems. An additional opportunity to practice skills taught in the program, in the form of an on-screen quiz, follows the selection of the appropriate solution to each problem. The instructional methods used in the program are demonstration, quizzing, repetition, rehearsal, recognition, and feedback for correct and incorrect answers. The program takes, on average, 2.5 hours to complete.

Except for the on-screen quiz, the non-interactive videotape intervention contains all the aforementioned problems shown in the interactive multimedia program. Due to its linear nature, the videotape intervention requires users to sequentially watch all the 26 scenes that include effective as well as ineffective solutions. This unique feature is different from the traditional

videotape parenting programs that only provided effective methods, and they are not usually shown in a realistic family context.

Program Evaluation

As stated earlier, the second purpose of this case study is to demonstrate how the PAW intervention is evaluated by multiple criteria. The evaluative evidence can also substantiate content validity of the program. While selecting criteria, we applied a modified model proposed by Kirkpatrick (1959a, 1959b, 1960). According to the modified model, three types of criteria were assessed: reaction, learning, and behavior. Reaction criteria are referred to outcome measures that assess users' affective reaction toward training programs; while learning criteria are the outcomes which focus on amount of knowledge or skills learned from the programs. Behavior criteria, the most important aspect in evaluation, indicate the extent to which trainees actually change (or transfer) outside of the programs. Evaluation of PAW (or other intervention programs) would not be complete without knowing if parents attain two ultimate goals: the parents are actually practicing the learned skills outside the intervention environment, and their children's behavior problems decrease.

Following Kirkpatrick's suggestion, we evaluated a reaction criterion that measured how parents liked or felt about PAW (a user acceptability and satisfaction measure). We also measured a learning criterion that gauged how effective and ineffective parenting skills were understood and absorbed by the parents during and immediately after the intervention. Finally, we assessed a series of behavior criteria (i.e., observed negative and expected positive behaviors, discipline strategies, and use of relevant skills) that showed how frequently and effectively parents used the skills outside of the intervention. A summary of the criteria assessed in each stage is presented in Table 1. Choices of criteria should meet, at least, two psychometric

requirements: relevance and reliability. These requirements are the cornerstones of sensible evaluation because the evaluative results with irrelevant or unreliable criteria will be inconclusive.

Insert Table 1 about here

Learning criterion. Parental knowledge of parenting-skills related to the PAW program was measured by the Parenting Knowledge Test (PKT). This measure is a 34-question, multiple-choice test which measures knowledge of parenting skills taught in the PAW program. The PKT was designed for the current study. A pilot study using only the PAW workbook showed that participants who read this workbook demonstrated increased performance on the PKT, compared to a control group receiving no such parenting skill information.

Reaction criterion. Users' acceptability of the program was surveyed by the Consumer Satisfaction Questionnaire (CSQ), which was modified from the original measure of Forehand and McMahon (1981). The CSQ assessed four aspects of users' acceptability: perceptions about child behavior improvement, teaching format difficulty, intervention usefulness, and perceived overall difficulty of the parenting-skills that were taught. The measure consists of 38 items with variations on a 7-point response format. Webster-Stratton (1989) reported internal consistency reliability estimates varying from .71 to .90.

Behavioral criteria. Parents' perceptions of child adjustment were assessed by the Eyberg Child Behavior Inventory (ECBI, Eyberg & Ross, 1978). The ECBI consists of 36 items pertaining to child misconduct behavior in a yes-no format. It has been shown to differentiate

clinic-referred and non-clinic populations (Robinson, Eyberg, & Ross, 1980). Reliability estimates were reported to be 0.86 (test-retest) and 0.98 (internal consistency).

Parents' observed negative and expected positive behaviors were assessed by the modified Parent Daily Report (PDR; Webster-Stratton, 1990), which was originally developed by Chamberlain and Reid (1987). The questionnaire consists of a list of 23 negative and 28 prosocial behaviors commonly displayed by children. During the intake, parents were asked to select those negative behaviors from the list that they felt were major problems and those positive behaviors that would be particularly pleasing to them if performed by their child. These individually tailored checklists were then used as the basis for the phone calls that were conducted weekly at post-test assessment. During phone calls, the checklist was read to the parent. The parent was asked to observe and report on the occurrence or nonoccurrence of the target behaviors in the previous 24 hours.

Parents' discipline strategies were recorded by the Daily Discipline Interview (DDI; Webster-Stratton & Spitzer, 1991). The DDI consisted of all the PDR behavior checklists on which parents described how they handled the particular child behaviors. The description was then rated to be either appropriate or inappropriate according to a coding system consisting of nine categories. This coding system was specifically designed for the current study. The ratio of the number of appropriate strategies to number of behaviors presented by the child is computed and used for the analysis. The mean percent interrater agreement reliability ranged from 60% to 88% ($\underline{M}$ = 80%), and split-half reliability coefficients ranged from .59 to .86 across all nine parenting skills.

Parents' use of the adaptive parenting skills taught in the PAW program was evaluated by the Parent Behavior Questionnaire (PBQ). This measure was developed for use in the current

study. The PBQ is an 8-item measure which assesses how often in the past month parents implemented parenting skills (e.g., active listening, assertive discipline, contingency management, and positive reinforcement) taught in the PAW program.

Criteria of selecting participants. Forty-two parents were recruited through three outpatient mental health clinics and the local community. They were selected on the basis of the following entry criteria. First, they needed to have a child between the ages of 11 and 18 living in the home, as the children portrayed in the video scenes were within this age range. In addition, they were not participating in any other parenting program during the study. Third, their primary problem was related to negative child behaviors (e.g., noncompliance) and they reported more than the average number of behavior problems on the Eyberg Child Behavior Inventory (ECBI; Eyberg & Ross, 1978). Finally, they had to have completed at least the 5th grade (as the text in the PAW program is written at a 5th grade reading level), and be accessible by phone. Only one parent per family participated in the study. They were paid $30 for their participation. Mean age and year of education of the parents were 42 and 13. Ninety percent of the parents were females, and 45% were single. Eighty-six percent reported incomes lower than $25000.

Intervention Procedure. Once parents met entry criteria they were randomly assigned to one of the two intervention conditions: non-interactive videotape (NV) or interactive multimedia (IM). Parents in the NV group viewed the same video scenes as those in the IM group. Prior to the beginning of the intervention, parents completed pretreatment questionnaires consisting of demographic data, one learning criterion (i.e., PKT), and four behavior criteria (i.e., ECBI, PDR, PBQ, DDI). Beginning one week after the intervention, all participants were contacted by phone twice a week for the following three weeks and asked to respond to the PDR and DDI. In the

final week, they also completed a mailed package containing the ECBI, PKT, PBQ, and the CSQ.

Parents in the IM group were instructed to interact with the program delivered by a computer until they had observed and interacted with all nine problems and their solutions. It took an average of 2.5 hours to complete the interactive multimedia program. Due to the nature of the non-interactive videotape technology only a summary of the major points of the critiques of the solutions were presented using on-screen text and voice-over narration. The on-screen quiz was not present in the NV format. To control for any content differences, all parents were referred to the accompanying workbook that contains all the printed information in the same sequence employed in the program. It took parents in the NV group 1.5 hours to observe all nine problems and the correspondent solutions with critiques from the videotape.

Evaluative results. Comparisons between the two groups on the dependent variables (i.e., ECBI, PBQ, PKT, PDR positive behaviors, PDR negative behaviors, DDI positive behaviors, and DDI negative behaviors) at the pretreatment stage, the Hotelling T^2 showed no significant mean differences, $F(7,32)=0.12$ based on Wilks' criterion. Due to the interactive nature of the multimedia presentation, we also assessed if there was any difference on number of problems reviewed by both groups of subjects. No significant differences were found ($\chi^2= .17$). In general, parents were satisfied with both programs with respect to teaching format, usefulness, and parenting techniques. The ratings ranged from 5.16 to 6.11 on a 7-point scale. No significant differences on the above aspects were found between the groups.

Based on a 2 (Treatment: IM vs. NV) by 2 (Time: pre vs. post) mixed design, we conducted a multivariate analysis of variance on the seven dependent variables. According to Wilks' criterion, there was a significant Time main effect, $F(7,32)=16.67$, and an interaction

effect, F(7,32)=3.27. No Treatment main effect was found, which did not support our expectation. The significant Time main effect supported the validity of the PAW program. As described in Table 2, the univariate analyses showed significant improvements from pre- to post-treatment on four of the dependent variables in the IM group and six of the dependent variables in the NV group. Although the significant interaction effect suggested the formats have differential effects on the dependent variables at the post-treatment stage, the univariate analyses were not able to support this interaction due to low statistical power.

Insert Table 2 about here

Supplementary evaluative approaches. Given the practical limitations of using a large sample size and a control group in the present case, we utilized two additional approaches, which can offer insights about the effectiveness of the intervention. The first approach focuses on to what extent parents' behavior change meets the ultimate goal of the intervention. In other words, are the parents who receive the intervention meeting our goal (i.e., their children's behavior was similar to the functional population). The second approach compares effect size and effect ratio (i.e., effect size per hour of contact time) of the current study to those reported in the literature. Effect sizes offer another perspective to quantify how effective the intervention is without relying on statistical significance tests.

Three indexes were used to evaluate if the parents meet the goal of the intervention: reliable change index (RCI), movement closer to the functional population mean, and clinical significance. According to Jacobson and Truax (1991), RCI indicates whether an individual's pre- to post-change score on a particular outcome measure is more parsimoniously explained by

the error inherent in the outcome measure (i.e., standard error of measurement, SEM) or by the effect of treatment. Specifically, reliable change is defined as pre-to post-difference scores that exceed 1.96 SEM. The second criterion requires an individual to demonstrate movement from a "dysfunctional" to a "functional" population distribution. In other words, we estimated the extent to which the parents' children moved into the functional population by significantly moving away from the mean of the dysfunctional population. Finally, both reliable change and movement closer to the functional population mean had to be met before an individual's pre-to-post change was considered clinically significant.

Based on the ECBI norms and reliability estimates, we assessed the RCI and movement closer to the functional population mean in both IM and NV groups. Overall, more than 50% of parents' children in both groups showed reliable change and movement away from the mean of the dysfunctional population. With respect to clinical significance, 33% in the IM group and 48% in the NV group met this stringent standard. These results, as reported in Table 3, provide further evidence of the validity of the programs.

Insert Table 3 about here

To further examine the clinical utility of the PAW program, the effect size based on Cohen's *d*, suggesting the improvement from the pre- and post-test ECBI and PDR scores, was compared to those reported in the literature (see Tables 4 and 5). The PAW program took an average of 1.5 to 2.5 hours for parents to complete, yet the results of the present study are similar to those obtained through longer (10-12 hours) programs (Patterson, Chamberlain, & Reid, 1982; Webster-Stratton, 1984; Webster-Stratton et al., 1988; Webster-Stratton, 1992). Figure 2 depicts

the effect ratio (i.e., effect sizes per hour of contact time). The PAW program (both formats) shows a substantially greater effect ratio than the other parenting interventions.

Insert Figure 2, Tables 4 and 5 about here

Conclusion

In the present case study, we described the process of developing the PAW program based on the content validation model. The model offers conceptual guidelines that can be used to identify relevant contents in various computer-aided interventions. Identification of relevant contents are requirements that guarantee the effect of any interventions. Based on the model, content relevance can first be verified based on empirical literature, followed by experts' judgment. Quality of experts' judgment can further be assessed by examining an agreement among them (e.g., rater-agreement index, Burry-Stock, Shaw, Laurie, & Chissom, 1996) and be evaluated by the criteria, particularly by learning criteria. .

Besides relying on traditional statistical significance tests (e.g., MANOVA in this case), we demonstrated how other quantitative evaluative approaches (e.g., effect size, effect ratio, targeted goals met) can be used to validate the intervention contents in practice. Different approaches offer unique aspects for different purposes. For instance, it would be in the best interest of a trainee to know if he/she reaches the goals set up by an intervention program (e.g., their children's behavior was similar to the functional population). It will also be useful for any involved party (e.g., health agency) to know how efficient (e.g., effect ratio) and effective (e.g., effect size) an intervention program is. Based on the effect ratio, the PAW program shows a substantially greater cost-benefit ratio than the other parenting interventions. This finding has

implications for dealing with at-risk families who are difficult to engage and retain in treatment. The costs of intervening are less with such programs as PAW than with traditional, clinician delivered programs; thus, the same amount of funds can be used to treat more families. The very brief time commitment families need to make to use of the program increases its access and appeal. The relatively immediate benefits parents experience with such a program further enhances its appeal.

A unique aspect of the current evaluation is to include three types of criteria, reaction, learning, and behavior. Information derived from the above multiple criteria allows us to identify where users have improved or need to improve in the future. The criteria focusing on different aspects are relevant to the parties who are involved in developing the intervention. For instance, the level of satisfaction with the program and the amount of knowledge improvement provide critical cues for program designers and content developers. Given the ultimate goal of an intervention is to transfer what users have learned from the intervention back to their actual behavior changes in reality, it is imperative to include behavior criteria in any program evaluation.

The current study also revealed an interesting finding. The effectiveness difference between the two PAW formats was not clearly shown according to the independent univariate analyses, although it seemed that the NV format slightly outperformed the IM format. Although there is no reason to expect the NV format to perform better, it is possible that the parents in the NV format watched through the effective as well as the ineffective solutions because they had no other choice. On the other hand, once the parents in the IM group chose the best solution, they were rewarded with praise and given the option to move on to the next problem. Future research and software development should examine to what extent user-control can inhibit or facilitate

learning in the interactive multimedia platform. Research on interactive video formats has shown the superiority of this medium over traditional methods on long-term retention and performance (Fletcher, 1990; Niemiec and Walberg, 1987). We may not have allowed enough time for these formats to show differences in this study, so we recommend a six-month to one year follow up. The IM format in clinical settings would produce more attention and involvement compared to the NV format. Agencies are not likely to closely monitor parents who are given the video (NV format), so fewer parents would see the whole video, be paid for watching it and completing evaluation forms, and then show the treatment gains we found here. In addition, the IM format offers opportunities for interaction among family members and should increase the implementation of skills among those members. Many agencies are using an advanced version of the PAW program (the *Parenting Wisely* CD-ROM) in this manner, and take it into homes for family members to use together, as well as using it with groups of parents. Engaging high risk families, in community centers or in the home, with this CD-ROM format has proven successful (Gordon, 2000). The NV format does not lend itself to this kind of participation.

Professionals who provide services to families at risk for child behavior problems often face both external and internal constraints such as health care reimbursement limits and larger caseloads of families who need services. Interventions such as the multimedia modeling of parenting skills described here, may assist these frontline professionals in bringing cost-effective and evidenced based interventions to their child and parent clients. The PAW program has shown promise in producing the effect sizes comparable to those found in the past literature. Due to the simplicity of implementation, the assurance of treatment integrity, and the low cost, this type of program can facilitate the dissemination of research-based methods among

practitioners.

References

Austad, C. S. & Hoyt, M. F. (1992). Psychotherapy in a staff model health maintenance organization: Providing and assuring quality care in the future. Special issue: The future of psychotherapy. Psychotherapy, 29, 119-129.

Burda, P. C., Starky, T. W., & Dominguez, F. (1991). Computers administered treatment of psychiatric inpatient. Computers in Human Behaviors, 7, 1-5.

Burry-Stock, J. A., Shaw, D. G., Laurie, C., & Chissom, B. S. (1996). Rater agreement indexes for performance assessment. Educational & Psychological Measurement, 56, 251-262.

Chamberlain, P., & Reid, J. B. (1987). Parent observation and report of child symptoms. Behavioral Assessment, 9, 97-109.

Cohen, V. B. (1984). Interactive features in the design of videodisc materials. Educational Technology, 24, 16-20.

Dumas, J. D. (1989). Treating antisocial behavior in children: Child and family approaches. Clinical Psychology Review, 9, 197-222.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. Psychological Methods, 1, 170-177.

Eyberg, S. M., & Ross, A. W. (1978). Assessment of child behavior problems: The validation of a new inventory. Journal of Clinical Psychology, 16, 113-116.

Fletcher, J. D. (1990). Effectiveness and cost of interactive video in defense training and education. (Report No. P2732). Institute for Defense Analysis.

Ford, J. K., Quinones, M. A., Sego, D. J., & Sorra, J. S. (1992). Factors affecting the opportunity to perform trained tasks on the job. Personnel Psychology, 45, 511-527.

Forehand, R. L. & Long, N. (1988). Outpatient treatment of the acting-out child: Procedures, long-term follow-up data, and clinical problems. Advances in Behavior Research and Therapy, 10, 129-177.

Forehand, R. L., & McMahon, R. J. (1981). Helping the Noncompliant Child: A Clinician's Guide to Parent Training. The Guilford Press, New York, N.Y.

Goldstein, I. L. (1993). Training in organizations (3rd ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.

Gordon, D. A. (2000). Parent training via CD-ROM: Using technology to disseminate effective prevention practices. Journal of Primary Prevention, 21, No. 2, 227-251.

Hapeshi, K, & Jones, D. M. (1992). Interactive multimedia for instruction: A cognitive analysis of the role of audition and vision. International Journal of Human-Computer Interaction, 4, 79-99

Jacobson, N. S.; &Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting & Clinical Psychology, 59, 12-19

Kazdin, A. E. (1988). Child Psychotherapy. Pergamon Press, Elmsford, NY.

Kazdin, A. E. (1991). Effectiveness of psychotherapy with children and adolescents. Journal of Consulting and Clinical Psychology, 59, 785-798.

Kirkpatrick, D. L. (1959a). Techniques for evaluating training programs. Journal of American Society for Training and Development, 13, 3-9.

Kirkpatrick, D. L. (1959b). Techniques for evaluating training programs: Part 2 - Learning. Journal of American Society for Training and Development, 13, 21-26.

Kirkpatrick, D. L. (1960). Techniques for evaluating training programs: Part 3 - Behavior. Journal of American Society for Training and Development, 14, 13-18.

McMahon, R. J. & Wells, K. C. (1989). Conduct disorders. In Mash, E. J. & Barkley, R. A. (Ed), Treatment of Childhood Disorders. The Guilford Press, New York, NY.

Miller, G. E. & Prinz, R. J. (1990). Enhancement of social learning family interventions of childhood conduct disorder. Psychological Bulletin, 108, 291-307.

Nay, R. W. (1976). A systematic comparison of instructional techniques for parents. Behavior Therapy, 6,14-21.

Nimiec, R., & Walberg, H.J. (1987). Comparative effects of computer-assisted instruction: A synthesis of reviews. Journal of Educational Computing Research, 3, 19-37.

Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of a parent-training program. Behavior Therapy, 13, 638-650.

Patterson, G. R. & Forgatch, M. S., (1987). Parents and Adolescents Living Together. Castalia Publishing Company, Eugene, OR.

Robinson, E. A., Eyberg, S. M., & Ross, A. W. (1980). The standardization of an inventory of child conduct problem behaviors. Journal of Clinical Child Psychology, 8, 22-29.

Webster-Stratton, C. (1984). Randomized trial of two parent-training programs for families with conduct-disordered children. Journal of Consulting and Clinical Psychology, 52, 666-678.

Webster-Stratton, C. (1989). Systematic comparison of consumer satisfaction of three cost-effective parent training programs for conduct problem children. Behavior Therapy, 20, 103-115.

Webster-Stratton, C. (1990). Enhancing the effectiveness of self-administered videotape parent training for families with conduct-problem children. Journal of Abnormal Child Psychology, 18, 479-492.

Webster-Stratton, C. (1992). Individually administered videotape parent training: "Who benefits?" Cognitive Therapy and Research, 16, 31-52.

Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. Journal of Consulting and Clinical Psychology, 56, 558-566.

Webster-Stratton, C. & Spitzer A. (1991). Development, reliability, and validity of the Daily Telephone Discipline Interview. Behavioral Assessment, 13, 221-239.

Authors Note

This article is based on the first author's doctoral dissertation under the direction of the third author. We retain responsibility for the conclusions reported here, which do not represent views of our affiliations. We thank Peg Rothwell for the editorial comments and Paul Spector for the statistical suggestions. This research was supported by the Office of Juvenile Justice and Delinquency Prevention's Formula Grants program in Ohio, grant #JJIN9404. We thank Thomas Mallory for his assistance as project officer.

Table 1
Summary of evaluative criteria assessed in different stages

Types of Criteria		1st pretest	2nd pretest	1st posttest	2nd posttest
Reaction	Consumer Satisfaction Questionnaire				X
Learning	Parenting Knowledge Test	X			X
Behavior	Eyberg Child Behavior Inventory	X			X
	Parent Behavior Questionnaire	X			X
	Parent Daily Reports	X	X	X	
	Daily Discipline Interview		X	X	

Table 2

Mean, standard deviation, F statistic value, effect size (η^2), and power for pre and post comparison within the NV and IM groups.

Interactive Multimedia Intervention (IM)					
	Pretest Mean/ SD	Posttest Mean/ SD	F	η^2	Power
ECBI	17.15 / 6.63	11.85 / 6.91	13.08**	.41	.93
PBQ	20.85 / 8.43	26.0/ 6.21	16.95**	.47	.97
PKT	26.45/ 4.22	30.45 / 3.47	17.18**	.47	.98
NPDR	3.63 / 1.87	1.80 / 1.01	26.91**	.59	1.00
PPDR	10.08 / 3.74	9.44 / 4.50	1.19	.06	.18
NDDI	.41 / .29	.48 / .29	.54	.03	.11
PDDI	.30 / .18	.34 / .24	.33	.02	.08

Linear Videotape Intervention (NV)					
	Pretest Mean / SD	Posttest Mean / SD	F	η^2	Power
ECBI	14.9 / 5.87	9.2 / 7.88	22.20**	.54	.99
PBQ	18.3 / 6.82	25.2 / 7.36	14.34**	.43	.95
PKT	25.45 / 4.58	28.35 / 3.70	15.53**	.45	.96
NPDR	3.53 / 2.56	1.96 / 1.93	24.71**	.57	1.00
PPDR	6.9 / 4.44	8.16 / 3.14	7.44*	.28	.73
NDDI	.21 / .25	.55 / .26	33.46**	.64	1.00
PDDI	.20 / .25	.22 / .12	.19	.01	.07

Note. **p < .01, *p < .05. ECBI = Eyberg Child Behavior Inventory. PBQ = Parent Behavior Questionnaire. NPDR = Parent response to negative targeted behaviors assessed by Parent Daily Reports. PPDR = Parent response to positive targeted behaviors assessed by Parent Daily Reports. NDDI= Ratio of the number of appropriate strategies to number of negative targeted behaviors. PDDI = Ratio of the number of appropriate strategies to number of positive targeted behaviors.

Table 3

Frequencies and percentages of children's reliable change, movement into the functional distribution, and clinical significance.

	IM (%)	NV (%)
Reliable Change Index	10 (48)	12 (57)
Functional Distribution	11 (52)	15 (71)
Clinical Significance	7 (33)	10 (48)

Note. IM= Interactive Multimedia Intervention; NV=Non-interactive Videotape Intervention.

Table 4

Mean, standard deviation, and effect size (Cohen's d) for the pre- and post- Eyberg Child Behavior Inventory.

Reference	Description	Length	Pretest Mean (SD)	Posttest Mean (SD)	d
Webster-Stratton (1984)	Videotape with Group Discussion	10-12 hours	21.80 (6.9)	9.70 (5.9)	1.89
Webster-Stratton et al. (1988)	Videotape Modeling with Group Discussion	10-12 hours	21.00 (5.4)	12.77 (8.4)	1.19
Webster-Stratton et al. (1988)	Videotape modeling	10-12 hours	20.11 (5.8)	11.7 (5.7)	1.46
Webster-Stratton (1992)	Videotape modeling	10-12 hours	20.74 (5.39)	12.34 (11.69)	.98
Present study	Interactive multimedia	2.5 hours	17.15 (6.63)	11.85 (6.91)	.78
	Non-interactive videotape	1.5 hours	14.9 (5.87)	9.2 (7.88)	.83

Table 5

Mean, standard deviation, and effect size (Cohen's d) for the pre- and post- child's targeted negative behaviors on the Parent Daily Report.

Reference	Intervention	Length	Pretest Mean (SD)	Posttest Mean (SD)	d
Webster-Stratton (1984)	Videotape with Group Discussion	10-12 hours	7.10 (2.5)	3.20 (1.3)	2.05
Webster-Stratton et. al. (1988)	Videotape Modeling with Group Discussion	10-12 hours	6.50 (2.5)	3.10 (1.9)	1.55
Webster-Stratton et. al. (1988)	Videotape modeling	10-12 hours	6.30 (3.0)	3.60 (2.5)	.98
Webster-Stratton (1992)	Videotape modeling	10-12 hours	5.89 (2.54)	4.00 (2.47)	.75
Patterson et al. (1982)	Individual Parent Training	17 hours	3.19 (n.a.)	1.66 (n.a.)	.74
Present Study	Interactive multimedia	2.5 hours	3.63 (1.87)	1.80 (1.01)	1.27
	Non-interactive Videotape	1.5 hours	3.52 (2.56)	1.96 (1.93)	.69

Note: n.a. = Not available. [a]Cohen's d is estimated based on F(1,17)=20.35, reported by Patterson et al. (1982, p. 656), with the formula, $d = t_c [2(1-r)/n]^{1/2}$, where r = correlation across pairs of measure (Dunlap, Cortina, Vaslow, & Burke, 1996). Since the correlation was not available in Patterson et al. (1982), it was estimated based on the present study.

Figure 1

Conceptual content validation model.

Figure 2

Effect size per hour (Effect Ratio) across studies reported in Tables 4 and 5.

Figure 1

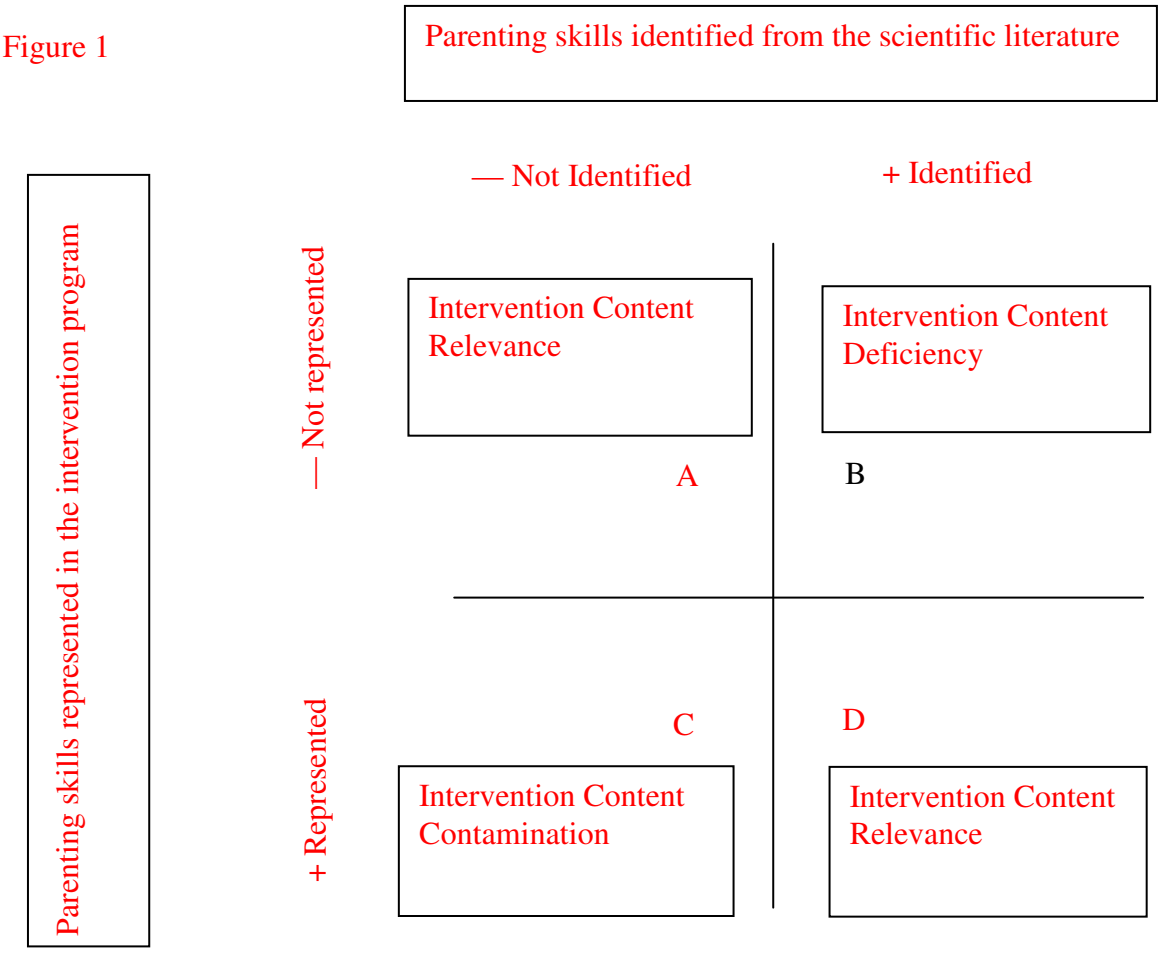
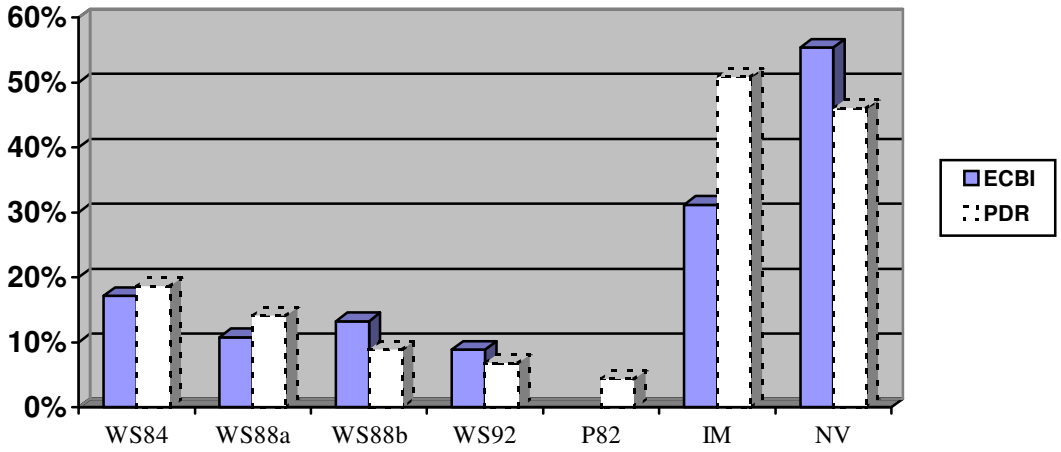

Parenting skills identified from the scientific literature
— Not Identified
+ Identified
Parenting skills represented in the intervention program
— Not represented
+ Represented
Intervention Content Relevance
Intervention Content Deficiency
A
B
C
D
Intervention Content Contamination
Intervention Content Relevance

Figure 2



Note: ECBI = Eyberg Child Behavior Inventory; PDR = Parent Daily Report; WS84 = Videotape with group discussion in Webster-Stratton (1984) study; WS88a = Videotape modeling with group discussion in Webster-Stratton et al. (1988); WS88b = Videotape modeling in Webster-Stratton et al. (1988); WS92 = Videotape modeling in Webster-Stratton (1992); P82 = Individual parent training in Patterson et al. (1982); IM = Interactive multimedia in the current study; NV = Non-interactive videotape in the current study.