

# The Importance of Understanding Dosage When Evaluating Parenting Programs: Lessons from a Pilot Study

**Pamela B. Payne**

*Weber State University*

**Daniel A. McDonald**

*The University of Arizona*

*As government resources for community programs diminish, it is vital that Cooperative Extension make greater efforts to show program efficacy. Assessing the appropriate amount of an intervention optimal for reaching desired outcomes can help inform program development and provide for a more efficient use of limited resources. The current pilot study (funded by CYFAR, NIFA, USDA award #2008-41520-04810) focuses on dosage and its effect on outcomes in parenting education delivered in four states.*

*Keywords:* evaluation, program evaluation, parenting, common measures, dosage

## Introduction

In 2005, the Children, Youth and Families At Risk (CYFAR) Sustainable Communities Evaluation Task Force was formed to assess the degree to which CYFAR projects were achieving their intended outcomes and to gauge the impact that CYFAR projects were having on the status of children, youth, and families across the United States. A performance-monitoring framework was developed to track progress toward global indicators specified for each outcome cluster. The initial outcome clusters included the following areas: youth citizenship, parenting, healthy lifestyles, workforce preparation, and literacy and communication. While global indicators provided a first step towards gauging the impact of CYFAR projects across the country, there were substantial limitations inherent in this method. The global indicators continued to be measured using a variety of methods and instruments, which inhibited the Task Force from making any conclusive judgments concerning the impact projects were having collectively. In 2009, a pilot study was launched with the goal of determining the extent to which CYFAR projects could use identical instruments as part of their evaluation efforts. CYFAR Sustainable Communities Projects (SCP) associated with the parenting cluster were selected for the pilot study and invited to participate. The current study evaluated 4 programs across the United States in the domain of parenting through the CYFAR SCP funded by USDA National Institute of Food and Agriculture.

---

Direct correspondence to Pamela Payne at [pamelapayne@weber.edu](mailto:pamelapayne@weber.edu)

One main question addressed through this pilot study was whether or not there were common elements, such as dosage or length of program, that predicted changes in parenting behaviors, such as communication, involvement, anger management, and associated parenting skills over time and across varied educational classes and curricula. Previous research has suggested that dosage is a key factor in being able to demonstrate efficacy of community programs (Braverman & Engle, 2009; Payne & McDonald, 2012) and that dosage, in fact, is a key element in program integrity (Dane & Schneider, 1998; Duerden & Witt, 2012). This paper explored the impact of dosage in parenting program evaluation.

## Parenting Programs

### Program Descriptions

A challenging feature of this evaluation is that it involved a variety of parenting programs while employing a single common evaluation instrument. Of the four CYFAR SCP states that chose to participate, three utilized the *Strengthening Families Program for Parents* (SFPP), which is a nationally recognized curriculum (Kumpfer & DeMarsh, 1983). SFPP is geared towards small groups of families so that parents and children can be brought together to learn in environments that strengthen the entire family as a unit. The program is designed to be 14 sessions, but we found that some participants received greater or fewer sessions in their specific program. The remaining state used a combination of two recognized programs, *1-2-3 Magic* (Phelan, 1996) and *Common Sense Parenting* (Burke & Herron, 1992). The *1-2-3 Magic* program (Phelan, 1996) is a series of DVDs and parenting trainings that focuses on utilizing effective discipline. The program examined in this pilot study used *Common Sense Parenting* (CSP) with parents of young children. CSP focuses on improving child behavior, parent attitudes, family satisfaction, and parent problem-solving abilities. This program is designed to be implemented in 2-hour sessions once a week for 6 weeks (Burke & Herron, 1992).

### Methods

The parenting domains selected for evaluation were parent involvement, communication, anger management, and parenting skills. Respondents were asked to refer to one child when completing the survey, but the child could range in age from early childhood through adolescence. Items were selected from the Intervention Targeting Parent Behavior scale (ITPB) (Spoth, Redmond, Haggerty, & Ward, 1995), as well as the Parent Education Survey (PEPWT), which combined items from the 3 Cities Study and the National Longitudinal Survey of Youth (NLSY) child supplement (Cornell Cooperative Extension, 2009).

The ITPB is a 13-item measure that has been used in Native American, Latino, and African-American populations. Three items on addiction and substance use behavior were eliminated

from the common measure as they were not the intended focus of the study and were not relevant for some of the target participants. This measure had good established reliability ranging from  $\alpha = .68$  to  $\alpha = .87$  in previous studies (Spoth et al., 1995). The PEPWT is a 10-item measure with items combined from the 3 Cities Study and the National Longitudinal Survey of Youth Child Supplement form. The PEPWT includes questions asked of nationally representative samples (Baker & Mott, 1989). The PEPWT measure had good established reliability ranging from  $\alpha = .50$  to  $\alpha = .90$  in previous studies (AZREACH, 2014).

Analyses presented here coming from the ITPB (Spoth et al., 1995) and the forced choice section of the PEPWT (Cornell Cooperative Extension, 2009). The open-ended portions of the PEPWT yielded inconsistent results from participant responses, and therefore, were not used for this study (discussed further in the limitations section).

The survey, combining portions of the ITPB and PEPWT, was used as the common measure and was offered in a pre/post format or retrospective format to accommodate variation in program length, cycle, and frequency and was also made available in both paper/pencil format or online (via Survey Monkey and a website). Two states implemented a pre/post format, while the other two sites used the retrospective format. Although the online formats were available, all sites elected to use a pencil/paper method and conduct data entry onsite. Data were collected between 2009 and 2011.

Given the variability in program content, length, and location, it was important to assess the “dosage” of the program that participants received. Dosage was assessed through questions that asked participants how many sessions they attended of a specific program, how often the sessions were held, how many hours each session lasted, in addition to the specific dates that participants started and discontinued attending the program. In order to examine the total number of hours more efficiently, groups were established based on the number of hours a participant received parenting education. To yield balanced groups, the low group ( $N = 61$ ) consisted of participants who spent nine hours or less in a program, the medium group ( $N = 52$ ) consisted of those who spent between 10 and 12 hours in a program, and the high group ( $N = 51$ ) consisted of those who spent 13 or more hours in a specific program.

## Results

Table 1 shows the demographics of parenting programs combined and by state, including sample size, participant sex, age, ethnicity, marital status, mean number of children, and average age of the children.

**Table 1. Parenting Program Demographics By State**

| Participant Sex (%)                   |      | Participant Age (%) |      | Ethnicity (%)   |      | Marital Status (%)              |      | Mean Number of Children (SD) | Mean Age of Children (in years) |
|---------------------------------------|------|---------------------|------|-----------------|------|---------------------------------|------|------------------------------|---------------------------------|
| <i>State (N)</i>                      |      |                     |      |                 |      |                                 |      |                              |                                 |
| <i>All Parenting States (N = 361)</i> |      |                     |      |                 |      |                                 |      |                              |                                 |
| Male                                  | 29.8 | Under 20            | 8.1  | White           | 22.3 | Married                         | 41.5 | 1.7<br>(1.0)                 | 9.7                             |
| Female                                | 70.2 | 20-29               | 10.9 | Black           | 12.7 | Never married                   | 23.6 |                              |                                 |
|                                       |      | 30-39               | 48.7 | Hispanic        | 54.2 | Single, but living with someone | 18.9 |                              |                                 |
|                                       |      | 40-49               | 26.2 | Native American | 8.7  | Separated                       | 1.3  |                              |                                 |
|                                       |      | 50-59               | 5.8  | Other           | 2.2  | Divorced                        | 7.9  |                              |                                 |
|                                       |      | 60-69               | 0.3  |                 |      | Remarried                       | 1.3  |                              |                                 |
| <i>Washington (N = 250)</i>           |      |                     |      |                 |      |                                 |      |                              |                                 |
| Male                                  | 36.4 | 20-29               | 5.6  | White           | 7.8  | Partner/Spouse                  | 89.0 | 1.4<br>(1.6)                 | 11.9                            |
| Female                                | 63.6 | 30-39               | 57.2 | Black           | 0.9  | No Partner/Spouse               | 11.0 |                              |                                 |
|                                       |      | 40-49               | 30.4 | Hispanic        | 77.1 |                                 |      |                              |                                 |
|                                       |      | 50-59               | 6.8  | Native American | 12.4 |                                 |      |                              |                                 |
|                                       |      |                     |      | Other           | 1.8  |                                 |      |                              |                                 |
| <i>Tennessee (N = 51)</i>             |      |                     |      |                 |      |                                 |      |                              |                                 |
| Male                                  | 1.2  | Under 20            | 56.9 | White           | 29.4 | Married                         | 19.6 | 1.6<br>(0.83)                | 2.9                             |
| Female                                | 98.0 | 20-29               | 31.4 | Black           | 70.6 | Never married                   | 43.1 |                              |                                 |
|                                       |      | 30-39               | 11.8 |                 |      | Single, but living with someone | 29.4 |                              |                                 |
|                                       |      |                     |      |                 |      | Divorced                        | 7.8  |                              |                                 |
| <i>Iowa (N = 34)</i>                  |      |                     |      |                 |      |                                 |      |                              |                                 |
| Male                                  | 21.2 | 20-29               | 9.1  | White           | 72.7 | Married                         | 42.4 | 3.2<br>(1.3)                 | 10.2                            |
| Female                                | 78.8 | 30-39               | 36.4 | Black           | 3.0  | Never married                   | 6.1  |                              |                                 |
|                                       |      | 40-49               | 45.5 | Hispanic        | 21.2 | Single, but living with someone | 12.1 |                              |                                 |
|                                       |      | 50-59               | 6.1  | Other           | 3.0  | Separated                       | 9.1  |                              |                                 |
|                                       |      | 60-69               | 3.0  |                 |      | Divorced                        | 27.3 |                              |                                 |
|                                       |      |                     |      |                 |      | Remarried                       | 3.0  |                              |                                 |
| <i>Alaska (N = 25)</i>                |      |                     |      |                 |      |                                 |      |                              |                                 |
| Male                                  | 40.0 | 20-29               | 24.0 | White           | 76.2 | Married                         | 90.9 | 2.3<br>(1.3)                 | 8.9                             |
| Female                                | 60.0 | 30-39               | 56.0 | Black           | 9.5  | Never married                   | 4.5  |                              |                                 |
|                                       |      | 40-49               | 12.0 | Native American | 4.8  | Single, but living with someone | 4.5  |                              |                                 |
|                                       |      | 50-59               | 8.0  | Other           | 9.5  |                                 |      |                              |                                 |

Parenting subscales from the ITPB (Spoth et al., 1995) are shown in Table 2. Scale items are presented by subscale (e.g., communication, involvement, and anger management). Cronbach's alphas are presented on each subscale for both pre- and posttests (Cronbach, 1951; Cronbach & Shavelson, 2004). Scale reliability is comparable to previous studies using these measures (Spoth et al., 1995).

**Table 2. ITPB Parenting Scale Items**

|  | <b>Scale Reliability</b> |
|--|--------------------------|
| <i>Communication Subscale</i>  |                          |
| -I have discussed my child's goals and dreams with him/her on several occasions.   | Pretest $\alpha = 0.82$  |
| -I often tell my child how I feel when he/she misbehaves.  | Posttest $\alpha = 0.80$ |
| -When my child tells me something important, I let him/her know that I am trying to understand what he/she is feeling.               |                          |
| -I let my child know I care about him/her while setting limits and consequences.   |                          |
| -I have discussed our family values with my child on several occasions.  |                          |
| <i>Involvement Subscale</i>  |                          |
| -I find ways to keep my child involved in fun activities with our family.  | Pretest $\alpha = 0.90$  |
| -I find ways to keep my child involved in family work activities (e.g., chores).   | Posttest $\alpha = 0.70$ |
| -I find ways to keep my child involved in family decisions about fun and work activities in a manner appropriate for his/her age.    |                          |
| <i>Anger Management Subscale</i>   |                          |
| -I am able to control my anger and frustration with my child.  | Pretest $\alpha = 0.70$  |
| -I work hard with my child on ways to express and control his/her anger and frustration.   | Posttest $\alpha = 0.50$ |
| Response Set: 1 = <i>Strongly Disagree</i> ; 2 = <i>Disagree</i> ; 3 = <i>Neutral</i> ; 4 = <i>Agree</i> ; 5 = <i>Strongly Agree</i> |                          |

Items from the Parent Education Survey (PEPWT; Cornell Cooperative Extension, 2009) were responded to on the follow metrics as shown in Table 3.

**Table 3. PEPWT Items**

| <b>Item Name</b> | <b>Item Detail</b>   | <b>Response Set</b>  |
|------------------|--|--|
| No Patience*     | I do not have as much patience with my child(ren) as I should.                   | 1 = <i>Strongly Disagree</i> ; 2 = <i>Disagree</i> ; 3 = <i>Neutral</i> ; 4 = <i>Agree</i> ; 5 = <i>Strongly Agree</i> |
| Rules            | I try to make rules that take my child's individual needs into consideration.    | 1 = <i>Strongly Disagree</i> ; 2 = <i>Disagree</i> ; 3 = <i>Neutral</i> ; 4 = <i>Agree</i> ; 5 = <i>Strongly Agree</i> |
| Skills           | I honestly believe that I have the skills necessary to be a good caregiver.      | 1 = <i>Strongly Disagree</i> ; 2 = <i>Disagree</i> ; 3 = <i>Neutral</i> ; 4 = <i>Agree</i> ; 5 = <i>Strongly Agree</i> |
| Reason Rules     | I try to explain the reasons for the rules that I make. Would you say this is... | 1 = <i>Definitely True</i> ; 2 = <i>Sort of True</i> ; 3 = <i>Sort of False</i> ; 4 = <i>Definitely False</i>          |

\*Reverse scored

Table 4 presents results of a *t*-test analysis for paired samples for subscales from the ITPB (Spath et al., 1995). Subscale change was computed by calculating the difference between the mean of a subscale at Time 1 (pretest) and the mean of that same subscale at Time 2 (posttest). Paired *t*-tests were used as indicated since participants were the same at pretest and posttest. These results indicated that there was significant change in the desired direction in all three subscale means from Time 1 to Time 2.

**Table 4. ITPB Aggregated Parenting Scales Paired Samples *t*-tests (T1-T2)**

| Scale Name       | Time 1        | Time 2        | <i>df</i> | <i>t</i> |
|------------------|---------------|---------------|-----------|----------|
|                  | <i>M (SD)</i> | <i>M (SD)</i> |           |          |
| Communication    | 3.4 (1.1)     | 5.3 (0.6)     | 189       | 20.2***  |
| Involvement      | 3.5 (1.1)     | 5.2 (0.7)     | 192       | 17.0***  |
| Anger Management | 3.6 (1.0)     | 5.0 (0.7)     | 189       | 15.3***  |

Table 5 presents the results of a *t*-test analysis for paired samples for items from the PEPWT. Item change was assessed in the same way as subscale change, calculating the difference between the mean of an item at Time 1 (pretest) and the mean of that same item at Time 2 (posttest). Participants were paired at Time 1 and Time 2. These results indicated that although the change was in the expected direction for items “no patience” and “rules,” the change was not significantly different between Time 1 and Time 2. The results for the items “skills” and “reason rules” showed significant change, but opposite of the expected direction.

**Table 5. PEPWT Items Paired Sample *t*-tests (T1-T2)**

| Item Name    | Time 1        | Time 2        | <i>df</i> | <i>t</i> |
|--------------|---------------|---------------|-----------|----------|
|              | <i>M (SD)</i> | <i>M (SD)</i> |           |          |
| No Patience  | 3.0 (1.1)     | 3.2 (1.1)     | 118       | n.s.     |
| Rules        | 4.5 (1.1)     | 4.2 (1.3)     | 119       | n.s.     |
| Skills       | 4.9 (0.9)     | 4.3 (1.5)     | 119       | -4.2***  |
| Reason Rules | 2.2 (1.2)     | 1.5 (0.9)     | 119       | -7.2***  |

Tables 6 and 7 and Figure 1 show the results of a one-way ANOVA analysis for total number of hours a participant spent attending a program of low, medium, or high dosage of the intervention based on the ITPB (Spath, et al., 1995). The grouping variable was used in a one-way ANOVA to examine whether or not there were group differences in subscales at either Time 1 or Time 2 (Table 4). By examining the group differences at three points of measurement, Time 1, Time 2 and change during entire program, we can see not only where groups differ at the start and conclusion of programming, but also the amount of change yielded based on program dosage.

Table 6 shows the group differences at Time 1 and Time 2. Significant differences in the desired direction were detected in two of the three subscales at Time 2, while none were detected at Time 1.

**Table 6. ITPB Parenting Scales ANOVA by Total Hours in Program in Low, Medium, High Groupings at Time 1 and Time 2**

| Scale                   | Time 1   |           |                | Time 2   |           |                |
|-------------------------|----------|-----------|----------------|----------|-----------|----------------|
|                         | <i>M</i> | <i>SD</i> | <i>F</i> Value | <i>M</i> | <i>SD</i> | <i>F</i> Value |
| <i>Communication</i>    |          |           |                |          |           |                |
| Low                     | 3.3      | 1.2       | 1.5            | 5.1      | 1.2       | 3.8*           |
| Medium                  | 3.0      | 0.8       |                | 5.5      | 0.4       |                |
| High                    | 3.3      | 1.1       |                | 5.4      | 0.7       |                |
| <i>Involvement</i>      |          |           |                |          |           |                |
| Low                     | 3.5      | 1.1       | 2.8            | 5.0      | 1.1       | 3.3*           |
| Medium                  | 3.0      | 0.8       |                | 5.4      | 0.6       |                |
| High                    | 3.4      | 1.1       |                | 5.3      | 0.7       |                |
| <i>Anger Management</i> |          |           |                |          |           |                |
| Low                     | 3.7      | 1.1       | 2.8            | 5.0      | 0.8       | 0.3            |
| Medium                  | 3.3      | 0.8       |                | 5.1      | 0.6       |                |
| High                    | 3.5      | 1.0       |                | 5.0      | 0.6       |                |

Number of Hours in Program Groupings: Low = 9 or fewer hours, Medium = 10-12 hours, High = 13 hours or more.  $N = 149$ ; \* $p < .05$

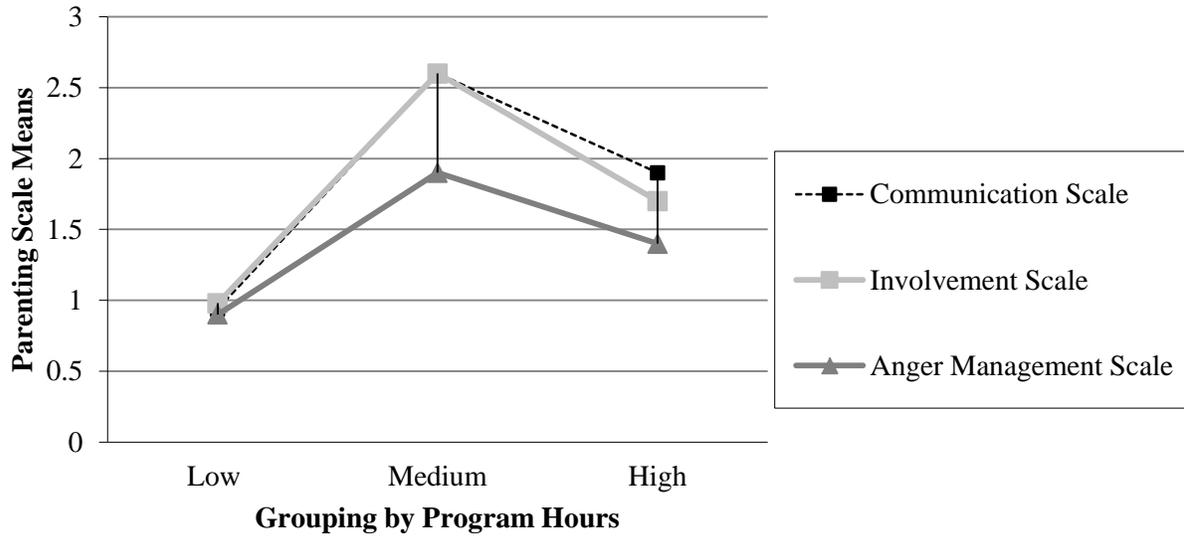
Table 7 shows the results of a one-way ANOVA examining the total amount of change in scores from Time 1 to Time 2 for each scale by total number of hours in the program (e.g., low, medium, high) on the ITPB (Spoth et al., 1995). All subscales showed significant differences between groups. The direction of the differences is shown visually in Figure 1.

**Table 7. ITPB Parenting Scales Change ANOVA by Total Hours in Program in Low, Medium, High Groupings for Entire Program Duration**

| Scale                   | <i>F</i> Value | Group Difference               | Direction of Difference<br>(See Figure 1) |
|-------------------------|----------------|--------------------------------|---|
| <i>Communication</i>    | 13.3**         | Low & Medium;<br>Low & High    | Low < Medium > High                       |
| <i>Involvement</i>      | 11.4**         | Low & Medium;<br>Medium & High | Low < Medium > High                       |
| <i>Anger Management</i> | 3.5*           | Low & Medium                   | Low < Medium > High                       |

Total Number of Hours in Program Groupings: Low = 9 or fewer hours, Medium = 10-12 hours, High = 13 hours or more.  $N = 149$ ; \* $p < .05$ , \*\* $p < .01$

**Figure 1. Total Number of Hours in Program in Groups by ITPB Scale Group Means**



Tables 8 and 9 show the results of a one-way ANOVA examining change between Time 1 and Time 2 in parenting skills by total hours in the program (using the low, medium, high grouping variable) on the PEPWT (Cornell Cooperative Extension, 2009). Again, the grouping variable was used in a one-way ANOVA to examine whether or not there were group differences in subscales at either Time 1 or Time 2 (Table 5).

Table 8 (next page) shows group differences at Time 1 and Time 2. Table 8 shows results of a one-way ANOVA examining parenting skills at Time 1 and Time 2 based on hours in programming using the PEPWT (Cornell Cooperative Extension, 2009). Results indicated significant difference between groups on patience at Time 1, but not Time 2. Further results indicated that “rules” and “skills” differed by hours in program at Time 2, but not at Time 1. In addition, “reason for the rules” showed differences in groups at both Time 1 and Time 2.

**Table 8. PEPWT Items ANOVA by Total Hours in Program in Low, Medium, High Groupings at Time 1 and Time 2**

| Items               | Time 1   |           |                | Time 2   |           |                |
|---------------------|----------|-----------|----------------|----------|-----------|----------------|
|                     | <i>M</i> | <i>SD</i> | <i>F Value</i> | <i>M</i> | <i>SD</i> | <i>F Value</i> |
| <i>No Patience</i>  |          |           |                |          |           |                |
| Low                 | 3.6      | 1.3       | 3.9*           | 3.4      | 1.6       | 1.3            |
| Medium              | 2.8      | 1.1       |                | 3.7      | 1.0       |                |
| High                | 3.5      | 1.2       |                | 3.0      | 1.0       |                |
| <i>Rules</i>        |          |           |                |          |           |                |
| Low                 | 4.8      | 1.0       | 1.0            | 4.5      | 1.4       | 7.8***         |
| Medium              | 5.1      | 0.9       |                | 2.9      | 0.8       |                |
| High                | 4.7      | 1.2       |                | 3.5      | 1.4       |                |
| <i>Skills</i>       |          |           |                |          |           |                |
| Low                 | 5.2      | 0.9       | 0.3            | 4.9      | 1.4       | 12.4***        |
| Medium              | 5.3      | 0.7       |                | 2.7      | 1.0       |                |
| High                | 5.1      | 1.2       |                | 3.6      | 1.5       |                |
| <i>Reason Rules</i> |          |           |                |          |           |                |
| Low                 | 2.3      | 1.5       | 5.1**          | 2.3      | 1.8       | 6.4**          |
| Medium              | 1.5      | 0.6       |                | 1.2      | 0.4       |                |
| High                | 1.6      | 0.6       |                | 1.4      | 0.6       |                |

Number of Hours in Program Groupings: Low = 9 or fewer hours, Medium = 10-12 hours, High = 13 hours or more.  $N = 149$ ; \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

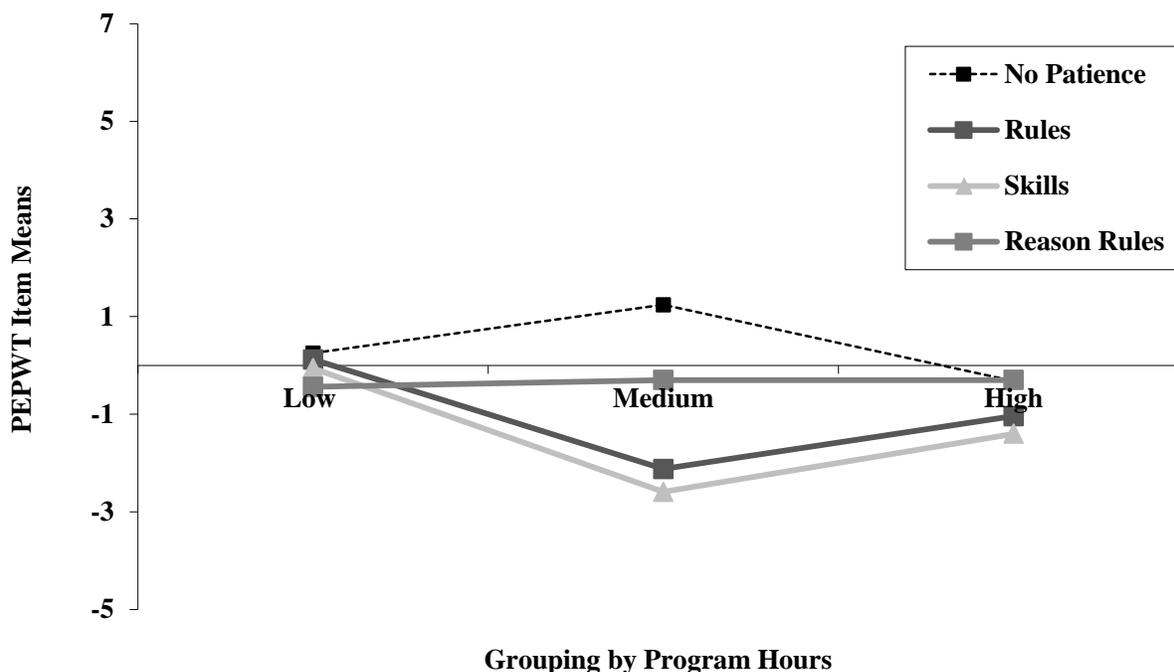
Table 9 shows results of a one-way ANOVA examining the change in parenting skills between Time 1 and Time 2 on parenting skills using the PEPWT (Cornell Cooperative Extension, 2009). Results indicated significant change on the skill of patience, setting age appropriate rules, and the belief in having skills to be an effective caregiver; however, the changes were not in the anticipated or desired direction (further discussion in the limitations section). The direction of the differences is shown visually in Figure 2.

**Table 9. PEPWT Change ANOVA by Total Hours in Program in Low, Medium, High Groupings for Entire Program Duration**

| Item                | <i>F Value</i> | Group Difference                             | Direction of Difference<br>(See Figure 2) |
|---------------------|----------------|--|---|
| <i>No Patience</i>  | 5.8**          | Medium & High                                | Low < Medium > High                       |
| <i>Rules</i>        | 10.6***        | Low & Medium; Low & High;<br>Medium & High   | Low > Medium < High                       |
| <i>Skills</i>       | 12.8***        | Low & Medium; Low and High;<br>Medium & High | Low > Medium < High                       |
| <i>Reason Rules</i> | 0.2            | None Significant                             | Low < Medium > High                       |

Total Number of Hours in Program Groupings: Low = 9 or fewer hours, Medium = 10-12 hours, High = 13 hours or more.  $N = 59$ ; \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

**Figure 2. Total Number of Hours in Program in Groups by PEPWT Means**



**Effect Size**

One way to examine the strength of the statistical results pertaining to intervention outcomes is to look at effect size. In examining the effect size of change subscale variables, it was found that all ITPB subscales resulted in large effect sizes (as indicated by a Cohen’s *d* of 0.8 or greater) (Cohen, 1988). This suggests that these programs were having a large effect on participant outcomes, in this case learning about parenting communication, anger management and involvement. PEPWT results also showed large effect size change in skills and reasons for the rules that are made, but only small effects for patience and rules that take into consideration individual child needs.

Tables 10 and 11 show the size of the effect of change in the parenting subscales using ITPB (Spoth et al., 1995) and PEPWT (Cornell Cooperative Extension, 2009). Using Cohen’s *d*, we found that there was a large effect of change on participants over time.

**Table 10. Effect Size by Parenting Change Scales on the ITPB**

| Scale                         | <i>t</i> -value | <i>df</i> | Cohen’s <i>d</i> | Size of Effect |
|-------------------------------|-----------------|-----------|------------------|----------------|
| Communication Change Scale    | 20.2            | 188       | 2.9              | Large          |
| Involvement Change Scale      | 17.0            | 191       | 2.5              | Large          |
| Anger Management Change Scale | 15.3            | 188       | 2.2              | Large          |

Cohen’s *d*: 0.2-0.3 = small effect size, 0.5 = medium effect size, 0.8+ = large effect size

Table 11 shows the size of the effect of change on the skills and beliefs assessed using the Parenting Education Survey (Cornell Cooperative Extension, 2009). Using Cohen's *d*, we found a small effect of change in relation to participant's patience and creation of rules that take into account children's individual needs and a large effect of change in relation to their belief that they have the skills to be good caregivers and to create rules and explain the reasons for those rules.

**Table 11. Effect Size by Parenting Skill Change on PEPWT**

| Item              | <i>t</i> -value | <i>df</i> | Cohen's <i>d</i> | Size of Effect |
|-------------------|-----------------|-----------|------------------|----------------|
| No Patience       | 1.0             | 118       | 0.2              | Small          |
| Rules             | 1.9             | 119       | 0.3              | Small          |
| Skills            | 4.2             | 119       | 0.8              | Large          |
| Reasons for Rules | 7.2             | 119       | 1.3              | Large          |

Cohen's *d*: 0.2-0.3 = small effect size, 0.5 = medium effect size, 0.8+ = large effect size

### Implications for Evaluation

This pilot study showed some very interesting results in terms of dosage (number of hours spent in the program) as it predicts change from initial time in program to the conclusion of the program. We found that participants appear to reach a point of diminishing returns, at least in the current sample, after 10-12 hours of programming on the ITPB. This might suggest that more is not necessarily better, at least for those in this sample. Participants may reach a plateau when learning levels off and actually decreases after receiving more than 12 hours of program content. While there was no significant difference in the three ITPB subscale items among low, medium, and high level participants at Time 1, there were differences detected at Time 2. Furthermore, we checked if the participants in the groups represented a particular program, and found that they did not. More research on the optimal dosage of parenting classes needs to be conducted, but the results of this pilot study suggest that more content does not always equal more learning.

Results from the PEPWT also indicate a decrease in the knowledge related to the skill of being an effective caregiver, making rules that account for individual differences in children, and being able to explain the reasons for the rules that are made. There are two possible explanations for these results: (1) participants did not, in fact, benefit in these specific domains from the program; or (2) at the onset of the program, Time 1, participants overestimated their skills, and throughout the course of the program, they learned that they did not, in fact, engage these behaviors as often as previously thought and adjusted their responses at Time 2 to reflect a more knowledgeable assessment of their behavior and skills. It has been shown that pretest/posttest methodologies have resulted in underestimation of outcomes in the evaluation of parenting programs (Pratt, McGuigan, & Katzev, 2000). It is well known in the area of divorce that parents often overestimate their ex-spouses' parenting skills (Madden-Derdich & Leonard, 2002). While the

present evaluation used a pre/post methodology at two of the sites and a retrospective methodology at two others, perhaps the phenomenon noted in prior research is occurring here; thus, the effect is not related to dosage. Further research is needed to explore possible reasons why parent-perceived efficacy in these domains declined with educational programming. For example, in the future, we intend to explore the difference between sites that utilized a pre/post format compared to those who used a truly retrospective format.

This pilot study suggests that it is important to pay particular attention to dosage participants received, as there is a satiation threshold, which needs to be considered not only in evaluation but program design and development.

### References

- Baker, P., & Mott, F (1989). *NLSY Child Handbook 1989*. Columbus, OH: Center for Human Resources Research, The Ohio State University.
- Braverman, M. T., & Engle, M. (2009). Theory and rigor in Extension program evaluation planning. *Journal of Extension*, 47(3), Article 3FEA1. Retrieved from [http://www.joe.org/joe/2009june/pdf/JOE\\_v47\\_3a1.pdf](http://www.joe.org/joe/2009june/pdf/JOE_v47_3a1.pdf)
- Burke, R. V., & Herron R. W. (1992). *Common sense parenting: A practical approach from Boys Town*. Boys Town, NE: Father Flanagan's Boys' Home.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cornell Cooperative Extension. (2009). *Parent Education Survey*. Retrieved from: [https://cyfernetsearch.org/sites/default/files/psychometricsFiles/Parent%20Education%20Survey%20\(parents%20of%20children%20all%20ages\)%20\(cornell%20extension\)\\_0.pdf](https://cyfernetsearch.org/sites/default/files/psychometricsFiles/Parent%20Education%20Survey%20(parents%20of%20children%20all%20ages)%20(cornell%20extension)_0.pdf)
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. doi:10.1007/BF02310555
- Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64(3), 391–418. doi:10.1177/0013164404266386
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23–45. doi:10.1016/S0272-7358(97)00043-3
- Duerden, M. D., & Witt, P. A., (2012). Assessing program implementation: What it is, why it's important, and how to do it. *Journal of Extension*, 50(1), Article 1FEA4. Retrieved from <http://www.joe.org/joe/2012february/a4.php>
- Kumpfer, K. L., & DeMarsh, J. P. (1983). *Strengthening Families Program: Parent training curriculum manual*. Salt Lake City, UT: Social Research Institute, Graduate School of Social Work, University of Utah.

- Madden-Derdich, D. A., & Leonard, S. A. (2002). Shared experiences, unique realities: Formerly married mothers' and fathers' perceptions of parenting and custody after divorce. *Family Relations, 51*(1), 37–45.
- Payne, P. B., & McDonald, D. A. (2012). Using common evaluation instruments across multi-state community programs: A pilot study. *Journal of Extension, 50*(4), Article 4RIB2. Retrieved from [http://www.joe.org/joe/2012august/pdf/JOE\\_v50\\_4rb2.pdf](http://www.joe.org/joe/2012august/pdf/JOE_v50_4rb2.pdf)
- Phelan, T. W. (1996). *1-2-3 Magic: Effective discipline for children 2-12* (2nd ed.). Glen Ellyn, IL: Child Management.
- Pratt, C. C., McGuigan, W. M., & Katzev, A. R. (2000). Measuring program outcomes: Using retrospective pretest methodology. *American Journal of Evaluation, 21*(3), 341–349. doi:10.1016/S1098-2140(00)00089-8
- Spoth, R., Redmond, C., Haggerty, K., & Ward, T. (1995). A controlled parenting skills outcome study examining individual difference and attendance effects. *Journal of Marriage and the Family, 57*(2), 449–464. doi:10.2307/353698

*Pamela B. Payne* is an Assistant Professor of Child and Family Studies at Weber State University in Ogden, Utah. She completed her Ph.D. in Family Studies and Human Development in 2010 at the University of Arizona.

*Daniel A. McDonald* is an Associate Agent with Pima County Cooperative Extension at the University of Arizona in Tucson, Arizona. He completed his Ph.D. in Family Studies and Human Development in 2000 at the University of Arizona.

### **Acknowledgements**

The authors would like to thank all the programs associated with this pilot study for their important and helpful contributions. This study was supported by the Children, Youth, and Families At Risk Program: Sustainable Community Projects with funding from the USDA National Institute of Food and Agriculture (NIFA Award #2008-41520-04810).