

Grant # 5R42HD051184-03

First Three Years Online: Improving Access to Early Parent Education and Support

IRIS Educational Media and Oregon Research Institute

7/10/2009 to 6/30/2013

Principal Investigators: Carol Metzler, PhD and Laura Backen Jones, PhD

Summary:

This project developed and evaluated a three-part online program to assist parents of infants, one-year olds and two-year olds in promoting their children's development. IRIS Educational Media, Oregon Research Institute and Parenting Now collaborated to develop the "First 3 Years Online" (F3Y Online) program, a self-administered, online adaptation of Parenting Now's "Parenting: The First 3 Years" parenting group curriculum. In F3Y Online, parents receive a series of email messages that link to the program website with (a) video vignettes that model targeted skills for each topic, (b) interactive exercises, (c) printable worksheets and tip sheets, and (d) opportunities for discussion with other parents. In Phase I, we developed and evaluated the feasibility of the first segment of the F3Y Online series, "Incredible Infants Online." Positive results from Phase I paved the way for the Phase II refinement of the Infant program (Aim 1) and the development of the "Wonderful Ones Online" and "Terrific Twos Online" (Aim 2). For Aim 3, we experimentally evaluated the efficacy of the entire three-part series on parents' reports of parenting skills, self-efficacy, pleasure, stress, and social support; satisfaction with the programs; and, for the Ones and Twos programs, children's socio-emotional adjustment.

Evaluation and Results:

The First 3 Years Online program was experimentally evaluated in a randomized block 2x3x2 design. The between-subjects factors were (a) immediate treatment vs. waitlist control, (b) age group (mothers of 0 – 4 month-olds received Infants Online vs. mothers of 15–23 month-olds received Ones Online vs. mothers of 24–36 month-olds received Twos Online), and (c) minority status (non-Hispanic Caucasian vs. minority). Mothers meeting eligibility criteria (see below) were assessed at baseline (T1), stratified by age of child and race/ethnicity, and then randomized to immediate exposure to the intervention (TX) or to a Waitlist Control (WC) condition. All parents were then assessed again approximately 3 months later at immediate posttest (T2). After WC participants completed their T2 assessments, they received the age-appropriate F3Y Online program.

The study design initially called for a six-month follow-up (T3). Due to the time required to recruit our target sample size, however, it was not possible within the study timeframe to complete T3 for enough subjects to have adequate power to detect intervention effects. Thus, it was deemed prudent to forego T3 and instead focus on ensuring a full sample for T1 and T2, to allow maximum power to detect effects at T2.

Subjects

Subjects were 303 mothers of first-born children either 0-9 months old ($n=120$), 13-22 months old ($n=101$), and 23-32 months old ($n=82$). These age groupings correspond to the target ages of each of the F3Y Online program series.

Mothers were eligible to participate if they (a) were the biological or adoptive mother of a first-born child age 0 – 9 months or 13 – 32 months; (b) had custody of the child; (c) were at least 20 years old; (d) were English-proficient; (e) had high-speed access to the internet through a computer (any kind) or android mobile device (Apple mobile devices could not play the videos); (f) were willing to receive parenting information through their internet account; (g) checked email at least twice per week; and (h) in response to screening questions, reported that they were experiencing feelings of stress, uncertainty, lack of self-confidence in their parenting, and/or social isolation; and (i) not otherwise receiving parenting support services.

A total of 20% of the sample was of minority ethnicity or race. With regard to ethnicity, 93% of the sample was Non-Hispanic and 6% of the sample was Hispanic or Latino (1% unknown). With regard to race, 86% were White, 6% multi-racial, 5% Asian, 1% Black, and less than 1% each American Indian/Alaska Native, and Native Hawaiian/Pacific Islander, and 2% unknown. Most subjects (81%) were college graduates, and household income was generally in the middle- to upper-middle income range. Table 2 provides descriptives of baseline characteristics.

Table 2. Baseline Descriptives and Equivalence of Sample Characteristics and Outcome Measures at Baseline

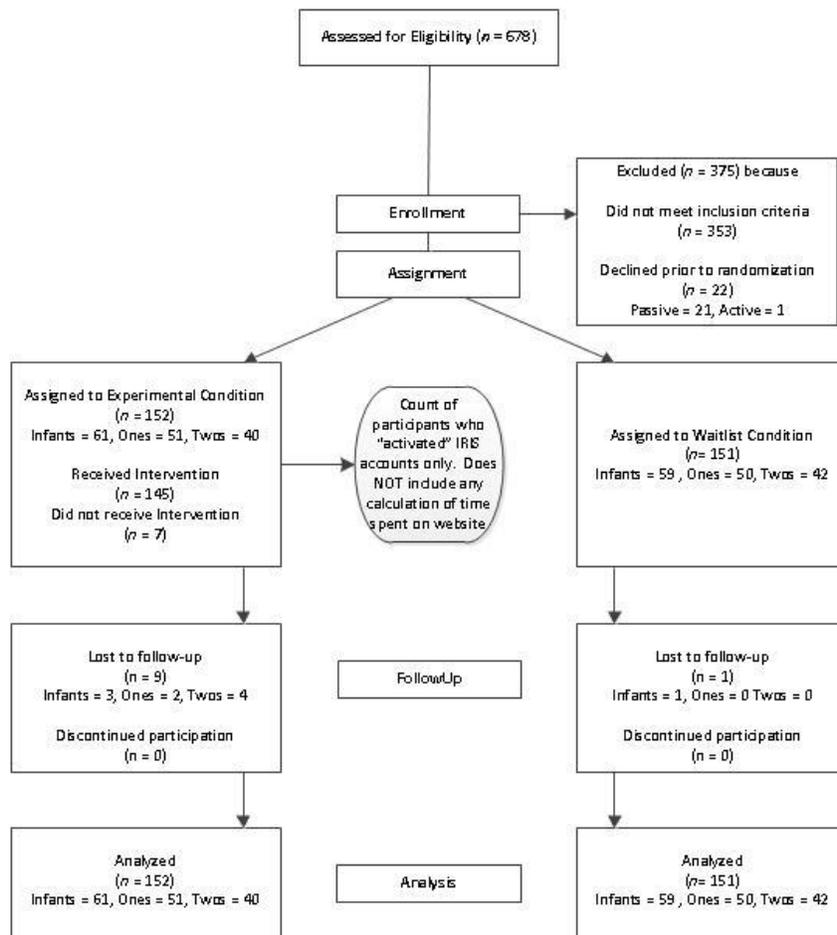
	WL	Int	Test Statistic	<i>p</i> -value
Child Characteristics				
Age, <i>n</i> (%)			0.09	.957
0 to 9 months	59 (39%)	61 (40%)		
13-22 months	50 (33%)	51 (34%)		
23-32 months	42 (28%)	40 (26%)		
Male, <i>n</i> (%)	82 (54%)	89 (59%)	0.56	.456
Minority, <i>n</i> (%)	37 (25%)	36 (24%)	0.03	.868
Mother Characteristics				
Age, <i>M</i> (<i>SD</i>)	31 (4.93)	31 (5.03)	1.85	.175
Minority, <i>n</i> (%)	28 (19%)	29 (19%)	0.03	.861
Household Characteristics				
Biological Father in home, <i>n</i> (%)	137 (91%)	138 (91%)	0.00	.985
Co-parenting, <i>n</i> (%)	136 (90%)	136 (90%)	0.03	.984
Working, <i>n</i> (%)	92 (61%)	94 (62%)	0.06	.811
Highest level of education, <i>n</i> (%)			3.15	.532
Grades 0-11	0 (0%)	1 (1%)		
High School or GED	5 (3%)	4 (3%)		
Some College	31 (21%)	22 (15%)		

College graduate	51 (34%)	58 (38%)		
Post College	64 (42%)	67 (44%)		
Household Income, n (%)			0.03	.998
Less than \$20,000	10 (7%)	10 (6%)		
\$20k-\$39,999k	26 (17%)	25 (16%)		
\$40-\$69,999	50 (33%)	51 (34%)		
\$70k or more	63 (42%)	64 (42%)		
Unspecified	2 (1%)	2 (1%)		
Outcome Measures at Baseline				
Child				
Child Displayed Total Problems	1.92 (0.30)	1.84 (0.29)	3.45	.065
Parent				
Pleasure in Parenting	19.34 (5.06)	18.33 (4.40)	3.40	.066
Perceived Stress Scale	27.81 (5.81)	27.28 (6.46)	0.58	.446
Parenting Skills, Infant	0.78 (0.09)	0.82 (0.08)	5.63	.019
Parenting Confidence, Infant	2.49 (0.61)	2.39 (0.63)	0.72	.398
Parenting Skills, Ones	0.84 (0.07)	0.84 (0.08)	0.01	.925
Parenting Confidence, Ones	2.30 (0.64)	2.39 (0.63)	3.35	.070
Parenting Skills, Twos	0.84 (0.08)	0.89 (0.06)	8.73	.004
Parenting Confidence, Twos	2.42 (0.52)	2.33 (0.68)	0.43	.513
Dysfunctional Parenting Practices	2.63 (0.61)	2.35 (0.59)	9.39	.003
Covariates				
Temperament, Difficultness	21.42 (5.80)	20.82 (6.21)	0.74	.390
Depressive Symptoms	6.07 (4.19)	5.49 (4.48)	1.32	.251
Social Support for Parenting, Satisfaction	5.54 (1.11)	5.67 (1.09)	1.14	.287
Social Support for Parenting, Unmet need	1.02 (1.09)	.84 (0.90)	2.27	.133

M: Mean, *SD*: Standard Deviation. Percentages can exceed 100% as a result of rounding to the nearest whole number.

Subject participation. Of the 303 mothers, all mothers completed pre-assessment questionnaires and were randomly allocated to a waitlist control (n=151) and intervention (n=152) condition. The post-intervention assessment was completed by 295 mothers (waitlist control n = 150, intervention n = 145), representing a very high retention rate of 97.4% (CONSORT, Figure 1).

ParentVES Consort Diagram



Measures and Assessment Procedures

Assessment procedures. All T1 and T2 assessments were conducted online, through the ORI Qualtrics tool. As indicated above, all participants who completed the online consent form were automatically routed to the T1 questionnaire within the Qualtrics system. Those who had not completed the T1 questionnaire within 4-5 days received prompts and reminders until the T1 assessment window closed after about 8 weeks. The invitation to the T2 post-intervention assessment was sent out via email to all study subjects 16 weeks after randomization; those who had not completed the T2 questionnaire within 4-5 days were prompted and reminded until the T2 assessment window closed after 8 weeks. The T2 questionnaire was identical to the T1 questionnaire in many respects, except that it included program satisfaction and usability questions for those in the immediate intervention condition.

Parent outcome measures. *Parenting skills* were measured with a Video Analog Test (VAT); parents viewed a series of 10 video vignettes depicting common parenting situations and were asked to identify the problem and the appropriate strategy for handling that situation, among multiple choices. Responses were scored on percent correct. Each age grouping (Infants, Ones, and Twos) had a different age-appropriate VAT measure.

For those mothers of children aged 13-32 months old at entry into the study, parenting skills were also measured with the Overreactivity and Laxness subscales of the Parenting Scale (PS; Arnold et al., 1993). The PS is a measure of dysfunctional discipline style, 20 items on which parents rate the nature of their discipline practices. These scales have shown good internal consistency, test-retest reliability, and discriminant and construct validity. The PS has been effectively used to identify at-risk parents, and scale scores are related to direct observations of dysfunctional discipline practices and child behavior problems (Arnold et al., 1993; Reitman et al., 2001).

Parenting self-efficacy was measured in two ways. First, after viewing each VAT vignette, parents indicated (on a scale of 1-5) the degree to which they felt confident that they could successfully handle a similar situation in their own lives. Parents of Ones and Twos also indicated their level of confidence (rating scale 1–10, from (1) Certain I can't do it to (10) Certain I can do it) in handling 20 negative child behaviors on the Child Adjustment and Parent Efficacy Scale (CAPES; Morawska & Sanders, 2010). Possible range for the Total CAPES efficacy score is 21–210. This scale has shown good internal consistency and construct validity.

The Social Support for Parenting Scale (adapted from Barrera et al., 1981 and Telleen, 1985) was used to measure *social support*. Mothers reported, on a 7-point scale, the degree to which they (a) had desired advice and information about raising a baby in the past month, (b) received helpful advice and information on raising a baby during that time, and (c) were satisfied with that advice and information. "Met need for advice" was measured by calculating the difference between advice desired and helpful advice received.

Parents' stress was measured with the Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983). The 10-item PSS assesses the degree to which situations in one's life are appraised as stressful and is rated 5-point frequency scale (never to very often) for the last month. The PSS has been shown to be a reliable and valid assessment of perceived stress.

Mothers' pleasure in parenting was measured with the Pleasure in Childrearing Schedule (Fagot, 1995). This 10-item questionnaire asks parents to rate, on a 5-point scale, the degree of pleasure they experience in a variety of routine childrearing tasks.

Child outcome measures. Mothers of Ones and Twos only reported on their *children's behaviors* with the Child Adjustment Scale of the CAPES (Morawska & Sanders, 2010). This scale samples children's emotional and behavioral problems (e.g. throwing tantrums, hurting other people, whining), with 30 items rated on a 4-point scale. The scale shows good internal consistency and construct validity.

Moderator measures. Child age, child temperament, baseline maternal depression, and various family demographics (mother working, income, and minority status) served as moderator variables. *Children's temperament* was measured at baseline for the different age groupings with three age-appropriate versions of the Infant Characteristics Questionnaire (ICQ; Bates, Freeland, & Lounsbury, 1979). The ICQ incorporates a 7-item response continuum to assess children's temperament on several different dimensions. Child "difficultness" is associated with concurrent and long-term problems in adjustment, so this factor was included as a moderator. Parents of all three age groups completed the 6-item child "difficultness" subscale.

In addition, *maternal depression* at baseline was measured with the *Patient Health Questionnaire* (PHQ-8; Kroenke et al., 2009). The PHQ-8 is a self-report version of the Primary Care Evaluation of Mental Disorders (PRIME-MD) to assess eight of the nine DSM-IV major depression symptoms (excluding the suicidal behavior symptom) and has been shown to be a reliable and valid brief depression assessment tool for the general population (Kroneke et al., 2009).

Process measures. Unobtrusive program use data from the IRISedOnline website tracked participant visits, duration, and completion of program components. A measure of *exposure to program components* was created by using the mean of the z-score transformations of website visit days (number), duration (minutes), and components completed (number).

Participant reports of program satisfaction were obtained at posttest. Participants were presented with a set of items to rate overall satisfaction, satisfaction of specific website features, and the extent to which the program met their needs, addressed important topics, and was interesting, entertaining, and useful. Lastly, participants reported if they would recommend the program and their willingness to pay \$20 to access the program.

Website usability was assessed with the System Usability Scale (SUS; Brooke, 1996). The SUS is an industry standard measure of website usability. The ten-item measure is rated on a 5-point scale with response options of *strongly disagree* to *strongly agree*.

Analytic Strategy

Baseline equivalence and predictors of attrition. To evaluate the effectiveness of randomization and identify potential confounds, ANOVA and Chi-Square analyses were used to evaluate baseline reports of sample characteristics, outcome measures, and predictors of outcome measures by condition (Immediate Treatment vs. Waitlist Control), using available data across age groupings. Participant attrition by condition was evaluated using Fisher's Exact Test.

Analysis of outcomes. A program-wide evaluation was conducted using the outcomes available across all age groups. An age-specific evaluation of the Incredible Infants Online program was also performed; this analysis was limited to the outcomes measured in this subsample. Program effects were evaluated using multiple regression models estimated using Mplus (Version 7.1), with parameters and standard errors calculated using full information maximum likelihood estimation, in an intent-to-treat model. To control for individual differences at baseline, the baseline value of the outcome measure was included as a covariate. The resulting models were outcome measures regressed on condition assignment (0 = WL, 1 = Int) and baseline values. Alpha was set to $p < .05$, two-tailed, for all tests. To complement the analysis, a Hedge's g was computed using the intervention effect divided by the posttest pooled standard deviation, as a measure of effect size.

Moderation. Moderation analyses were explored to identify differential intervention outcomes by sample characteristics at baseline. Moderation analyses were conducted with multiple regression models predicting outcomes with baseline values, condition, moderator, and the moderator-by-condition interaction term. Moderators of interest included participant characteristics such as child difficult temperament, mother depressive symptoms, child gender (0 = female, 1 = male), mother

minority status (0 = white Non-Hispanic, 1 = Non-white, Hispanic, or multiracial), income (0 = household income greater than or equal to \$30k, 1 = household income less than \$30k), mother work status (0 = does not work, 1 = works one or more hours per week), and mother perceived stress. In addition to the sample characteristics, the effects of the individual age-focused series were evaluated by comparing outcomes for the Terrific Twos Online versus the Incredible Infants Online and Wonderful Ones Online. A reduced set of outcomes were examined for this comparison based on item availability for each age grouping.

To reduce the likelihood of spurious results, domain-specific adjusted p-values were used. The three identified domains, the number of tests per domain and the adjusted p values are as follows: Child domain, one test, $p < .05$; Parent Adjustment domain, five tests, $p < .01$, Parenting domain, two tests, $p < .025$.

Participant engagement. Dose-response analyses were conducted to examine the degree to which treatment engagement predicted change in the outcome measures at the posttest. Partial correlation analyses were used to examine the association between the dosage measure and baseline-adjusted outcome measures. Multiple regression analysis was used to examine the predictors of program engagement, based on the baseline demographic characteristics, parent outcome measures, and child outcome measures.

Program usability and acceptability. Qualitative feedback from participating mothers and descriptive statistics for measures of usability and satisfaction are provided.

RESULTS

Attrition analysis

Differences in assessment completion rates by condition were analyzed but none were detected.

Baseline Equivalence

As seen in Table 2, participating mothers tended to be somewhat older ($M = 32$, $SD = 5.0$), Caucasian (90%), and co-parenting (90%). Sixty-one percent reported they were working, with a range of education levels and annual incomes. The children of participating mothers were predominately Caucasian (92%) and fairly balanced across genders: 56% were boys and 44% were girls. An analysis of baseline sample characteristics, outcomes, and predictors of outcomes by condition detected three significant differences by condition: Parenting Skills for mothers of infants $F(1,118) = 5.63$, $p < .05$, (Int: $M = .82$, $SD = .08$; WL $M = .78$, $SD = .09$) and two year olds $F(1,80) = 8.73$, $p < .01$, (Int: $M = .89$, $SD = .06$; WL $M = .84$, $SD = .08$) and dysfunctional parenting practices $F(1,181) = 9.39$, $p < .01$ (Int $M = 2.35$, $SD = 0.59$; WL $M = 2.63$, $SD = 0.61$). The three tests indicated that mothers in the intervention condition had better parenting skills and less dysfunctional parenting practices at baseline than the control condition.

Intervention Outcomes

For the evaluation of First Three Years Online program-wide effects, an analysis that evaluated measures across age groups, Table 3 presents the descriptive statistics of the outcome measures at the two assessment points by condition. The test statistics, p-value and effect sizes for the program effects are

also reported. As can be seen in Table 3, small-to-large effects were obtained for two outcomes. First, a significant effect for the First Three Years Online was obtained on the amount of unmet need for social support (Hedges' $g = .22$). Program participants reported a reduction in unmet need for social support, compared to the control group. In addition, a significant effect on parenting skills was obtained for the Incredible Infants Online series (Hedges' $g = .34$). No other significant effects were detected, either for the overall program or for any age-specific series.

A post-hoc analysis of paired t-tests performed on the intervention condition alone revealed a significant amount of change between baseline and post-test on most of the outcome measures for the intervention condition; the fact that most outcome measures saw pre-to-post change supported the evaluation of a dose-response analysis.

Moderator Analyses

An exploratory look at differential intervention effects revealed six significant effects out of 97 tests conducted. The six effects detected were for some outcomes with some moderators, with no consistent pattern to make large claims. Specifically, mothers with fewer depressive symptoms obtained greater reductions in child problem behaviors, $\beta = .022$, $p < .01$, 95% CI [.011, -.034]. Non-minority mothers obtained greater reductions in perceived stress, $\beta = 4.96$, $p < .01$, 95% CI [2.58, 7.27]. Those with household incomes of \$30,000 or more obtained greater reductions in dysfunctional parenting practices, $\beta = 0.46$, $p < .01$, 95% CI [0.19, 0.74]. For those in the Terrific Twos Online, greater increases in parenting skills were obtained for parents of females, $\beta = 0.06$, $p < .05$, 95% CI [0.01, 0.09], and those with household incomes greater than or equal to \$30,000, $\beta = -0.07$, $p < .01$, 95% CI [-0.12, -0.04]. For those in the Incredible Ones Online, greater increases in parenting skills were obtained for mothers who did not work, $\beta = -0.06$, $p < .05$, 95% CI [-0.11, -0.02]. It should be noted that caution should be paid to the investigation of dichotomous moderators and the age-specific outcomes of parenting skills, as inferences were made using reduced sample sizes.

Engagement and Dose-Response Analyses

Unobtrusive program use data indicated that mothers were engaged with the program. The average mother visited the site on four separate days ($SD = 3.4$), on which they spent an average of 86.4 minutes on the website ($SD = 77.3$), completing 26.9 components ($SD = 16.4$). Included in these descriptive statistics are seven mothers (infants = 4, ones = 1 and twos = 2) who did not access the site at all and were assigned values of zero for visits, duration and components completed. Program engagement was greater in households where a biological father was present $F(1,150) = 8.54$, $p < .01$, where co-parenting occurred $F(1,150) = 6.60$, $p < .05$, and where household income was equal to or greater than \$30k $F(1,149) = 4.79$, $p < .05$.

Exposure to the First 3 Years Online program was positively related to increased pleasure in parenting, partial $r(140) = .17$, $p < .05$, to improved parenting skills for the Incredible Infants Online, partial $r(56) = .39$, $p < .01$, and to increased parental confidence for the Wonderful Ones Online, partial $r(46) = .41$, $p < .001$.

Program Satisfaction and Usability

Participant satisfaction items collected at post-test indicated that 59% of the mothers reported being “quite” or “very satisfied” with the First 3 Years Online series. Participants reported that the program met their needs (“somewhat,” “quite,” or “very”: 81%), addressed important topics (83%), was interesting and entertaining (82%), and was found useful (80%). Similar percentages were achieved for satisfaction with specific program features, such as the videos (“somewhat,” “quite,” or “very”: 75%) and worksheets (66%). Seventy-four percent of mothers said that they would recommend the program to other parents and 22% would pay \$20 to access the program.

Results on the System Usability Scale administered at post-test provided a quantitative measure of website usability. The average System Usability Scale score was 73.3 ($SD = 16.6$), which translates to an “above average” usability grade of “B-” for the First 3 Years Online series (Sauro, 2011; Sauro & Lewis, 2012).

Parent comments reflected their general satisfaction with the program. One parent remarked that that the program “built confidence, [I] felt less alone.” Parents also appreciated the program format, “It is well organized, easy to use and helpful!” and the relevance of the content to their own lives, “It couldn't have come at a better time in my life. It was so informational and gave me some much encouragement that I was doing a good job...” “It really helped me realize I actually had good parenting instincts. It gave me more confidence & was very educational. I also felt less alone.”

Although parents were generally enthusiastic, some parents expressed challenges in finding the time to fit the program into their busy lives “[The program] really covered a lot of the issues we are dealing with in our house. The problem was finding the time to complete it with a TWO who doesn't nap!” and “...would be better if it was shorter. Working parents don't have time to do all this stuff.” Parents suggested that having a flexible format that worked for portable devices such as tablets or phones would have afforded more access. One parent stated, “I would love if I could watch the videos on my phone, I don't often have time to get on the internet at home” and “I wish the videos worked on my iPad. I don't have access to a real computer often so I wasn't able to easily access the program.”

Among parents who expressed dissatisfaction with the program content, some parents commented that the program focus was too broad for them. “The information was a bit too general for me. I would like to see more information on sleep habits and training and more hands on tips and information” and other parents felt that the program content was geared toward parents with less experience, “It's a good program for the first time parent” and “very helpful, especially to first time parents with limited baby experience.” Other parents said that they would prefer a group format, “This kind of program would have a huge benefit if it was done in a classroom form, but my feeling is that it has only moderate benefit being done online only” and “...Personally, I'd prefer to be in a class setting, interacting with a facilitator and other participants.”

With regard to paying for the program, one parent commented that “Having gone through Wonderful ONEs, I would pay \$20 to continue and participate in Twos and Threes,” while another parent commented, “Before I participated and knew how helpful the program was, I would probably not have chosen to pay for it.”

Discussion

In summary, modest effects were found for the First Three Years Online program. Significant effects were found on reductions in unmet need for social support across all programs, and on improved parenting skills for the Incredible Infants Online program. No other main effects were found. Some scattered moderation effects were found, but there was no consistent pattern of interactions. In the dosage-response analyses, those who got greater exposure to the program reported greater improvements in parenting pleasure overall, greater improvements in parenting skills for the Incredible Infants Online program, and greater improvements in parental confidence for the Wonderful Ones Online program.

In general, parents reported high levels of satisfaction with the program, the information it offered, and the online format in which it was presented. Some parents gave specific points of constructive feedback. The lack of stronger program effects might be attributable to several factors. First, the sample was fairly well educated and middle- to high-middle income. Their need for the kind of information offered in the program may not have been particularly high, and a higher-risk sample may have benefitted more from the program. Indeed, ceiling effects were observed in some measures across both conditions. In addition, pre-to-post changes were observed in both conditions, suggesting that time and learning experience between the baseline and post-assessment, rather than the program itself, may have accounted for at least some of the changes seen in parents' and children's outcomes. These findings raise questions about the appropriate target audience for the program. The program is designed as a universal intervention, appropriate for those experiencing mild to moderate levels of parenting stress, lack of parenting confidence, and social isolation. Although the sample was screened for these characteristics and they appeared to enjoy the program, it is possible that their level of need was not great enough to show significant program effects. Evaluating the program on a more at-risk sample would provide a test of the question of whether those experiencing greater need might benefit from the program more.