Comparison of a Pretest Versus a Retrospective Pretest with Parents and Children Involved in the Strengthening Families Program

Ray M. Merrill PhD, MPH*
Dillon Simmons EdD
Jaynie L. Brown BA

Abstract

A retrospective pretest (a survey technique in which participants are asked to indicate their level of understanding, attitudes, and skills prior to the intervention, even though they are asked after the intervention occurred) for identifying a self-reported program's effect may be more accurate than a pretest due to bias toward more favorable behavior reporting at baseline. This study compares a pretest (or preintervention) and a retrospective pretest for evaluating the effect of the Strengthening Families Program (SFP) for parents and youth. Analyses are based on 135 parents and 137 youth who received the SFP intervention and completed a pretest and retrospective pretest. Data was collected in 2023. Overestimation bias in the pretest was identified for several items and corresponding constructs related to parenting and family functioning (bonding, boundaries, monitoring, and child good behavior). This overestimation may be due to lack of trust in the service delivery, lower mindfulness about adequate parenting and family functioning, and a desire to be more socially appropriate. The level of overestimation bias was not dependent on who the child lives with (mother, father, both, neither), sex, age, race, or ethnicity. A retrospective pretest can help lower overestimation bias and provide a more accurate indication of program efficacy based on a within group design.

Keywords: Bias, pretest, retrospective pretest; Strengthening Families Program

***Corresponding author can be reached at Ray_Merrill@byu.edu

Introduction

Establishing whether a public health program is satisfying its intended purpose often involves within and between group designs to determine its efficacy. However, the common use of a pretest may lack criterion validity. Inaccurate responses may be due to insufficient knowledge, attitudes, and ability to accurately perceive things before participation in the intervention (Brook et al., 2016; Piwowar & Theil, 2014, English & Horowitz, 2002). In response, a retrospective pretest has been considered for acquiring baseline knowledge, attitudes, and perceptions (Piwowar & Thiel, 2014, English & Horowitz, 202, Pratt et al., 2000, Walk 1956; Deutsch & Collins, 1951). A more recent study compared the use of a pretest versus retrospective pretest with child-welfare involved families (Brook et al., 2016). The study showed that the Strengthening Families Program (SFP), which separates the parents from youth for skills training and then brings them together for later behavior skills practice (Kumpfer & Hansen, 2014; LoBraico et al., 2019), positively impacted family, child, and parent functioning. Consideration of a retrospective pretest had some impact on effect size computations.

Prior to building a trusting relationship within the service delivery process, the authors of one study noted that parents may overestimate their caregiving abilities and practices (Kumpfer et al., 2010). The authors also found that social desirability bias was greater in child welfare families who were motivated to higher function than other groups. Further, prior to the intervention the authors believed that parents would be less mindful of their parent and family functioning skills. That is, parents would rate their parenting and family functioning more favorably in a pretest, but after gaining better skills, knowledge, and more trust in the service delivery process, they would be more willing to reveal areas of challenge in their family. A response shift may also be because the intervention, intended to help them better understand desirable parenting behavior, changed their awareness or perceptions of their own parenting, causing them to now report their previous parenting skills less favorably (Brook et al., 2016). A similar response shift may appear in children.

To provide additional theoretical grounding, this study utilizes the Family Resilience Framework (Walsh, 2003) alongside Kumpfer's foundational insights. The Family Resilience Framework emphasizes adaptive family processes and resilience-building skills developed during interventions. Walsh suggests that properly designed interventions focused on recovery and growth help families reshape their perceptions of their abilities by promoting reflection and skill competency (2003). This research utilizes Walsh's (2003) three key processes (Family Belief Systems, Organizational Patterns, and Communication / Problem-Solving) alongside Kumpfer's observations about response shifts due to increased self-awareness and trust to provide a stronger lens for interpreting differences between pretest and retrospective pretests.

The purpose of this study was to compare whether differences occurred in a pretest versus a retrospective pretest across 32 items and seven constructs used to measure parenting skills, family cohesion, organization, communication, and social skills. The study was conducted among nationally representative parent-and-youth dyads participating in different learning modalities in the SFP. While this study primarily focused on response shifts as a marker of intervention impact, it is important to acknowledge that the pretest also provided valuable descriptive information about participants' initial self-perceptions. Understanding the full range of these baseline responses can inform future tailoring of interventions and outreach strategies. The efficacy of the intervention, determined in relation to a comparison group, has been established.

Methods

Population

Family participants were obtained through 24 SFP recruiters who were racially and demographically diverse across the United States. Each recruiter had a benchmark of racially/ethnically diverse families to recruit. Each recruiter received $250 when they identified the families and $250 when the families completed the study. Each family received $450 for their participation, divided between parents and youth. Recruitment and participation in the study occurred in 2023.

Sample

A sample of 245 parents and 244 youths were recruited for the study. Each of these individuals completed the pretest prior to randomization to the intervention arms of the study. The pretest included personal and demographic information and several

questions about family behaviors and dynamics. Participants were then randomly assigned one of three arms of the study (Game, Video, or Control). Of the initial participants, 217 (88.6%) parents and 218 (89.3%) youth completed the study. 217 parents were randomly assigned to the SFP Online Game (n = 79), SFP Video (n = 56), and control group (n = 82). 218 youth were assigned to either the SFP Online Game (n = 81), the SFP Video (n = 56), or the control group (n = 81). Continued participation was not statistically significantly associated with parent and youth age, sex, race, and ethnicity, or parent education or income (Chi-square $p > 0.05$). These individuals did not significantly differ from those in the Control group (82 for parents and 81 for youth) with respect to parent and youth age, sex, race, or ethnicity, or parent education or income (Chi-square $p > 0.05$). Hence, randomization produced groups with similar distributions of these demographic variables.

Another study found that both the Game and Video groups statistically significantly showed more favorably results than the control group, but there was no significant difference between the Game and Video groups. Because only the Game and Video groups completed the retrospective pretest and the efficacy of the intervention was similar between the Game and Video groups, the control group was not included in the current study and the Game and Video groups were combined in the current analysis.

Interventions

The Strengthening Families Program was developed in the early 1980s for the children of parents with substance use issues under a National Institute of Drug Abuse grant (Kumpfer et al., 2010). The assignment in the grant was to discover which skills parents needed to keep their kids from using alcohol and drugs. Essential skills were identified and put in lesson format for classroom instruction for both parents and youth, who then practiced the skills together during the second hour of class. As it turned out, SFP skills training also reduced youth depression and delinquency.

In this study, the intervention groups received updated SFP lesson material (SFP 7-17) via an SFP 11-lesson video series, or a new 10-session online gamified version of SFP. The gamified intervention was delivered online remotely with the app maintained on a server at Imagen Multimedia and the evaluation component was conducted by Qualtrics, an online survey methodology company. An introductory email sent to parents provided a hyperlink where parents consented to participate and gave permission for their child to participate and provided their child's email address to also assent and participate. The hyperlink provided an electronic "registration packet" that explained the study.

Participants in the SFP Online Game group received a 10-lesson online family-based intervention with separate, but intersecting, tracks for parents and youth. Each lesson contained three 10-minute mini-lessons per week and includes behavioral skills training and interactive multimedia lessons with gamified learning interactions that target parenting skills, family cohesion, organization, communication, social skills, and drug prevention for youth. In between these mini-lessons, users completed family activities made up of reflective form-based questions that were forwarded to a family leader to help foster meaningful family discussions over key skills. Self-correcting quizzes were used to assess learning and engagement. Each lesson concluded with game rewards and activities that family members were encouraged to use for further discussion and skill practice in their weekly SFP family meetings.

Participants in the Video group received an 11-session home-use DVD or coupon code for viewing the SFP videos online at home. These contain the same SFP skills and lesson content as the online Game. Participants used the Internet or a DVD player to view lessons together as a family in a self-paced format. The video lesson material is more static without the interaction required in the gamified version of SFP Online.

At the end of the intervention, participants in the Game and Video arms of the study received a hyperlink URL to complete both a retrospective pretest and a posttest. In the retrospective pretest they were asked to respond according to how they would prior to completing the intervention. Their views measured in the surveys about family cohesion, organization, communication, and social skills likely changed because of the new skills they learned and practiced, and the acquisition of possibly more accurate perspectives on parenting and instructions on how to build loving relationships they gained in the intervention. In the pretest, participants may also have felt less trust in the system prior to participating in the intervention and thus overstated positive skills and behaviors (Kumpfer et al., 2010). These intervention activities align with the core concepts of the Family Resilience Framework, which focus on fostering reflection, family structure, and collaborative problem-solving skills (Walsh, 2003).

Constructs

For biological parents, caregivers, or legal guardians, each lesson entails learning nurturing parenting skills that strengthen family bonds, setting clear boundaries with positive discipline, and monitoring youth's social activities and emotional well-being. The youth lessons also train youth in skills to nurture family bonds, accept boundaries, stress the importance of parental monitoring, and promote mental health. Three other constructs were also considered: good behavior, bad behavior, and family dysfunction.

Seven constructs are identified, made up of 32 items. These items were measured through survey responses at the pretest and retrospective pretest. The bonding construct involves items measuring care, love, respect, and quality time. The boundaries construct involves items measuring clear firm rules against antisocial behavior and substance use. The monitoring construct involves items measuring their children's emotional well-being and their activities, to see that they stay in an alcohol-free social environment; and the mental health construct involves items measuring whether the youth is sad, depressed, unhappy, or irritable. Good behavior includes those actions that increase the likelihood of success in academic, social, and family settings. Bad behavior includes actions that are inappropriate, disruptive, or harmful in some way. Family dysfunction is lack of structure, misbehavior, or abuse between family members.

The constructs measured in our surveys align conceptually with the Family Resilience Framework, specifically targeting the framework's three key processes outlined by Walsh (2003). The bonding and mental health constructs reflect aspects of family belief systems (attitudes toward nurturing and emotional care). Boundaries and monitoring represent organizational patterns (family rules, structure, and monitoring practices). Lastly, constructs measuring communication and collaborative skills closely parallel Walsh's domain of communication and problem-solving.

Hypotheses

We hypothesize that the pretest will be inflated for items related to bonding,

boundaries, monitoring, good behavior, bad behavior, and family dysfunction. The mental health construct is more factual and unlikely to change from pretest to retrospective pretest. We further hypothesize that items involving socially acceptable personal behaviors are more likely to be inflated in the pretest.

Statistical Techniques

Demographic variables were classified according to categories, the levels of which were described using numbers and percentages. The distributions of these variables were compared between the groups.

Item responses range from "never" to "almost always" or "very wrong" to "not at all wrong." These variables were re-coded using a scale from 0 to 4, with mean scores, along with standard deviations reported at the pretest and (10-weeks later) retrospective pretest. Differences in scores were evaluated using the paired t-test.

The degree that specific items contributed to underlying constructs of each of these types of measures was evaluated using factor analysis and Cronbach alpha. For the items providing an underlying construct for each of the measures, new summary variables (constructs) were derived.

Regression analysis evaluated the association between the items and construct scores with pretest and retrospective pretest and the demographic variables. Interaction terms were evaluated in the models involving the pretest versus retrospective pretest variable and each of the demographic variables. Two-sided tests of statistical significance were based on the 0.05 level against a null hypothesis of no association. Data were evaluated using the Statistical Analysis System (SAS) software package (Base SAS 9.4; SAS Institute, Inc., 2016). Findings were then analyzed with consideration for how changes in perceptions can influence response shifts in line with Walsh's (2003) Family Resilience Framework.

Results

Self-reported demographic information of parents and youths appear in Table 1. Most parents who completed the survey were ages 41-50 years, female, White, and non-Hispanic, had a college degree, and had an annual income of ≥$100,000. Most youth were ages 13-14 years, male, White, non-Hispanic, and lived with both parents.

Parent – Pretest versus Retrospective Pretest

Mean pretest and retrospective pretest scores among parents completing the SFP Online Game or the SFP Video are shown for 32 items and seven corresponding constructs in Table 2. A single construct was identified for each of the groups of items at pretest and retrospective pretest. Cronbach's alpha scores are generally good, although the scores for boundaries show lower reliability. Statistically significantly higher pretest scores compared with retrospective pretest scores appear for the bonding, boundaries, and monitoring constructs. No statistically significant differences resulted in mental health, good behavior, bad behavior, or family dysfunction constructs.

Table 1. Demographic Characteristics for Parents and Youths

Parent			Youth		
	No.	%		No.	%
Age			Age		
18-30	2	1.5	9-10	3	2.2
31-40	38	28.1	11-12	28	20.4
41-50	75	55.6	13-14	53	38.7
≥ 51	20	14.8	15-16	36	26.3
			17-18	17	12.4
Sex					
Male	11	8.1	Male	72	52.6
Female	124	91.9	Female	65	47.5
Race			Race		
White	115	85.2	White	113	82.5
Black	7	5.2	Black	8	5.8
Other	13	9.6	Other	16	11.7
Hispanic			Hispanic		
No	125	92.6	No	120	87.6
Yes	10	7.4	Yes	17	12.4
Education			Child lives with …		
HS or GED	6	4.4	Both parents	107	78.1
Post HS	25	18.5	Single parent	26	19.0
College Grad	104	77.0	Other†	4	2.9
Annual income					
< 40K	17	12.6			
40K - < 80K	29	21.5			
80K - < 100K	15	11.1			
≥ 100K	74	54.8			

†Lives with relatives, joint or shared custody, or other.

For parents, mean scores for the different constructs measured at pretest and retrospective pretest were adjusted for the demographic variables appear in Table 3. A summary of the results is as follows:

- For bonding, there was a statistically significantly higher adjusted mean score in the pretest compared with the retrospective pretest. None of the demographic variables were statistically significantly associated with bonding.
- For boundaries, in addition to there being a statistically significantly higher adjusted mean score in the pretest (vs. the retrospective pretest) group, the adjusted mean score was significantly higher for Hispanics (vs. non-Hispanics) (mean 0.34, $p = 0.0189$) and significantly lower for those living with one parent (vs. both) (mean -0.30, $p = 0.0044$).
- For monitoring, in addition to there being a statistically significantly higher adjusted mean score in the pretest (vs. the

retrospective pretest) group, the adjusted mean score was significantly higher for females (vs. males) (mean 0.34, p = 0.0126).

- For mental health, there was not a statistically significant difference in adjusted mean scores between the pretest and the retrospective pretest. The adjusted mean mental health score was significantly lower for ages 31-40 (vs. 18-30) (mean -0.79, p = 0.0374), Blacks (vs. Whites) (mean -0.44, p = 0.0574), and higher for those living with one parent (vs both) 1.25, p < .0001).
- For good behavior, in addition to there being a statistically significantly higher adjusted mean score in the pretest (vs. the retrospective pretest) group, the adjusted mean score was significantly lower for those with a HS/GED degree (vs. College Graduate) (mean -0.63, p = 0.0007) and for those living with one parent (vs. both) (mean -0.54, p < .0001) or other (vs. both) (-0.52, p = 0.0010).
- For bad behavior, there was not a statistically significant difference in adjusted mean scores between the pretest compared with the retrospective pretest. The adjusted mean bad behavior score was significantly lower for ages 31-40 (vs. 18-30) (mean -0.76, p 0.0032), 41-50 (vs. 18-30) (mean -0.69, p 0.0064), and ≥ 51 (vs. 18-30) (-0.78, p = 0.0030), and higher for those living with one parent (vs. both) (mean 0.22, p = 0.0079) and other (vs. both) (mean 0.34, p = 0.0053).
- For family dysfunction, there was not a statistically significant difference in the adjusted mean scores between the pretest and the retrospective pretest. The adjusted mean family dysfunction score was significantly lower for ages 31-40 (vs. 18-30) (mean -0.91, p 0.0102), 41-50 (vs. 18-30) (mean -0.86, p 0.0140), and ≥ 51 (vs. 18-30) (-1.16, p = 0.0015).

While some of the demographic variables were associated with the constructs, as described above in the adjusted models, interaction terms involving the pretest versus the retrospective pretest and the demographic variables were not statistically significant. In other words, none of the demographic variables modified the associations between the constructs and the pretest versus the retrospective pretest variable.

Youth – Pretest versus Retrospective Pretest

The average pretest and retrospective pretest scores among youth completing the SFP Online Game or the SFP Video are shown for 32 measures and seven corresponding constructs in Table 4. As found in the parent data, a single construct was identified for the selected item responses at pretest and retrospective pretest. Cronbach's alpha scores are good for bonding, mental health, good behavior, and family dysfunction, and acceptable for boundaries, monitoring, and bad behavior. Statistically significantly higher pretest scores compared with retrospective pretest scores appear for the bonding, boundaries, monitoring constructs, and good behavior. No statistically significant differences result in mental health, bad behavior, or family dysfunction constructs.

Table 2. Parent Pretest versus Retrospective Pretest Scores

	Pretest			Retrospective Pretest			
Variable	**Factor Pattern**	**Mean**	**Std Dev**	**Factor Pattern**	**Mean**	**Std Dev**	**Paired t p value**
Bonding							
I praise my youth when he/she behaves well.	0.73	2.84	0.81	0.72	2.31	0.73	<.001
I let my youth know I really care about him or her.	0.79	3.16	0.79	0.80	2.85	0.90	0.004
I am loving and affectionate with my youth.	0.77	2.96	0.88	0.77	2.85	0.90	0.340
I listen respectfully to my youth's point of view.	0.61	2.81	0.71	0.61	2.50	0.79	0.001
I spend quality time (one-on-one play time) with my youth.	0.70	2.19	0.83	0.70	1.93	0.87	0.013
I talk to my youth about his/her feelings.	0.69	2.51	0.79	0.73	2.25	0.85	0.010
Bonding Construct		2.74	0.59		2.45	0.57	<.001
Cronbach Alpha		0.81			0.82		
Eigenvalue	3.09			3.15			
Boundaries							
I follow through with reasonable consequences when rules are broken.	0.52	2.64	0.93	0.56	2.42	0.88	0.046
Our family has clear, firm rules about no youth alcohol and drug use.	0.74	3.80	0.60	0.72	3.53	0.96	0.005
I talk to my youth about how to avoid using alcohol, tobacco, and drugs.	0.78	2.98	1.04	0.76	2.66	1.19	0.020
My youth knows how to say "No" to stay out of trouble.	0.65	3.06	0.82	0.64	2.99	0.90	0.479
Boundaries Construct		3.12	0.56		2.90	0.68	0.004
Cronbach Alpha		0.60			0.58		
Eigenvalue	1.85			1.81			
Monitoring							
I talk to my youth about his or her plans for the next day or week.	0.77	2.82	0.90	0.77	2.52	0.90	0.006
I talk to my youth about his or her friends.	0.80	2.79	0.85	0.79	2.55	0.85	0.019
I check daily to see if my youth completes his/her homework.	0.70	2.48	1.21	0.70	2.39	1.18	0.508
I know where my youth is and who he/she is with, and what they are doing.	0.46	3.70	0.52	0.49	3.59	0.65	0.123
Monitoring Construct		2.95	0.61		2.76	0.61	0.011
Cronbach Alpha		0.63			0.63		
Eigenvalue	1.93			1.94			

Mental Health							
My youth looks sad or down.	0.88	1.44	0.74	0.88	1.44	0.81	0.937
My youth appears depressed.	0.89	1.01	0.81	0.89	1.01	0.93	0.999
My youth is irritable.	0.73	1.64	0.80	0.73	1.73	0.89	0.429
Mental Health Construct		1.37	0.66		1.40	0.74	0.718
Cronbach Alpha		0.78			0.78		
Eigenvalue	2.11			2.11			
Good Behavior							
My youth helps with chores, errands, and other work.	0.51	2.81	1.00	0.48	2.52	0.94	0.015
My youth gets good grades (A's and B's)	0.77	3.56	0.81	0.78	3.53	0.84	0.711
My youth pays attention.	0.85	2.95	0.88	0.86	2.82	0.86	0.235
My youth stays on task until completed.	0.81	2.60	0.88	0.81	2.49	0.92	0.313
Good Behavior Construct		2.98	0.66		2.84	0.66	0.079
Cronbach Alpha		0.71			0.71		
Eigenvalue	2.22			2.24			
Bad Behavior							
My youth yells at others.	0.74	1.20	0.83	0.76	1.22	0.94	0.837
My youth gets in fights.	0.65	0.45	0.71	0.66	0.46	0.76	0.934
My youth loses his or her temper.	0.78	1.45	0.84	0.78	1.39	0.92	0.581
My youth breaks rules.	0.73	1.17	0.75	0.71	1.17	0.85	1.000
My youth tells lies.	0.59	1.22	0.91	0.57	1.15	0.97	0.518
My youth skips school.	0.36	0.14	0.43	0.31	0.14	0.46	1.000
Bad Behavior Construct		0.94	0.48		0.92	0.54	0.781
Cronbach Alpha		0.73			0.72		
Eigenvalue	2.58			2.56			
Family Dysfunction							
I yell or shout when my youth misbehaves.	0.67	1.63	0.99	0.63	1.59	1.11	0.774
People in my family insult or yell at each other.	0.85	1.47	0.84	0.84	1.54	0.97	0.545
People in my family have serious arguments.	0.88	1.21	0.75	0.86	1.24	0.92	0.718
We fight a lot in our family.	0.85	1.33	0.80	0.85	1.41	0.89	0.390
Family Dysfunction Construct		1.42	0.63		1.52	0.76	0.268
Cronbach Alpha		0.83			0.81		

Eigenvalue	2.66			2.57			

Scale for each item: 0 = Never to 4 = Almost Always. Factor analysis showed one construct for each of the reported measures. Factor analysis, using the method of principle components, showed one construct for each of the corresponding measures, based on eigenvalues greater than 1.

Table 3. Multiple Regression Evaluation Comparing Parent Pretest versus Retrospective Pretest Scores for Each Construct

	Pretest	Retrospective Pretest	
Variable	**Adjusted Mean**	**Adjusted Mean**	**t p value**
Bonding Construct	2.66	2.34	<.001
Boundaries	3.10	2.87	0.003
Monitoring	2.62	2.43	0.010
Mental Health	1.60	1.63	0.710
Good Behavior	2.47	2.33	0.055
Bad Behavior	1.24	1.23	0.774
Family Dysfunction	1.50	1.60	0.255

Note. Means were adjusted for age, sex, ethnicity, parent, and education. Scale for each item: 0 = Never to 4 = Almost Always.

Table 4. Youth Pretest versus Retrospective Pretest Scores

	Pretest			Retrospective Pretest			
Variable	**Factor Pattern**	**Mean**	**Std Dev**	**Factor Pattern**	**Mean**	**Std Dev**	**Paired t p value**
Bonding							
My parent/caregiver praises me when I behave well.	0.68	2.88	0.85	0.75	2.47	0.91	0.000
My parent/caregiver lets me know they really care about me.	0.83	3.37	0.77	0.74	3.04	0.82	0.001
My parent/caregiver is loving and affectionate with me.	0.81	3.43	0.81	0.81	3.10	0.92	0.002
I listen respectfully to my parent's/caregiver's point of view.	0.53	3.01	0.86	0.65	2.72	0.94	0.006
My parent/caregiver spends one-on-one time playing with me.	0.65	2.17	0.98	0.70	1.91	1.05	0.038
My parent/caregiver talks to me about my feelings.	0.74	2.60	1.05	0.67	2.31	1.09	0.025
I enjoy spending time with my parent/caregiver.	0.75	3.32	0.84	0.66	3.06	1.00	0.019
Bonding Construct		2.97	0.62		2.66	0.66	<.001
Cronbach Alpha		0.83			0.83		
Eigenvalue	3.61			3.56			
Boundaries							
My parent/caregiver gives me fair, reasonable consequences when I don't follow family rules.	0.62	3.14	0.87	0.62	2.80	1.02	0.003
Our family has clear, firm rules about no youth alcohol and drug use.	0.62	3.91	0.37	0.70	3.74	0.69	0.013
My parent/caregiver talks to me about how to avoid using alcohol, tobacco, and drugs.	0.59	3.27	0.98	0.73	3.00	1.15	0.038
I know how to say "No" to stay out of trouble.	0.71	3.43	0.75	0.67	3.18	0.90	0.014
Boundaries Construct		3.44	0.47		3.18	0.65	0.000
Cronbach Alpha		0.48			0.60		
Eigenvalue	1.62			1.87			
Monitoring							
My parent/caregiver talks to me about my plans for the next day or week.	0.73	2.81	1.03	0.75	2.40	1.15	0.002
My parent/caregiver talks to me about my friends.	0.78	2.61	1.06	0.73	2.32	0.98	0.019
My parent/caregiver checks daily to see if I complete my homework.	0.63	2.80	1.20	0.64	2.41	1.30	0.011

When I am not at home, my parent or caregiver knows where I am, who I am with, and what I am doing.	0.20	3.72	0.53	0.53	3.42	0.85	0.001
Monitoring Construct		2.98	0.62		2.64	0.73	<.001
Cronbach Alpha		0.48			0.58		
Eigenvalue	1.58			1.78			
Mental Health							
I feel sad or down.	0.89	1.67	1.02	0.89	1.77	1.01	0.439
I feel depressed.	0.88	1.04	1.04	0.90	1.09	1.12	0.696
I am happy most of the time.	0.84	1.01	0.97	0.84	1.09	0.92	0.444
Mental Health Construct		1.24	0.88		1.32	0.90	0.470
Cronbach Alpha		0.84			0.85		
Eigenvalue	2.28			2.31			
Good Behavior							
I help with chores, errands, and other work.	0.58	3.24	0.83	0.49	2.92	0.95	0.003
I can control my anger.	0.60	2.54	1.06	0.59	2.34	1.17	0.145
I follow rules.	0.76	3.27	0.83	0.74	3.11	0.93	0.132
I pay attention in school.	0.82	3.45	0.73	0.82	3.15	0.90	0.002
I get good grades (A's & B's)	0.78	3.60	0.84	0.77	3.44	0.85	0.115
I stay on a task until it is completed.	0.72	2.80	0.88	0.76	2.40	1.03	0.001
Good Behavior Construct		3.15	0.58		2.89	0.64	0.001
Cronbach Alpha		0.79			0.78		
Eigenvalue	3.07			2.98			
Bad Behavior							
I yell at others when I get angry.	0.84	1.86	1.08	0.83	1.68	1.06	0.159
I get in fights.	0.70	0.95	0.97	0.74	0.89	1.04	0.632
I tell lies.	0.66	1.30	0.92	0.66	1.46	1.03	0.174
I skip school.	0.31	0.18	0.53	0.21	0.13	0.40	0.330
Bad Behavior Construct		1.07	0.58		1.05	0.60	0.721
Cronbach Alpha		0.56			0.54		
Eigenvalue	1.73			1.71			
Family Dysfunction							
My parent/caregiver yells or shouts at me when I misbehave.	0.58	1.63	0.99	0.61	1.59	1.11	0.774

People in my family insult or yell at each other.	0.86	1.69	1.05	0.86	1.81	1.10	0.369
People in my family have serious arguments.	0.81	1.32	0.93	0.85	1.42	0.99	0.502
We fight a lot in our family.	0.86	1.50	1.03	0.87	1.58	1.12	0.590
Family Dysfunction Construct		1.53	0.78		1.60	0.87	0.499
Cronbach Alpha		0.78			0.81		
Eigenvalue	2.47			2.59			

Scale for each item: 0 = Never to 4 = Almost Always. “I am happy most of the time” was reverse scored. Factor analysis, using the method of principle components, showed one construct for each of the corresponding measures, based on eigenvalues greater than 1.

For youth, mean scores for the different constructs measured at the pretest and retrospective pretest were adjusted for the demographic variables (Table 5). A summary of the results is as follows:

- For bonding, in addition to there being a statistically significantly higher adjusted mean score in the pretest (vs. the retrospective pretest) group, the adjusted mean bonding score was significantly lower for those living with one parent (vs. both) (mean -0.27, $p = 0.0053$) or other (vs. both) (-0.74, $p = 0.0011$).
- For boundaries, in addition to there being a statistically significantly higher adjusted mean score in the pretest (vs. the retrospective pretest) group, the adjusted mean score was significantly lower for those living with one parent (vs. both) (mean -0.34, $p < .0001$) or other (vs both) (-0.61, $p = 0.0021$).
- For monitoring, in addition to there being a statistically significantly higher adjusted mean score in the pretest (vs. the retrospective pretest) group, the adjusted mean score was significantly higher for females (vs. males) (mean 0.22, $p = 0.0072$).
- For mental health, there was not a statistically significant difference in the adjusted mean score in the pretest compared with the retrospective pretest. However, the adjusted mean mental health score was significantly higher for youth ages 17-18 (vs. 9-10) (mean 0.72, $p = 0.0484$), females (vs. males) (mean 0.36, $p = 0.0005$), Hispanics (vs. non-Hispanics) (mean 0.41, $p = 0.0081$), and for those living with other (vs both) 1.25, $p < .0001$).
- For good behavior, in addition to there being a statistically significantly higher adjusted mean score in the pretest (vs. the retrospective pretest) group, the adjusted mean score was significantly lower for those living with one parent (vs. both) (mean -0.59, $p < .0001$) or other (vs both) (-1.15, $p < .0001$).
- For bad behavior, there was not a statistically significant difference in the adjusted mean score in the pretest compared with the retrospective pretest. The mean score was significantly higher in those living with other compared with both parents (0.93, $p < .0001$).
- For family dysfunction, there was not a statistically significant difference in the adjusted mean score in the pretest compared with the retrospective pretest. The adjusted mean family dysfunction score was significantly higher for females (vs. males) (mean 0.22, $p = 0.0252$), other race (vs. Whites) (mean 0.33, $p = 0.0356$), and for those living with only one parent (vs both) 0.33, $p = 0.0097$).

Interaction terms involving the pretest versus the retrospective pretest and the demographic variables were not statistically significant. Hence, beyond some demographic variables being possible confounders, none were modifiers.

Table 5. Multiple Regression Evaluation Comparing Youth Pretest versus Retrospective Pretest Scores for Each Construct

	Pretest	Retrospective Pretest	
Constructs	**Adjusted Mean**	**Adjusted Mean**	**t p value**
Bonding	2.62	2.30	<.001
Boundaries	3.18	2.93	0.000
Monitoring	2.92	2.57	<.001
Mental Health	1.58	1.66	0.424
Good Behavior	2.69	2.43	<.001
Bad Behavior	1.33	1.31	0.709
Family Dysfunction	1.83	1.90	0.485

Note. Means were adjusted for age, sex, ethnicity, and parent. Scale for each item: 0 = Never to 4 = Almost Always. "I am happy most of the time" was reverse scored.

Discussion

It is important to distinguish between two analytic goals: identifying the range of baseline perceptions (using the pretest) and measuring perception changes (through comparison with the retrospective pretest). This study emphasizes the latter to evaluate how the intervention shifted participants' understanding of their behaviors and family functioning.

For many of the items, mean response scores were significantly higher in the pretest than the retrospective pretest. As previously suggested, this is likely due to an initial lack of trust in the service delivery, lower mindfulness of what adequate parenting and family entail, and/or a desire to appear more socially appropriate (Kumpfer et al., 2010). However, another plausible explanation, reinforced by Walsh's Family Resilience Framework (2003), is that the intervention itself shifted participants' understanding of family support systems, leading them to reassess their earlier behaviors more critically at the retrospective pretest. In a study involving child-welfare participating in the SFP, all but one of 18 items statistically significantly overstated their caregiving behaviors and abilities (Brook et al., 2016). The results in the current study are consistent with these results in that parents consistently provided more favorable parenting and family functioning at the pretest. However, we found a smaller percentage of statistically significant differences, which previous research has suggested is likely because our study population did not focus on child-welfare involved families who are more prone to social desirability responses (Kumpfer et al., 2010).

These findings are consistent with suggested relationships using the Family Resilience Framework (Walsh, 2003). Participants' retrospective evaluations indicated greater awareness and a more realistic perception of their family belief systems, organizational patterns, and communication skills after the intervention. This heightened awareness may reflect not only reduced bias but also the intervention's success in reshaping participants' internal frameworks for evaluating their family relationships, particularly through the reflective and skills-based components emphasized in the Family Resilience Framework. This aligns with Walsh's (2003) assertion that quality interventions cannot simply focus on skill acquisition but should

also account for how different families may respond to stress and adapt to new challenges.

For the parents, the largest difference between the pretest and retrospective pretest was for the bonding, boundaries, monitoring, and good behavior (marginally insignificant) constructs. The intervention was specifically designed to improve parenting skills related to bonding, boundaries, and monitoring, where bonding reflected family cohesion, boundaries reflected family rules and organization, and monitoring and the other two constructs reflected communication and social skills. The lower score in the retrospective pretest versus the pretest for child good behavior may be because parents better recognized such behavior as they progressed through the intervention. The child's mental health, child's bad behavior, and family dysfunction constructs did not statistically significantly differ between the pretest and retrospective pretest. The mental health construct reflects more of what exists and is low to begin with. The child's bad behavior and family dysfunction constructs began low, with less room to go lower. Nevertheless, we expected that the parent's perception of what is socially acceptable would have caused the pretest score to be comparatively higher than the retrospective pretest score.

For the youths, the largest difference between the pretest and retrospective pretest was for the bonding, boundaries, monitoring, and good behavior constructs. The intervention was specifically designed to influence these constructs, suggesting that the difference is due to a possible combination of initial lack of trust in the service delivery, lower mindfulness of what is appropriate, and/or a desire to appear more socially appropriate. The lower score in the retrospective pretest for child good behavior may be because the child better recognized what constitutes good behavior. The child's mental health, child's bad behavior, and family dysfunction constructs did not statistically significantly differ between the pretest and retrospective pretest. Again, the mental health construct reflects a state of fact and is relatively low. The child's bad behavior and family dysfunction constructs also were relatively low, with less room to become lower.

Social-desirability bias in survey responses is over-reporting "good" behavior or under-reporting "bad" behavior due to self-perception concerns (Krumpal, 2011). Of interest is whether social-desirability bias played a different role in the parents than their youth, depending on the construct. Differential social-desirability bias was not generally observed. However, an exception was that parents assigned lower mental health scores for their children, which persisted at both pretest and retrospective pretest. This result suggests that parents tend to think their child's mental health reflects upon themselves (Karimzadeh et al., 2017). In addition, at the pretest (vs. the retrospective pretest) parents had more favorable monitoring scores compared with their youth and youth had more favorable child good behavior scores compared with their parent's. These results are consistent with parents having the primary responsibility for monitoring and youth having the primary responsibility for their behavior. As for the specific items, in the pretest youth statistically significantly overestimated certain socially acceptable behaviors that they had primary responsibility for, whereas the parents did not overestimate these items in their assessment of their youth. These items were knowing how to say "no" to stay out of trouble, their parent/caregiver checking daily to see if they completed their homework, paying attention in school, and staying on a task until it is completed.

The difference in construct scores between the pretest and retrospective pretest

was not dependent on any of the demographic variables. In other words, overestimation at the pretest for some of the constructs was not dependent on the level of the demographic variables. Thus, overestimation bias due to factors like trust, social desirability, and mindfulness appears to be a similar threat regardless of who the child lives with (mother, father, both, neither), sex, age, race, and ethnicity. The possibility that intervention participation itself fundamentally enhanced families' resourcefulness, which Walsh (2003) suggests helps families manage perceived risks and vulnerabilities, must also be considered.

Limitations

Certain limitations should be considered. First, despite 24 SFP recruiters from agencies that were racially and demographically diverse across the United States who were given the task of selecting racially/ethnically diverse families into the study, non-Hispanic Whites were overrepresented and Blacks, Hispanics, and other racial groups were underrepresented. For example, the U.S. Census reported that in 2022 the population consisted of 58.9% Whites, 12.6% Blacks, and 19.1% Hispanics (USA Facts, 2022). Second, participation was voluntary, such that the predominately White parents may have been more motivated than the parents in the general population because of unique parenting and family functioning characteristics. However, comparison of the demographic variables between those recruited who went on to complete the intervention versus those who did not show statistically significant differences. Third, we could not determine whether factors identified as causing overestimation at the pretest may have persisted through the retrospective pretest. Fourth, the magnitude of overestimation in the pretest could have differed among different subpopulations, although our tests of interaction did not suggest this to be the case.

Conclusions

In self-reported surveys measuring items related to social desirability, there is a tendency to give responses that are more socially acceptable. This may be particularly true in measuring parenting and family functioning. Prior to receiving an intervention, responses may lead to greater social desirability due to lack of trust in the service delivery, lower mindfulness about adequate parenting family functioning, and desire to appear more socially appropriate. While retrospective pretests help reduce overestimation bias, the observed changes may also reflect meaningful shifts in how participants view their family dynamics because of the intervention—suggesting both a methodological correction and a substantive psychological impact. Future research might explore how presenting both the range of pre-intervention responses, and the degree of post-intervention change can offer a more comprehensive picture of program impact and participant diversity.

Implications for Health Behavior Research

The findings of this study suggest that retrospective pretests should be a core consideration for any intervention-based research relying on self-reported survey items. By providing a more reliable baseline, researchers can gain greater confidence in the true impact of interventions designed to address sensitive areas like parenting and family dynamics. This can lead to more effective tailoring of interventions and a better understanding of overall program outcomes. The value of retrospective pretests is likely even higher for any intervention-

based study comparing impacts across continually evolving digital learning environments.

Discussion Questions

For the parents and youth, why was the largest difference between the pretest and retrospective pretest among the constructs bonding, boundaries, monitoring, and good behavior?

At the pretest (vs. the retrospective pretest), why did parents have more favorable monitoring scores compared with their youth and youth have more favorable child good behavior scores compared with their parent's?

Why was the level of overestimation bias at the pretest due to factors like trust, social desirability, and mindfulness not dependent on who the child lives with (mother, father, both, neither), sex, age, race, and ethnicity?

References

Brook, J., Akin, B. A., Lloyd, M., Bhattarai, J., & McDonald, T. P. (2016). The use of prospective versus retrospective pretests with child-welfare involved families. *Journal of Child and Family Studies*, *25*, 2740–2752. https://doi.org/10.1007/s10826-016-0446-1

Deutsch, M., & Collins, M. E. (1951). *Interracial housing: A psychological evaluation of a social experiment*. Minneapolis: University of Minnesota Press.

English, J. L., & Horowitz, J. (2002). Retrospective pretesting in health education settings. *American Journal of Health Education*, *33*(1), 17–22. https://doi.org/10.1080/19325037.2002.10604715

Karimzadeh, M., Rostami, M., Teymouri, R., Moazzen, Z., & Tahmasebi, S. (2017). The association between parental mental health and behavioral disorders in pre-school children. *Electron Physician*, *9*(6), 4497-4502. https://doi.org/10.19082/4497

Krumpal, I. (2011). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity*, *47*(4), 2025-2047. https://doi.org/10.1007/s11135-011-9640-9

Kumpfer, K., Whiteside, H., Greene, J. A., & Allen, K. C. (2010). Effectiveness outcomes of four age versions of the Strengthening Families Program in statewide field sites. *Group Dynamics: Theory, Research, and Practice*, *14*(3), 211–229. https://doi.org/10.1037/a0020602

Kumpfer, K. L, & Hansen, W. B. (2014). Family-based prevention programs. In: Scheier LM, Hansen WB, eds. *Parenting and teen drug use: The most recent findings from research, prevention, and treatment.* New York, NY: Oxford University Press; 166–192.

LoBraico, E. J., Fosco, G. M., Crowley, D. M., Redmond, C., Spoth, R. L., & Feinberg, M. E. (2019). Examining intervention component dosage effects on substance use initiation in the Strengthening Families Program: For parents and youth ages 10-14. *Prevention Science*, *20*, 852–862. https://doi.org/10.1007/s11121-019-00994-7

Piwowar, V., & Thiel, F. (2014). Evaluating response shift in training evaluation: Comparing the retrospective pretest with an adapted measurement invariance approach in a classroom management training program. *Evaluation Review*, *38*(5), 420–448. https://doi.org/10.1177/0193841X14546932

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.

USA Facts. *Our Changing Population: United States*. 2022. Retrieved from https://usafacts.org/data/topics/people-society/population-and-demographics/our-changing-population/?endDate=2022-01-01&startDate=2010-01-01

Walk, R. D. (1956). Self ratings of fear in a fear-invoking situation. *The Journal of Abnormal and Social Psychology, 52*(2), 171–178. https://doi.org/10.1037/h0042978

Walsh F. (2003). Family resilience: A framework for clinical practice. *Family Process, 42*(1), 1–18. https://doi.org/10.1111/j.1545-5300.2003.00001.x