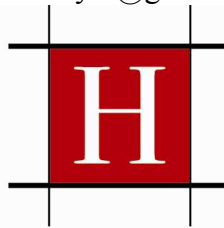


OUTCOME EVALUTION OF PUERTO RICAN FAMILY INSTITUTE'S
FATHERHOOD INVOLVEMENT GRANT ACTIVITIES

Prepared by:
ALLISON HYRA, Ph.D.
allison.hyra@gmail.com



H Y R A C O N S U L T I N G

For:
THE PUERTO RICAN FAMILY INSTITUTE

DO NOT CITE OR REPRODUCE WITHOUT AUTHOR PERMISSION

October 15, 2011

Acknowledgements: This evaluation was contracted by the Puerto Rican Family Institute and supported by ACF grant number #90FR0013. The author would like to thank all of the grant staff who helped shape this evaluation and report and who coded all of the data into spreadsheets. This list includes Marianela Cortes, Abigail Karic, Iris Ortiz, Luis A. Ortiz Jr., Yanet Rivas, and Vimary Serra. Although the aforementioned group provided feedback and critiques, all mistakes and errors are the author's.

ABSTRACT

In 2006, the Puerto Rican Family Institute was awarded a grant by the Administration for Children and Families to provide father involvement instruction to adult and adolescent fathers in the New York City area. The various curricula employed intended to increase knowledge about parenting, child development, healthy relationships, finances and employment, improve attitudes about the importance of fathers, appropriate discipline, marriage, employment and financial responsibility. This evaluation uses a single group, pre-/post-test design to test whether the program was able to achieve the aforementioned goals. Generally, the program demonstrated statistically significant changes in attitudes and knowledge in the desired direction, with one exception. The conflict resolution retreat had no measures that moved in the desired direction, and one measure showed statistically significant movement in the opposite direction. The results suggest that, overall, the program is achieving the stated goals.

Introduction

In 2005, the US Congress passed the Deficit Reduction Act which provided \$500 million dollars a year for five years to support healthy marriage and fatherhood involvement programming. The Puerto Rican Family Institute (PRFI), a multi-program family service non-profit located in New York City, received one of the grants. The grant provides relationship education and fatherhood education to couples, fathers and adolescents throughout the New York City metropolitan area. This evaluation describes the outcomes associated with the fatherhood and adolescent relationship education components; a separate evaluation will document the couples' relationship education component.

PRFI brought together a team of family strengthening service providers to implement the grant activities. PRFI Brooklyn and PRFI Jersey City provided the 24/7 Dad™ curriculum to men who were fathers (and a small subset of women who were interested in the services). Hoboken University Medical Center's Community Mental Health Center (Hoboken), which provides counseling and behavioral health services as part of Hudson County's largest health system, taught two components of the Love U2 adolescent relationship curriculum, Communication Smarts and Sex Smarts, to parenting and expectant adolescent boys at high schools, drug treatment centers and juvenile detention facilities on a voluntary basis. The Committee for Hispanic Children and Families (Committee), a non-profit that provides prevention services to Hispanic families to improve children's lives, taught a variety of voluntary programming to parenting and expectant adolescent male high school students, including Love U2's Relationship Smarts, Sex Smarts and Baby Smarts; Professional Development; Financial Literacy; a Conflict Resolution Retreat; and Child Discipline.

Curricula and Service Delivery

The 24/7 Dad™ curriculum was developed by National Fatherhood Initiative and is designed to provide men the skills and information they need to be great dads twenty-four hours a day, 7 days a week. Topics include family history, discipline, co-parenting, showing emotions, anger management and communication skills. PRFI teaches the curriculum as a 16 hour program, using various dosage schedules (i.e., over 3 days, or over 7 weekly sessions).

Love U2 is a values-based relationship curriculum for adolescents. Marlene Pearson developed the curriculum and it is available through the Dibble Institute. The sub-grantees in this evaluation are teaching different components of the program. Relationship Smarts focuses on self-awareness, future orientation, intimacy, trust, signs of healthy and unhealthy relationships, romantic attractions and the building blocks of love. Sex Smarts explores the social and emotional aspects of sexuality and the benefits of delaying sexual activity, including information on sexually transmitted infections (STIs). Baby Smarts teaches about infant and young child developmental needs and how

the quality and stability of the parental relationship is related to child outcomes. Communication Smarts is based on the adult relationship curriculum PREP and imparts communication and conflict resolution skills. Hoboken teaches Communication and Sex Smarts at 8-12 hour programs each, usually providing one hour of service per week. Committee teaches Relationship, Sex and Baby Smarts for 15 hours each.

The Professional Development curriculum was developed by Hands On Professional Insight Learning, Inc. The curriculum aims to prepare participants to complete an effective job search and secure employment. Program activities include helping participants discover their career interests, develop resumes and cover letters, prepare for a job interview and negotiate a job offer. The program is offered in four, 1.5 hour settings for a total of six curriculum hours.

The Financial Literacy curriculum was also developed by Hands On Professional Insight Learning, Inc. The program educates participants on both personal financial decisions and the larger financial system. Specific goals include learning about the origin of money; the banking system and the differences between credit unions and banks; and individual capacity to purchase goods and services. Committee teaches the Financial Literacy program in three, hour and a half sessions for a total of 4.5 hours.

The Conflict Resolution Retreat was developed by Ramapo for Children, a not-for-profit organization with the mission to give children the chance to experience success, develop healthy relationships and learn life skills. Ramapo for Children has a large campus in upstate New York where they provide retreats for school-aged children. The Conflict Resolution Retreat is taught by Ramapo staff and is designed to build trust and support among participants and teach communication and problem solving techniques. The retreat takes place over a two-day weekend, with 9 hours of instruction each day.

The Child Discipline program was developed in-house by Committee. The curriculum is designed to educate participants about inappropriate discipline techniques and definitions of child abuse and neglect. The curriculum also describes indicators of physical abuse, cultural differences in child discipline, shaken baby syndrome and ways to report suspected child abuse or neglect. The program is a total of two hours, taught over two classroom sessions.

Evaluation Design and Evaluation Questions

This evaluation has a single group pre-/post-test design, meaning that only the program participants were surveyed and data was collected using the same instrument at the start and completion of program activities. Since there was no comparison group that completed the evaluation forms but did not participate in the program, there can be no test of program effectiveness. Thus the evaluation cannot answer the “big” question, “Did it work?” This evaluation cannot discern whether participants became better fathers or made smarter relationship decisions. Instead, this evaluation focuses on what, in logic model terminology, is called “immediate outcomes.” Immediate outcomes are the things that should have changed as direct participation in the program and evident by program

completion. These things include increase in knowledge, and change in attitudes, opinions or behavioral intent (e.g., I plan to do X in the future).

This evaluation answers the following questions for each program.

- How many people completed the program?
- What demographic information can we provide about program completers?
- Did participants show statistically significant improvements in: a) knowledge, b) attitudes or c) behavioral intention?
- What lessons have been learned for future relationship/fatherhood education research?

Methods

This evaluation reports on data collected by each sub-grantee using questionnaires that each sub-grantee developed. With the exception of 24/7 Dad™, measures were not drawn from any pre-packaged evaluations that might have been provided with these commercial curricula, nor were they selected from other research reports. Additionally, items used in this evaluation have not been validated or cognitively tested.

Each site also followed their own procedures for distributing and collecting the pre- and post-tests. Most provided the pre-tests before the start of the curriculum on the first day of the program and the post-tests on the last day after the close of instruction. Program staff was in charge of receiving the completed forms and inputted the responses to selected items into a spreadsheet designed by the evaluation author.

Data were not coded for participants that did not complete the program or did not complete the evaluation. Thus demographic data captures information on program completers only. The data in this evaluation report was collected between 2007 and 2011.

Analyses were conducted using a two-tailed t-test with an alpha = .05 significance level. The report provides pre- and post-test mean values for all tests. Only those that are designated statistically significant should be interpreted as an actual shift between pre- and post-test means. The other measures should be interpreted as not having changed between the two waves of data collection; in other words the difference between the two means is zero.

Findings

The findings for each program provided by each site are provided below.

Hoboken University Medical Center Communication Smarts

Hoboken Communication Smarts had 163 completed pre- and post-tests. Average participant age was 16.3. The evaluation probed three aspects of communication knowledge: using a time-out during a discussion, effective ways to complain, and using the speaker-listener technique (see Table 1). Participants showed a statistically significant improvement in each measure. On the pre-test, 34% of participants knew how to use a time out; 88% knew how to on the post-test (n = 161). Slightly over half (51%) of participants chose an effective way to complain on the pre-test, 83% did on the post-test (n = 161). Finally, 69% of participants correctly identified how to use the speaker-listener technique on the pre-test. That number increased to 93% on the post-test (n = 162).

Table 1. Hoboken Communication Smarts Results

Item	Pre-test % Correct	Post-test % Correct (n)
Using a “time-out”	34%	88% (161)*
Effective complaining	51%	83% (161)*
Speaker-listener technique	69%	93% (162)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Hoboken University Medical Center Sex Smarts

Hoboken Sex Smarts had 94 completed pre and post tests. Average age was 17.0. Participants improved their knowledge about the lack of sexually transmitted infection (STI) symptoms; 45% of participants knew that not all STIs have symptoms before the program started; 98% did after (n = 93). Participants also improved the knowledge about the cause of cervical cancer – with a caveat. 46% of pre-tests that provided a response correctly identified HPV as the cause of cervical cancer; 98% of post-tests did (n = 56). As indicated, the sample for this question was 56, not 94. It was a fill-in-the-blank and 38 pre-tests had no response. It is likely that those were not answered because the participant did not know the correct response.

The evaluation also had two behavioral intention questions. One asked the likelihood that participants would check to see if their partner had protection when interested in sex, the second asked the extent to which participants felt comfortable saying no to sex. The response scales are 1 Never, 2 Sometimes and 3 Always. Thus, the program should drive post-test mean values higher than pre-test values. And in fact, post-test means were statistically significantly higher for both questions. The protection measure had a pre-test mean of 2.0 and a post-test mean of 2.4 (n = 91). The sex refusal measure had a pre-test mean of 2.0 and a post-test mean of 2.4 (n = 92).

Table 2. Hoboken Sex Smarts Results

Item (scale)	Pre-test Mean or % Correct	Post-test Mean or % Correct (N)
Symptom-free STI	45%	98% (93)*
Cervical cancer cause	46%	98% (56)*
Ask partner about protection (1-3)	2.0	2.4 (91)*
Refusal efficacy (1-3)	2.0	2.4 (92)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Committee for Hispanic Children and Families Professional Development

Sixty-three participants completed the pre- and post-tests for Committee’s professional development program. The mean participant age was 17.4. The evaluation asked participants if they had a resume and a cover letter. The number of participants that had either item showed a statistically significant increase from pre- to post-test. Only 14% of participants had resumes at pre-test and 100% did at post-test (n = 63). Additionally, 3% of participants had a cover letter at pre-test and 100% did at post-test (n = 63).

Table 3. Committee Professional Development Results

Item	Pre-test % Yes	Post-test % Yes (N)
Have a resume	14%	100% (63)*
Have a cover letter	3%	100% (63)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Committee for Hispanic Children and Families Financial Literacy

Fifty participants completed both pre- and post-tests for the financial literacy program. Their mean age was 17.8. The evaluation measured six aspects of financial knowledge: the definitions of fixed expenses, flexible expenses, and discretionary expenses; the definition of the acronym APR; aspects of a line of credit; and the definition of a credit report. Participants showed a statistically significant increase in knowledge in all six aspects. A third of participants (39%) correctly identified fixed expenses on the pre-test; all participants provided a correct answer on the post-test (n = 49). Participants provided a correct definition for flexible expenses 48% of the time on the pre-test and 98% of the time on the post-test (n = 50). Less than half (42%) of pre-tests correctly identified a discretionary expense, while 98% of post-tests did (n = 50). APR was correctly defined on 59% of pre-tests and all of the post-tests (n = 46). About half (54%) of participants knew about credit lines according to the pre-test; almost all (98%) did on the post-test (n = 48). Credit report knowledge change is a more complicated story, due to a large

number of missing responses. Only 34 participants completed the pre- and post-test credit report items. Among those 34, 32% answered the pre-test correctly while 76% had a correct post-test response.

Table 4. Committee Financial Literacy Results

Item	Pre-test % Correct	Post-test % Correct (N)
Fixed expense	39%	100% (49)*
Flexible expense	48%	98% (50)*
Discretionary expense	42%	98% (50)*
APR	59%	100% (46)*
Credit lines	54%	98% (48)*
Credit report	32%	76% (34)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Committee Conflict Resolution Retreat

Seventy-six participants completed the conflict resolution retreat pre- and post-test forms. Average participant age was 17.7. The evaluation probed two attitudes: the extent to which conflict is perceived as negative and the extent to which the participant believes that only his needs are important. The needs measure fell just short of statistical significance, meaning that the pre- and post-test scores are considered statistically equal – there was no change in that attitude ($n = 76$). Negative conflict attitudes did change, however. This measure (conflict is a negative experience) had five response categories – Almost never (1), Occasionally (2), Half of the time (3), Usually (4) and Almost always (5). Theoretically, since the program is designed to teach that conflict is an inevitable part of human interaction and can be dealt with in a calm and respectful manner, post-test score should decrease. However, the mean pre-test score was 3.4 while the mean post-test score was 4.1 ($n = 74$). In other words, participants increased in the frequency with which they felt conflict was negative.

Table 5. Committee Conflict Resolution Retreat Findings

Item (Scale 1-5)	Pre-test Mean	Post-test Mean (N)
Conflict is negative	3.4	4.1 (74)*
Only my needs important	3.0	2.9 (76)

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Committee Child Discipline Program

Seventy-six participants completed the pre- and post-tests for Committee’s Child Discipline program. Their average age was 17.7. The evaluation asked participants to rank their knowledge on: (1) the effects of child abuse on infants and toddlers; (2) how to handle possible abuse or neglect concerns; and (3) the causes of challenging

infant/toddler behavior. Response options were: 3- High, 2 – Medium, or 1 - Low levels of knowledge. Participants showed statistically significant increases in their subjective evaluation of their child abuse and behavior knowledge. The average pre-test child abuse knowledge score was 1.2 and the average post-test score was 2.3 (n = 76). The average pre-test score for knowledge about ways to handle suspected abuse was 1.2; the average post-test values for the same measure was 2.2 (n = 75). Finally the average pre-test score for knowledge about child behavior was 1.2 and the average post-test was 2.4 (n = 73).

Table 6. Committee Child Discipline Program Results

Item (Scale 1-3)	Pre-test Mean (N)	Post-test Mean (N)
Child abuse effects	1.2	2.3 (76)*
Handling abuse/neglect	1.2	2.2 (75)*
Child behavior causes	1.2	2.4 (73)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Committee Love U2 Relationship Smarts

One hundred thirty-five adolescents completed the pre- and post-test forms for Committee’s Love U2 Relationship Smarts program. Their average age was 17.4. The evaluation asked respondents three knowledge questions. One asked at what age male humans are fully developed. The second defined harassment and the third item asked whether simply finding the right person would make the respondent happy. That measure is designed to test whether participants understand that relationships are work rather than perfect matches. Forty-four percent of respondents knew the correct age of the completion of male development on the pre-test while 82% responded correctly on the post-test (n = 130). Less than 40% of adolescents correctly identified behavior as harassment on the pre-test; 96% of respondents did on the post-test (n = 124). On the pre-test, 29% of respondents said that finding the right person is not the key to happiness, that number increased to 81% on the post-test (n = 129).

Table 7. Committee Relationship Smarts Results

Item	Pre-test % Correct	Post-test % Correct (N)
Male development	44%	82% (130)*
Harassment behavior	37%	96% (124)*
Right person makes one happy	29%	81% (129)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Committee Love U2 Sex Smarts

One hundred fifteen participants completed the pre- and post-tests for Committee’s Love U2 Sex Smarts program. Their average age was 17.4. The evaluation measured three aspects of the adolescents’ sexually transmitted infection knowledge and two showed

statistically significant improvement. One item asked what body part was affected by HPV; another probed for symptoms of a herpes infection; the third asked whether sexually transmitted diseases could also be transmitted from oral sex. Participants showed a statistically significant improvement on their knowledge of HPV. Less than 40 percent (35%) provided a correct answer on the pre-test; 100% did on the post-test (n = 109). On the pre-test, 55% knew that STIs could be transmitted orally while 69% provided a correct answer on the post-test (n = 107). The difference between herpes symptoms questions on the pre- and post-tests was not statistically significant.

Table 8. Committee Sex Smarts Results

Item	Pre-test % Correct	Post-test % Correct (N)
Body part affected by HPV	35%	100% (109)*
Oral STI transmission	55%	69% (107)*
Herpes symptom	22%	20% (109)

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Committee Baby Smarts

One hundred and thirty-eight participants completed the pre- and post-tests for Committee's Love U2 Baby Smarts program. Their average age was 17.4. The evaluation measured knowledge about the possible negative effects of single parenting and step-parenting using six items. The items measured knowledge about (1) incarceration rates of men raised in single versus married parent households; (2) childhood poverty rates by family type; (3) adolescent mother high school graduation rates; (4) rates of adolescent childbirth by family type; (5) rates of childhood behavioral problems by family type; and (6) emotional health and academic behaviors by family type.

All six measures show statistically significant improvement. About half (45%) of participants knew that men from single parent homes had higher incarceration rates at pre-test; 95% did at post-test (n = 137). On the pre-test, 34% of participants knew that single parent homes were more likely to be poor, while 85% did on the post-test (n = 135). Forty percent of participants knew that teen mothers have a hard time finishing school on the pre-test; this improved to 86% on the post-test (n = 133). A third of participants (35%) knew that children of single parents are more likely to become teen parents than children of married parents on the pre-test; 67% answered correctly on the post-test (n = 135). On the pre-test, 33% of participants correctly answered the item on rates of child behavior problems by family types; on the post-test, 82% did (n = 121). Finally, 30% of participants answered the emotional health and academic behavior question correctly on the pre-test and 84% did so on the post-test (n = 129).

Table 9. Committee Baby Smarts Results

Item	Pre-test % Correct	Post-test % Correct (N)
Incarceration by family type	45%	95% (137)*
Poverty by family type	34%	85% (135)*
Teen mom graduation rate	40%	86% (133)*
Generational teen parent transfer	35%	67% (135)*
Child behavior by family type	33%	82% (121)*
Emotional health/academic behavior by family type	30%	84% (129)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

24/7 Dad™

The 24/7 Dad™ program was provided at two different sites – PRFI Jersey City and PRFI Brooklyn. A total of 785 participants completed the pre- and post-tests; 623 in Brooklyn and 162 in Jersey City¹. While the program was intended for men, women were allowed to participate and 34 did (4% of participants). The mean participant age was 38.9. The youngest participant was 16 and the oldest was 85. Two-thirds (67%) of participants identified as Hispanic. Another 25% were African-American; 5% were white, 2% identified as “other” and less than 1% said they were Native American or Asian.

The evaluation probed numerous aspects of parenting. We collected information on parenting knowledge, parenting attitudes, communication knowledge, co-parenting/father-mother relationship attitudes, and gender role attitudes. All items were measured on a 5-point scale.

Parenting Knowledge

Eight items measured an aspect of parenting knowledge and seven out of the eight measures showed statistically significant improvement. The item that asked the extent to which the participant agreed that soft discipline makes for spoiled kids moved from a pre-test mean of 3.0 to post-test mean of 3.2 when 3 equals uncertain and 4 equals disagree (n = 775). An item that probed the extent to which the participant agreed that parental expectations affect child self-worth moved from a pre-test mean of 2.2 to a post-test mean of 2.0 where 3 is uncertain and 2 is agree (n = 773). The item capturing the extent the participant agreed that praising oneself in front of children models self-worth had a

¹ The evaluation also included 31 women who participated in the program. The program was not allowed to discriminate on the basis of gender. Separate analyses were not done for the male only group, but given the small number of women in the program (less than 4%), it is highly unlikely that the inclusion of women affected the robustness of evaluation results.

pre-test mean of 2.9 and a post-test mean of 2.3 where 3 equals uncertain and 2 equals agree (n = 769). A measure about whether fathers can raise children as well as mothers had a pre-test mean of 3.7 and a post-test mean of 3.9 where 3 is uncertain and 4 is disagree (n = 760). A measure concerning whether put one's self in their children's shoes is a good way to feel empathy for ones' children moved from a pre-test mean of 1.8 to a post-test mean of 1.6 where 2 is agree and 1 equals strongly agree (n = 773). Another question asked if there was an "ideal" dad. That measure had a pre-test mean of 3.1 and a post-test mean of 3.4 where 3 equals uncertain and 4 is disagree (n = 762). The final significant item questioned whether children should help make family rules and had a pre-test mean of 2.4 and a post-test mean of 2.0 where 2 equals agree and 3 equals uncertain (n = 764). A measure about the extent to which culture affects parenting did not show a significant change from pre- to post-test.

Table 10. 24/7 Dad™ Parenting Knowledge Results

Item (Scale 1-5)	Pre-test Mean	Post-test Mean (N)
Soft discipline spoils	3.0	3.2 (775)*
Expectations affect self-worth	2.2	2.0 (773)*
Self-praise models self-worth	2.9	2.3 (769)*
Dads are as good as moms	3.7	3.9 (760)*
Putting self in kids' shoes shows empathy	1.8	1.6 (773)*
"Ideal" dads exist	3.1	3.4 (762)*
Children should help make family rules	2.4	2.0 (764)*
Culture affects parenting	2.5	2.4 (757)

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Communication Knowledge

The evaluation had four measures of communication knowledge and all four had statistically significant changes from pre- to post-test. The measure whether talking about anger is a waste of time moved from 4.1 to 4.3 where 4 is disagree and 5 is strongly disagree (n = 778). One item asked if it was "okay to keep feelings inside." The pre-test mean was 4.1 and the post-test mean was 4.2 with 4 meaning disagree and 5 meaning strongly disagree (n = 759). The measure that probed agreement with the statement that some feelings are good and others bad improved from a pre-test mean of 2.3 to a post-test mean of 2.8 with 2 equaling agree and 3 equaling uncertain (n = 759). An item that asked whether using a punching bag was a good way to express anger failed had a pre-test mean of 3.2 and a post-test mean of 3.4 where 3 was uncertain and 4 was disagree (n=774).

Table 11. 24/7 Dad™ Communication Knowledge Results

Item (Scale 1-5)	Pre-test Mean	Post-test Mean (N)
Talking about anger wastes time	4.1	4.3 (778)*
OK to hold feelings in	4.1	4.2 (759)*
Feelings are good and bad	2.3	2.8 (759)*

Punch bag to express anger	3.2	3.4 (774)*
----------------------------	-----	------------

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Parenting Attitudes

There were 5 parenting attitude measures in this evaluation and 3 showed statistically significant change. An item asking whether dads need to push their children moved from 2.9 on the pre-test to 3.1 on the post test when 2 equals agree, 3 equals uncertain and 4 equals disagree (n = 774). Participants improved their responses to the item asking the extent to which they agreed that children need to learn that “dads don’t mess around.” The pre-test mean was 1.9 and the post-test was 2.2 with 1 meaning strongly agree, 2 meaning agree and 3 meaning uncertain (n = 77). Another item asked whether children need to know right from wrong using whatever technique necessary. This measure moved from 2.6 on the pre-test to a mean of 3.0 on the post where 2 equals agree and 3 is uncertain (n = 776). Two measures, whether harsh punishments show that dads “mean business” and whether fathers who “lay down the law” command respect failed to show any statistically significant movement.

Table 12. 24/7 Dad™ Parenting Attitudes Results

Item (Scale 1-5)	Pre-test Mean	Post-test Mean (N)
Dad should push kids	2.9	3.1 (774)*
Kids know dads don’t mess around	1.9	2.2 (777)*
Kids know right/wrong using necessary technique	2.6	3.0 (776)*
Harsh punishment shows one means business	3.8	3.8 (778)
Laying down law gets respect	3.2	3.3 (762)

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Co-Parenting/Relationship Attitudes

The evaluation probed co-parenting and relationship attitudes using 10 items and 9 of those showed statistically significant improvement. Two probed marriage and love between parents. Participants showed improvement on an item asking if the best thing a father can do is love his children’s mother, moving from a pre-test mean of 2.4 to a post-test mean of 2.3 where 3 equals uncertain and 2 equals agree (n = 764). Additionally, a measure questioning agreement with the statement “men are better off married” moved from a pre-test mean of 3.2 to a post-test mean of 2.8 where 4 is disagree, 3 is uncertain and 2 is agree (n = 771).

Five items measured attitudes about the extent to which mothers and mothering and fathers and fathering differ. The measure of whether fathering is more important than

mothering had a pre-test mean of 4.0 and a post-test mean of 4.2 where 4 is disagree and 5 is strongly disagree (n = 772). When asked whether fathers are better disciplinarians, participants had pre-test means of 3.6 and post-test means of 3.8 where 3 is uncertain and 4 is disagree (n = 779). A measure probing whether fathers need to be the head of the household moved from a pre-test mean of 2.8 to a post-test mean of 3.0 where 2 equals agree and 3 is uncertain (n = 772). One item stated that “fathers work; mothers take care of children.” This measure had a pre-test mean of 3.8 and a post-test mean of 4.0 where 3 equals uncertain and 4 equals disagree (n = 759). An item measuring the extent to which a participant believes that the roles of mothers and fathers are clearly different did not show a statistically significant change between pre- and post-test.

Three items looked at participant opinions on the connection between work and fatherhood. One measure probed the extent to which participants’ believed the major role of a father is to be a provider. The pre-test mean was 2.9 and the post-test mean was 3.1 where 2 is agree, 3 is uncertain and 4 is disagree (n = 771). Mean responses for an item measuring whether work is more important for men than family moved from 4.2 on the pre-test to 4.3 on the post-test where 4 is disagree and 5 is strongly disagree (n = 766). Finally, the item that stated “balancing work and family is more important for women than men” had a pre-test mean of 3.5 and a post-test mean of 3.6 where 3 is uncertain and 4 is disagree (n = 762).

Table 13. 24/7 Dad™ Co-parenting/Relationship Attitudes Results

Item (Scale 1-5)	Pre-test Mean	Post-test Mean (N)
Best for kids if dad loves mom	2.4	2.3 (764)*
Men are better off married	3.2	2.8 (771)*
Fathering more imp. than mothering	4.0	4.2 (772)*
Dads better disciplinarians	3.6	3.8 (779)*
Dad must be head of home	2.8	3.0 (772)*
Dads work; Moms care for kids	3.8	4.0 (759)*
Mom and dad roles are different	2.4	2.4 (759)
Dads need to be providers	2.9	3.1 (771)*
Work more imp than family for men	4.2	4.3 (766)*
Balance work/family more important for women	3.5	3.6 (762)*

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Gender Role Attitudes

Nine items measured gender role attitudes such as how men and boys should behave. Five of the measures showed statistically significant changes between pre-test and post-test means. The item stating that it is wrong for men to show their feelings in public moved from a pre-test mean of 3.5 to a post-test mean of 3.8 where 3 is uncertain and 4 is disagree (n = 770). Quoting an old Timex watch ad, one measure stated that men should be able to “take a licking and keep on ticking,” presumably capturing an attitude that men

should be strong and impermeable to defeat. This measure has a pre-test mean of 3.2 and a post-test mean of 3.5 where 3 is uncertain and 4 is disagree (n = 762). A very similar item measured the extent to which participants felt that men should be strong no matter what. That item moved from a mean of 2.7 on the pre-test to a mean of 3.1 on the post-test with 2 meaning agree, 3 meaning not sure and 4 meaning disagree (n = 765). Another item stated that boys should be taught to “take it like a man” and this measure had a pre-test mean of 3.2 and a post-test mean of 3.7 where 3 is uncertain and 4 is disagree (n = 772). The final, statistically significant measure probed whether the participant believed that men and women should have different careers. This item had a pre-test mean of 3.8 and a post-test mean of 3.9 where 3 is uncertain and 4 is disagree (n = 772).

Four gender role attitude items failed to show a statistically significant difference between pre- and post-test. The items measured: whether boys should be taught to internalize feelings; whether one acts like a man by following traditional gender roles; whether real men don’t cry; and whether girls raised by their fathers become tomboys.

Table 14. 24/7 Dad™ Gender Role Attitudes Results

Item	Pre-test Mean	Post-test Mean (N)
Male emotions in public wrong	3.5	3.8 (770)*
Men should take a licking and keep on ticking	3.2	3.5 (762)*
Men strong no matter what	2.7	3.1 (765)*
Boys should take it like men	3.2	3.7 (772)*
Men and women should have diff careers	3.8	3.9 (772)*
Boys should internalize feelings	4.2	4.3 (780)
Men follow traditional gender roles	3.2	3.3 (768)
Real men don’t cry	4.2	4.2 (767)
Girls raised by dads are tomboys	4.3	4.3 (776)

* indicates that the pre-test and post-test means were statistically significant different at the $\alpha = .05$ level.

Discussion

Overall, the various programs evaluated in this report demonstrated significant positive change in their participants. The statistically significant changes for each program are described below as are study limitations and lessons learned.

Findings Synthesis

Hoboken Medical University’s Love U2 Communication Smarts and Sex Smarts programs both increased participants’ knowledge. Specifically, participants in the Communication Smarts program showed strong improvements in their knowledge of how

to properly use a “time-out” during an argument. A large percent of Sex Smart participants learned that STIs do not always have symptoms. Sex Smarts also increased participant behavioral intent to ask partner for contraception and to refuse unwanted sex.

The Committee for Hispanic Children and Families implemented six different programs. Their professional development program armed all of its program graduates with cover letters and resumes. The financial literacy program increased participants’ financial knowledge; almost 100% of program graduates got every knowledge question correct on the post-test. In contrast, the conflict resolution program did not see changes in participant attitudes towards conflict. Actually, participants became more negative about conflict, rather than seeing conflict as a normal part of human interaction that can be handled respectfully and calmly.

The evaluation of Committee’s child discipline program showed that participants believed they had more knowledge about child discipline and child abuse after the program. These measures asked participants how much they knew and did not test for correct or incorrect responses to an actual factual question (e.g., children should be potty trained by 18 months old; true or false?). As such, the results cannot be interpreted to say the program increased knowledge (though it may have), but rather perception of knowledge.

Committee also taught Love U2’s Relationship Smarts, Sex Smarts and Baby Smarts. Relationship Smarts showed an increase in participant knowledge regarding sexual development and relationships and had a very large increase in the number of participants that correctly identified harassment. Baby Smarts also increased participant knowledge about differential child outcomes by family type (by almost 50% for most measures). Their Sex Smarts program only showed improvement on knowledge about the causes of HPV; two other STI measures did not show improvement.

Even though both Hoboken and Committee implemented the Sex Smarts program, we cannot compare results across the two sites because they used different evaluation tools. It would seem that Hoboken’s program was more successful because they showed improvement on more measures, but because they implemented different data collection instruments, it is not clear if Committee’s participants would have had similar gains using identical instruments. Conversely, the two organizations could be serving markedly different adolescents or may deliver the same program differently (e.g., setting or facilitators).

The 24/7 Dad™ evaluation tested the largest number of knowledge and attitude items (36) and had the most number of items that demonstrated statistically significant change (27). The number of positive immediate outcome changes does not mean that this program was more effective or successful than some of the other programs in this evaluation, just that more questions were asked so more were available for analysis. The analysis showed a positive increase in 7 of 8 parenting knowledge questions; 3 of 4 communication/emotion knowledge questions; 3 of 5 parenting/fathering attitude questions; 9 of 10 co-parenting or relationship attitude questions; and 5 of 9 gender roles

attitudes. The participants demonstrated change on most of these measures despite having relatively positive pre-test means.

Large gains were seen on a subset of statistically significant items. The item asking whether some feelings are good and some feelings are bad showed a large gain in communication/emotion knowledge. The item “children need to know right from wrong” using whatever technique necessary had a large pre- to post-test movement, but other “harsh parenting” measures did not show any statistically significant change. Since it would make sense that three, very similar measures would move in conjunction with one another, it is not clear whether the larger gains on one measure are a statistical artifact or whether it measures some unseen aspect of harsh parenting not captured by the other items.

In terms of marriage/co-parenting attitudes, the item stating that men are better off married (presumably compared to being single) showed large gains. Four gender role items also had large, positive movement: that it is wrong for men to show feelings in public; that they should take a licking and keep on ticking; that men should be strong no matter what; and that boys should “take it” like a man. The magnitude of change for the gender role attitude items suggests that 24/7 Dad™ may have the biggest impact in terms of educating and enlightening participants about various, new ways to demonstrate masculinity without subscribing to rigid, unforgiving and possibly damaging “machismo” gender roles.

National Fatherhood Initiative (NFI) developed the 24/7 Dad™ program and has conducted an evaluation of it. A report entitled, “24/7 Dad™ A.M. and 24/7 Dad™ P.M. Outcome Evaluation Results 2005-2006,” is available on their website and uses many of the same measures as employed in this evaluation². Although the paper had a small sample size of under 50 fathers, had no Hispanics (their sample was roughly half white and half Black), and used a more lenient 90% confidence interval rather than this evaluation’s more stringent 95% confidence interval, it is informative to compare this paper’s results to theirs.

The NFI report does not provide actual pre- and post-test means. It only identifies which measures showed a statistically significant difference making magnitude and starting mean comparisons impossible. Both evaluations showed positive, statistically significant movement on seven items measuring relationship attitudes, fathering attitudes, parenting knowledge and emotional/communication knowledge. The previous evaluation had two other statistically significant items that were not tested in the current evaluation. Since the previous evaluation documented some, but not all, of the items tested that did not show statistically significant gains, it is impossible to directly discern which items the current evaluation showed improvement on that the NFI one did not. Regardless, this current evaluation had significant results on over a dozen items that were not significantly different in the other report.

² <http://www.fatherhood.org/Document.Doc?id=49> accessed on July 18, 2011.

Why did this current evaluation have so many more positive findings? Several hypotheses may explain this difference. One, the current sample is much larger than the previous one (over 750 compared to under 50), so this analysis has more power to identify differences. Two, the current sample is demographically different than NFI's. This paper's participants are largely Hispanic, while the previous evaluation's participants were mostly Black and white. Three, it could be that the current evaluation's participants had lower pre-test scores and thus more room for improvement. We cannot test that hypothesis because the previous evaluation did not provide pre- or post-test values. Four, possibly the curriculum has been improved and strengthened between the two evaluations and PRFI is teaching a newer version. Finally, evaluation differences could be due to the way PRFI is implementing the curriculum. They may have better trained facilitators capable of delivering the curriculum in a way that resonates with the participants or they may deliver it over a dosage schedule that better supports learning and retention. Regardless of the reason for the difference in findings, it is clear that PRFI's implementation of 24/7 Dad™ was at least as effective regarding immediate outcome change as the previous evaluation.

Limitations

As with all research, this evaluation suffers from several methodological deficiencies that may affect the reliability of the findings. All items were developed by the sites, with the exception of 24/7 Dad™. Those items apparently came from the previously released evaluation. To the author's knowledge, no measures were taken from published evaluations or research or were cognitively tested. These may or may not have been strongly connected to the program's learning objectives and may not have been worded in the most understandable manner. Additionally, some programs changed their answering format during data collection. For example, Committee's financial literacy program's forms were originally fill-in the-blank; they were then changed to multiple choice. Finally, we did not collect quality information on program dropouts so we cannot say the extent to which program participants were representative of all program attendees.

Lessons Learned

During the course of this evaluation, several lessons were uncovered. More meaningful comparisons across similar programs at different sites could have been made had they used identical items. Future work should employ identical items when possible. Additionally, those items should come from past research or be cognitively tested to ensure that participants understand the items the same way as the author intended. Finally, all evaluations should include measures of knowledge change (some of the current program evaluations did not), not only because these items seemed to show the largest improvements, but also because knowledge change is an important aspect of change that should be documented.

Conclusion

This report describes the findings from an outcome evaluation of several curricula provided to adult fathers and adolescent parenting or expectant fathers. The evaluation employed a single group pre- and post-test design to answer the following questions: how many people finished the program; what demographic information we know about them (e.g., age and race); what improvements were shown in knowledge, attitudes or behavioral intention; and what lessons can be reported for future evaluations? All but one program has statistically significant improvements in immediate outcomes (Committee's Conflict Resolution retreat did not show changes in the desired direction). This evaluation suggests that the programs funded under this grant are achieving the desired goals and educating fathers on important ways to improve their parenting and their relationships with their children.