HMRE
Local Evaluation Technical Assistance

Local Evaluation Findings from the 2020 Cohort of Healthy Marriage and Relationship Education Grant Recipients

July 2026

This brief presents selected results from local evaluations of Healthy Marriage and Relationship Education (HMRE) grant recipients serving adults and youth (Box 1). It is intended to share insights with practitioners, researchers, funders, and policymakers that shed light on promising practices that might help programs improve services to strengthen families. First, this brief provides an overview of the services offered by HMRE programs and the design of the local evaluations. Next, it summarizes the findings from the local evaluations, then presents the successes of evaluations and the challenges they faced. It concludes with a summary of key takeaways and next steps for the field.

Background on HMRE programming and evaluations

HMRE grant recipients support families by offering a range of services for adult individuals, adult couples, and youth (up to age 24) to increase their knowledge of healthy relationships and strengthen skills such as communication and co-parenting. Grant recipients typically provide programming through group-based workshops focused on relationship skills, parenting, and economic self-sufficiency. All grant recipients in this cohort provided at least 12 hours of programming—commonly using evidence-based curricula—which were usually spread out over multiple sessions and several weeks. Many HMRE programs also offered optional individual services such as case management, employment coaching, and financial literacy.

Box 1. Overview of the Healthy Marriage and Responsible Fatherhood Program

The Healthy Marriage and Responsible Fatherhood (HMRF) program, funded by the Administration for Children and Families, Office of Family Assistance, supports community organizations in their efforts to strengthen families. Through competitive grants, HMRF funds programs that offer relationship and marriage education, responsible fatherhood services, and related supports designed to help adults and youth build healthy relationships, strengthen parenting, and improve family stability. Grant recipients deliver services in ways that fit their communities and participants, often adapting content and delivery to local needs. Alongside service delivery, HMRF emphasizes learning and improvement by supporting evaluation and technical assistance so programs can understand what is working, strengthen their approaches, and share lessons with the broader field.

Mathematica
Progress Together

HMRF
HEALTHY MARRIAGE & RESPONSIBLE FATHERHOOD

ADMINISTRATION FOR CHILDREN & FAMILIES

HMRE programs can opt to conduct local evaluations of their program by working with an independent evaluator to design and execute a study to address questions of interest about grant recipients' services and the people they serve. Local evaluations allow HMRE programs to generate evidence about what works to improve participant outcomes such as relationships, parenting, or economic stability. The Administration for Children and Families (ACF) invests in local evaluations to help grant recipients improve program quality and build evidence to strengthen families. These evaluations also contribute to ACF's national research portfolio by identifying promising practices and shaping funding or program priorities.

In the 2020 cohort of HMRE grant recipients, 39 programs completed a local evaluation. Grant recipients used different evaluation types designed to address a range of questions about HMRE programs (Box 2). Although each evaluation is specific to the grant recipient's program, there are many commonalities in the programming and local evaluations, making it possible to summarize the findings across the evaluations.

Box 2. Understanding evaluation types

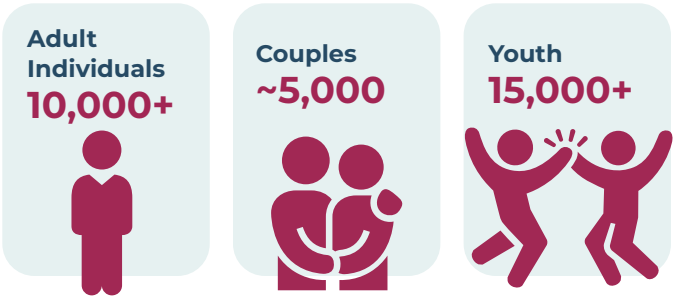
Descriptive studies describe what happened in a program and help explain who participated, how services were delivered, and how participant outcomes changed across time. They can point to patterns or associations but cannot show that a program directly caused an outcome. **Impact studies**, in contrast, are designed to test whether a program made a difference by comparing outcomes for participants who received the program with outcomes for a similar group who did not. This comparison helps us determine whether any changes we observed were likely due to the program itself. There are two types of impact studies—randomized controlled trials (RCTs) and quasi-experimental designs (QEDs). RCTs randomly assign participants to two groups for the study, whereas QEDs use statistical procedures to match two similar groups.

Thirty-nine grant recipients completed either an impact (12 RCTs and three QEDs) or descriptive (24 evaluations) study. Twelve impact studies focused on adults (individuals or couples) and three focused on youth. Eleven descriptive evaluations studied adults and 13 studied youth. Across all grant recipients conducting impact studies, more than 10,000 participants were enrolled in the evaluations, and for those conducting descriptive studies, almost 20,000 were enrolled in the evaluations.

What services did HMRE programs offer?

HMRE grant recipients designed evaluations to study services offered to three focal populations: adult individuals, adult couples, and youth. Nine of these grant recipients served adult individuals, 13 served adult couples, and 21 served youth. A total of almost 30,000 adults, couples, and youth were enrolled across the 39 grant programs conducting local evaluations (Figure 1 gives a specific breakdown by focal population).

Figure 1. Reach of HMRE local evaluations



The HMRE programs conducting local evaluations offered many of the same curricula that are widely known in the field of relationship education. The most frequently used workshop curricula for programs serving adults were the Prevention and Relationship Enhancement Program (PREP 8.0) (27 percent) and Within My Reach (18 percent). About one-third of adult programs (27 percent) used more than one curriculum in their workshops. For example, Montefiore Medical Center, which served parents in the Bronx, combined two curricula: Bringing Home Baby and PREP 8.0. For programs serving youth, the most used curricula were Love Notes (53 percent) and Mind Matters (40 percent). More than half of youth programs (56 percent) used more than one curriculum in their workshops—for example, Better Family

Life, which served high school students in St. Louis, Missouri, combined Love Notes, Mind Matters, and Skills to Pay the Bills. (Appendix A has more about the programs, including dosage and setting.)

Across the programs conducting a local evaluation, grant recipients offered on average 15 hours of workshop content over about eight weeks. The vast majority of programming for adults (82 percent) was offered in community settings, whereas the vast majority of youth programming (81 percent) was offered in schools (Appendix A has details). Some grant recipients gave participants flexibility in how they completed the workshops, such as offering condensed versions of the curricula over two consecutive Saturdays or delivering the program content in longer sessions over four weeks instead of shorter sessions over eight weeks. Planned workshops were delivered in person, virtually, or as a hybrid of both (activities contained virtual and in-person components). Twenty-two grant recipients offered hybrid services; 13 offered only in-person services; and three offered virtual services.[1,2]

How were the local evaluations designed?

The evaluation designs of impact and descriptive studies varied on features such as data collection processes and whether they were looking for differences across groups or within groups over time.

Impact study designs. Of the 15 impact evaluations conducted, nine collected follow-up data six months after the program ended to answer their primary research questions.[3] However, several evaluations collected data at shorter or longer intervals—three months, nine months, or 12 months after the program ended (two evaluations for each interval).[4]

To address different types of research questions, programs designed their impact evaluations in one of three ways: enhancement tests, equivalent effects tests, and program impact tests (Box 3 gives definitions of study designs).

First, 10 impact evaluations tested an enhancement of the services offered by the grant recipient. These evaluations were typically designed to offer the same core services—like a curriculum-based workshop—to two study groups, and offer an additional component—such as employment coaching—to only one of the two groups. Then, the evaluations tested whether the added service enhanced the benefit of the program for participants. For example, Champions for Children, serving parents in Tampa, Florida, did an enhancement test to ask whether there was an impact of adding the parenting curriculum Triple P–Positive Parenting Program to its established program Positive Parenting and Partnership (P3), comparing that to offering a stand-alone version of P3. Among these 10 enhancement tests, three impact evaluations tested curricula head-to-head to see if one curriculum improved participant outcomes more than the other. For example, Live the Life Ministries, which served couples in Tallahassee, Florida, tested its Adventures in Marriage curriculum delivered in person against a bundle of two other curricula delivered asynchronously through videotaped sessions.

[1] This brief focuses on describing the program services offered by the 39 HMRE grant recipients that conducted an evaluation.

[2] Four of the impact studies were counted as hybrid although the studies were testing whether the mode of services mattered for participant outcomes. Thus, for these evaluations, the treatment and comparison groups received services in only one mode, in person or virtual, instead of having access to both in-person and virtual activities the hybrid sites offered.

[3] All evaluations, regardless of design, could answer primary and secondary research questions. This brief focuses on the primary research questions because ACF provided standardized guidance to programs on using rigorous methods to analyze those questions, whereas the secondary research questions were considered exploratory and not subject to the same analysis standards.

[4] Guidance from the Office of Family Assistance required impact studies to use data collected at minimum three months after the end of the program for their primary research questions.

Second, three evaluations tested for equivalent effects. These evaluations were designed to offer the same content to two groups but deliver it through different methods, such as virtual versus in-person delivery. The evaluations asked whether the differences between the participant outcomes of two study groups were small enough to be considered equal, meaning no one way of delivering the program had a larger impact on participants compared to the other way. Montefiore Medical Center set out to conduct an equivalence test to find out if delivering its program to couples virtually or in person led to the same kinds of outcomes for participants. Finally, two impact evaluations tested the impacts of offering program services relative to offering no services. For this study design, the outcomes of participants who were offered the HMRE program services were compared to the outcomes of participants who were not offered program services. For example, Jewish Family & Children's Service of the Suncoast Inc., serving adults in Sarasota, Florida, tested whether offering Within My Reach to the program group led to better relationship quality outcomes for that group compared with the group that did not receive Within My Reach.

Box 3. Impact study designs: Enhancement, equivalent effects, and program impact tests

- **Enhancement tests.** Most enhancement tests examine whether adding a new service component improves outcomes beyond what a program's core services would do. In these studies, both groups receive the same main program, but one group receives an additional service or content, such as financial literacy, employment services, or job coaching supports, to assess whether the enhancement provides added benefit for participants. In some cases, these studies test two different combinations of services head-to-head to understand which approach leads to better participant outcomes.
- **Equivalent effects tests.** Typically, impact evaluations are designed to offer two or more sets of services and answer the question, "Is there a difference in participant outcomes?" But sometimes programs are interested in evaluating whether they can expect the same effects on participant outcomes from different sets of services. For example, given interest in offering programming virtually, two HMRE evaluations in this cohort conducted equivalent effects tests to examine whether completing the same program in person or virtually led to similar participant outcomes. In this case, evaluations must use equivalent effects testing to ask, "Is the difference in the outcomes of the groups small enough that it doesn't matter in practice?" To do this, researchers decide during the design phase how big a difference would need to be to be meaningful. Then they test whether the observed difference in effects is smaller than that amount—that is, whether the effects are equivalent.
- **Program impact tests.** Program impact tests assess whether a program or curriculum makes a difference by comparing outcomes for participants who receive services with outcomes for a group that does not receive services. This design is often used to answer the foundational question of whether a program causes improvements in participant outcomes.

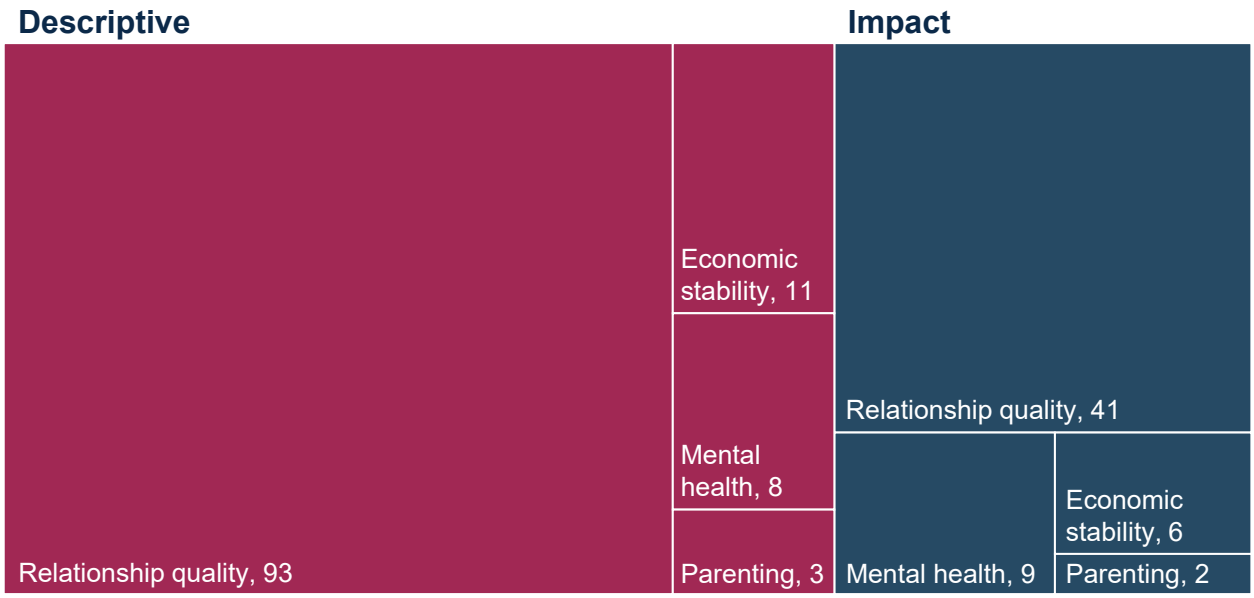
Descriptive study designs. All but one of the descriptive studies used a pre-post design, which is used to assess change in the participant outcomes of interest over time, from program entry ("pre") to a point in time after the program ends ("post"). Nearly all the descriptive evaluations (20 studies) collected data to answer their primary research questions at the time participants exited the program. A few studies used other time points, ranging from three months after the program ended to nearly 12 months after the program ended. One grant recipient conducting a descriptive evaluation did not use a pre-post design as its primary approach: Anthem Strong Families, which served adults in Dallas-Fort Worth, Texas. Anthem set out to implement and test a continuous quality improvement process to support the delivery of its Family Champions Project's services with the goal of improving services over time.[5]

[5] Because Anthem's design tested a continuous quality improvement process, Anthem is excluded on counts related to evaluation design and participant outcomes. Anthem did not test participant outcomes as primary research questions.

What participant outcomes did the evaluations test?

Each outcome measure the grant recipients tested was grouped into one of four outcome domains: relationship quality, economic stability, mental health, and parenting. For example, measures of relationship satisfaction, conflict, and communication were grouped under the relationship quality domain. Outcomes in the relationship quality domain were tested most frequently regardless of study design, with 41 outcomes tested across impact evaluations and 93 tested by descriptive evaluations (Figure 2). Studies across both evaluation types tested 17 outcomes in each of the economic stability domain and mental health domains. In contrast, evaluations tested the fewest outcomes in the parenting domain (five outcomes).[6] Notably, no evaluations of youth programming focused on parenting outcomes. This might be because none of the impact studies and only two descriptive studies had pregnant or parenting youth as a focal population, so the outcome measures in that domain were unlikely to align with the curriculum and content of those grant recipients' programs.

Figure 2. Number of outcomes tested in each domain, by study type



What evidence did the local evaluations find on participant outcomes?

The local evaluations focused on a range of outcome measures for each of the four primary domains (Box 4). Across local evaluations, regardless of the population served, evidence was most concentrated in outcomes reflecting relationship quality, whereas findings related to economic stability, mental health, and parenting were more limited. The findings on these three domains often depended on when outcomes were measured. Favorable findings, or those results that are statistically significant and in the expected direction, were found more often for outcomes measured right at the end of the program as opposed to those measured three months or more after the program ended. These findings highlight the notable differences in what evaluations can detect across domains.

[6] Six outcomes from descriptive studies are not reported because they did not map to any of the four domains of interest in this brief. Three of the six outcomes were the same measure tested at different time points by one grant recipient. The grant recipient combined several areas (anxiety, gratitude, and couple satisfaction) into a single measure that did not align with the four domains addressed in this brief. The remaining three outcomes were tested by two grant recipients and focused on general interpersonal skills that were not specific to a relationship context.

Box 4. Methods for summarizing findings

Each outcome measure the grant recipients tested was grouped into one of the outcome domains: relationship quality, economic stability, mental health, and parenting. For example, measures of satisfaction, conflict, and communication were grouped under the relationship quality domain. All evaluations tested statistical significance at the $p < .05$ level, meaning there would be less than a 5 percent chance of observing results of that size or larger if there were no true difference between groups in the outcomes measured. In the findings section, favorable findings refer to results that are statistically significant and in the expected direction. To support interpretation of practical significance, impact evaluations also reported effect sizes (using Hedges' g or Cox's Index), which help classify the difference as small (.20), medium (.50), or large (.80).

Impact evaluations with favorable findings

Seven of the 15 impact evaluations conducted by HMRE grant recipients reported favorable findings in at least one of the four outcome domains. Evaluations of adult programs demonstrated the most favorable findings within the relationship quality domain. In contrast, only one of the three impact evaluations of youth programs reported a favorable finding in any domain. Effect sizes for impact studies were between 0.16 to 0.36, which is within the expected range for HMRE evaluations (small to medium-sized effects).[7]

Seven impact evaluations reported favorable findings.

Of the seven evaluations with favorable findings, two were enhancement tests, three were equivalent effects tests, and two were program impact tests. Table 1 shows how many of the seven impact studies had favorable findings in each outcome domain.

Table 1. Distribution of favorable findings by outcome domain: seven impact studies

Domain	Enhancement tests	Equivalent effects tests	Program impact tests	Total
Relationship quality	1	2	2	**5**
Economic stability	1	0	1	**2**
Mental health	1	2	0	**3**

Note: Counts are not mutually exclusive. Table includes only impact evaluations with at least one favorable finding (seven evaluations). Evaluations may contribute to more than one outcome domain. Favorable findings are defined as those that are statistically significant and in the expected or hypothesized direction. In addition, for impact studies, favorable outcomes are counted only if they demonstrate baseline equivalence, meaning the groups being compared on that outcome were similar at the start of the study.

Enhancement tests. The two studies that were designed to test enhancements to their program and had favorable impacts served adults, offered content across eight weeks, and were established programs that offered similar services in their communities before they became grant recipients in this cohort. Notably, both programs had evidence from prior evaluations that their core services were effective in improving

[7] The effect size findings for the 2020 cohort are similar to those for the 2015 cohort. See Avellar, Sarah, Leah Shiferaw, Christine Ross, and Joanne Lee. "Supporting Healthy Relationships: Final Report on the 2015 Cohort of Healthy Marriage Grantees Serving Adults." OPRE Report 2021-170. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021.; Avellar et al. "Supporting Current and Future Healthy Families: Final Report on the 2015 Cohort of Healthy Marriage Grantees Serving Youth." OPRE Report 2021-171. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021.; and Lee, Joanne, Leah Shiferaw, Sarah Avellar, Christine Ross, and Alexandra Stanczyk. "Methodology and Supplemental Findings for Final Reports on the 2015 Cohort of Healthy Marriage and Responsible Fatherhood Grantees." OPRE Report 2021-177, Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021.

outcomes for participants. Building on this knowledge, the grant recipients wanted to know how adding service components could bolster participant outcomes. Both programs integrated robust curricula enhancements related to parenting and economic well-being to their core relationship education workshops.

For programs with demonstrated effectiveness of relationship education services, integrating additional content can bolster participant outcomes.

The evaluations found positive effects for the group receiving the enhanced program on three domains: relationship quality, economic stability, and mental health. One of the evaluations reported fewer depression symptoms (mental health domain) and the other evaluation reported improvements in relationship satisfaction and skills (relationship quality domain) and financial capability (economic stability domain). The latter, The Parenting Center, demonstrated how integrating curriculum-based content on employment and financial literacy led to enhancing impacts on economic stability outcomes for couples (Box 5).

Spotlight: Key findings for practice

Box 5. The Parenting Center integrated employment coaching and financial planning to improve outcomes for couples

The Empowering Families Project tested whether adding economics-focused services enhanced the effects of relationship education for low-income couples. Using a randomized controlled trial, the evaluation compared couples who received relationship education alone with couples who received the same relationship education plus financial planning and career coaching. Couples who received the enhanced program showed greater improvements in relationship satisfaction, positive communication, and financial capability one year later, whereas outcomes related to negative communication, income, and financial strain were similar across groups. Although the added effects were small, the findings demonstrate that layering economic services onto an existing relationship education program can provide additional benefits when services are well implemented and participants actively engage.

Equivalent effects tests. The three studies with favorable findings that conducted equivalent effects tests were interested in showing that alternative modes to delivering workshop content could lead to similar participant outcomes. One of the programs served high school youth in schools and the other two programs served adult couples in the community. The evaluation of the youth program examined whether using different facilitators to deliver the curriculum led to equivalent outcomes for high schoolers. The other two evaluations were tests of mode—the programs wanted to find out whether delivering the same set of services virtually or in person led to the same outcomes for couples.

Equivalent effects testing requires larger-than-normal sample sizes to adequately detect differences or similarities between study groups, which has implications for the program's goals for recruitment, enrollment, and retention.

The studies showed favorable results (equivalence) in two outcome domains. Auburn University's evaluation of youth showed favorable results in beliefs and attitudes about marriage (relationship domain) and self-regulation and mindfulness (mental health domain). The evaluation demonstrated that outcomes in these two domains were equivalent for youth regardless of whether the content was delivered by a college-aged peer or an experienced community educator (Box 6). The other two evaluations showed favorable findings on couple satisfaction (relationship quality domain) and self-regulation and mindfulness (mental health domain). Regardless of whether the content was delivered virtually or in person, these two evaluations of adult programs showed that a limited number of participant outcomes were essentially equivalent.

Two of these evaluations demonstrated equivalence on self-regulation and mindfulness. Both programs used curricula that highlighted mindfulness concepts and practice activities as a skill (Mind Matters and ELEVATE). These evaluation findings show promising evidence that content delivered to strengthen mindfulness and self-regulation skills may be particularly unaffected by delivery method or mode and could be an area of further study. Notably, a limitation of equivalent effects testing is that to achieve adequate statistical power, the sample size for the evaluation must be larger than it is for other designs, such as an enhancement study or comparing a program group to a no-services group. The decision by a program to evaluate equivalent effects of two (or more) modes of service delivery has implications for the program's recruitment, capacity to serve participants, and ability to retain participants in the program and evaluation. These three programs were established in their communities, which also played a role in grounding recruitment expectations based on historical trends and building on knowledge gained through prior evaluations of the program itself.

Promising evidence suggests programs can use various delivery methods or modes to successfully strengthen mindfulness and self-regulation skills.

Spotlight: Key findings for practice

Box 6. Auburn University's study demonstrates equivalent effects of two facilitation methods

Auburn University evaluated the Alabama Youth Relationship Education (AYRE) program using a quasi-experimental design to test whether two facilitation models, using college-age near-peer facilitators (NPs) or experienced community educators (CEs), produced equivalent effects on outcomes for high school students receiving the AYRE curriculum (combined from the Relationship Smarts PLUS 4.0, Mind Matters, and Money Habitudes for Teens curricula). Results showed that six months after the program ended, youth in both groups were equivalent on key outcomes related to tolerating dating aggression and self-regulation. Although having facilitators who were CEs was associated with slightly higher youth-reported knowledge about relationship skills, NPs fostered greater youth-reported self-efficacy in relationships. The findings indicate that multiple facilitation models can support positive youth outcomes in school-based relationship education. NPs may be well suited for building youth confidence and engagement, whereas CEs may bring strengths in delivering relationship knowledge. Programs can draw on these insights when selecting or combining facilitation approaches to best match their goals, resources, and local context.

Program impact tests. Two of the seven studies with favorable findings tested the effectiveness of the program services by comparing a group receiving services to a wait-listed comparison group, meaning the participants would receive services after their participation in the evaluation ended. Like the other evaluations with favorable findings, these two programs were established and had experience implementing and evaluating their services. Both programs served adult couples and offered services wholly or partly virtually; however one offered synchronous content and the other offered it asynchronously.

The two studies showed favorable results in the relationship quality and economic stability domains. One of the two studies used the Within My Reach curriculum, delivered in person, with an option to join classes virtually. The program offered a dedicated workshop session on financial literacy with optional financial literacy activities for interested participants. This evaluation's findings were well aligned with the offered content, showing positive effects on relationship conflict (relationship quality domain) and economic self-sufficiency. The other evaluation,

Virtual services provide a cost-efficient, effective alternative for participants who need flexibility in attendance options.

conducted by the University of Miami, offered virtual, asynchronous workshops to couples with the OurRelationship curriculum. The evaluation found higher relationship satisfaction and lower relationship conflict (relationship quality domain) three months after the program ended, showing evidence that online services can have positive sustained effects for couples and offer a cost-efficient way to deliver services (Box 7). Notably, this study found the largest effect sizes of any evaluation with favorable findings—0.36 for relationship conflict and 0.30 for relationship satisfaction.

Spotlight: Key findings for practice

Box 7. University of Miami demonstrates the effectiveness and potential cost savings of virtual service offerings for couples

The University of Miami conducted a randomized controlled trial of the OurRelationship program, delivering it nationwide in an online format to couples with low incomes, couples in the military, and veteran couples. The study found that couples assigned to the OurRelationship curriculum reported higher levels of relationship satisfaction and lower levels of communication conflict six months after enrolling in the program, in contrast with a comparison group that was not offered the program. Impacts were similar for military or veteran and non-military or non-veteran couples, and the program was delivered with high fidelity. The accompanying cost study showed that these positive impacts were achieved at a relatively low cost. During the evaluation period, the Miami team served more than 2,000 couples at an average cost of about $914 per couple (or $457 per individual) which is substantially lower than many in-person relationship education programs evaluated in prior studies. Together, the impact and cost findings present preliminary evidence that online relationship education programs can be a cost-efficient option for reaching couples and improving relationship outcomes, particularly when in-person services are difficult to deliver or scale.

Descriptive evaluations with favorable findings

20 descriptive evaluations reported favorable findings.

Twenty of the 24 descriptive evaluations had at least one favorable finding.[8] Most of the evaluations had at least one favorable finding in the relationship quality domain (16 studies), followed by five studies with favorable findings in mental health, four with favorable findings on economic stability, and one study with favorable findings on parenting (Table 2). About half of these 20 evaluations focused on grant recipients that served youth (five served youth in schools; four served youth in school and outside of school settings; and two operated programs outside of school settings only); and the other half served adults (five evaluations of programs for couples and four of programs for individuals).

In contrast with impact studies, descriptive outcomes were mainly collected at program exit (16 of the 20 evaluations). Although descriptive evaluations cannot conclusively point to the program as the only reason for observing change in participant outcomes, these studies often provide information about what kinds of outcomes are associated with program services and the adaptations made while implementing services. Thus, descriptive evaluations play a critical role in identifying promising evidence about what works in relationship education. These contributions inform future programs about whether to test services using an impact evaluation or dig deeper to produce more descriptive evidence.

[8] Only one of these descriptive evaluations had high attrition—or sample loss—in the range of 60 to 70 percent, depending on the outcome measure. This evaluation is included in all counts in this brief, but the high attrition makes the findings unlikely to reflect the broader population that enrolled in the evaluation.

Table 2. Distribution of favorable findings by outcome domain: 20 descriptive evaluations

Domain	Adult	Youth	Total
Relationship quality	7	9	**16**
Economic stability	2	2	**4**
Mental health	4	1	**5**
Parenting	1	0	**1**

Note: Counts are not mutually exclusive. Table includes only descriptive evaluations with at least one favorable finding (20 evaluations). Evaluations may contribute to more than one outcome domain. Favorable findings are defined as those findings that are statistically significant and in the expected or hypothesized direction.

Relationship quality domain. Sixteen evaluations reported favorable findings in the relationship quality domain, most often on measures of relationship skills (18 findings), beliefs and attitudes about marriage (12 findings), and relationship conflict (12 findings). The findings in this domain were robust across adult- and youth-focused evaluations. Most descriptive studies used a curriculum that aligned with improving relationship quality, skills, and attitudes; thus, the research questions posed and the number of findings follow directly from the content covered most often by programs.

Promising descriptive evidence suggests partnering with trusted community organizations and using culturally relevant relationship education helps youth build positive attitudes about marriage and strengthens core relationship skills.

For example, The Office of Samoan Affairs' (OSA) WRAP program combined and tailored three curricula—Love Notes, Mind Matters, and Money Habitudes for Teens—for youth in the community. The evaluation found improvements in relationship skills, particularly in attitudes about marriage and confidence in handling disagreements. Many participants also reported ending unhealthy relationships and feeling better prepared to apply healthy relationship skills in their own lives. OSA partly attributed the positive changes youth reported from the start to the end of the program to partnering with trusted community networks to deliver the tailored curriculum. Similarly, Be Strong International used the Prep 8.0 curricula to serve couples, also offering minimal content on parenting and financial literacy. This evaluation found favorable findings on relationships skills, satisfaction, and conflict (Box 8).

Spotlight: Key findings for practice

Box 8. Be Strong International shows how programs delivered virtually can improve relationship skills

Be Strong International's descriptive evaluation of the Marriage Matters program found that couples showed significant improvements in relationship skills, including marital self-efficacy, and reductions in negative relationship dynamics, even though most sessions were delivered virtually. Participants highlighted skilled facilitation, interactive discussions focused on communication and conflict resolution, and culturally responsive delivery in English and Spanish as key factors that helped them stay engaged and apply what they learned. The study provides promising evidence that virtual relationship education can maintain high engagement when facilitators prioritize interaction, create a warm and supportive environment, and intentionally design sessions as meaningful couple time.

Parenting domain. Only two of 24 descriptive evaluations examined parenting outcomes, and only one of the two reported favorable findings in this domain. The lack of programs studying parenting outcomes points to a significant gap and opportunity for learning more about how relationship education can improve those outcomes.

The one descriptive evaluation with favorable findings in this domain studied a program serving adults living in stepfamilies that offered services virtually and in person. The program explicitly included co-parenting skill-building into its relationship education workshops. The evaluation found that participants reported improvements in healthy co-parenting skills after completing the program, particularly among those who attended virtually. The study also reported favorable findings in the relationship quality domain on improving partner conflict resolution, and in the mental health domain on improving personal well-being. The program had success particularly with virtual participants, who attended more sessions and completed the program at higher rates than those attending in person.

Economic stability domain. Of the four descriptive evaluations with favorable economic stability findings, three were focused on adults and one on youth. These four descriptive evaluations studied programs that delivered content in person and in hybrid formats (two programs each). These programs also devoted large proportions of their service offerings to financial literacy or career and employment coaching, either as stand-alone workshop sessions or integrated into each workshop session. For example, in one program, six of the 20 workshop hours were devoted to economic stability content, using the Money Smart curriculum and delivered as separate sessions. In a different program, career readiness and employment-focused content were integrated into the relationship education workshop sessions.

These descriptive evaluations reported favorable outcomes on measures of job readiness and employment (two favorable findings) and financial well-being and capability (five favorable findings). For example, Imperial Valley Regional Occupational Program's descriptive evaluation of youth in school reported improvements in healthy financial behaviors one year after youth enrolled in the program. All participants received classroom-based workshops that included a money management component, along with life coaching focused on goal setting and decision making. A subset of youth also participated in more intensive, optional activities such as career readiness courses, work-based learning, mentoring, and paid work experiences, which the evaluation noted may have reinforced practical money skills over time. The findings suggest that offering financial literacy integrated into relationship education paired with real-world opportunities to practice financial decision making can lead to improvements in financial behavior among youth.

Integrated financial education and relationship education paired with hands-on activities, such as paid job experiences or career readiness courses, improves the longer-term financial skills of youth.

Mental health domain. Five evaluations had favorable findings in the mental health domain, which were concentrated in outcome measures such as depression symptoms or psychological distress (six findings). One of these five evaluations, a study of Bethany Christian Services' (BCS) youth program, was also the only descriptive study to report favorable findings on self-regulation and mindfulness. BCS intentionally embedded social emotional learning throughout the Love Notes curriculum to improve youth's future decision making in relationships and real-world settings. The study found that youth showed improvements in social and emotional learning skills, including self-regulation, conflict management, and empathy.

In addition to analyzing participant outcomes, BCS's evaluation included focus groups with youth to provide a fuller picture of their program experiences. Participants described applying self-regulation and mindfulness-related skills in their relationships and paid work experiences, suggesting that consistently practicing these skills in their classrooms and at work can help youth incorporate these skills into everyday behavior.

What were common evaluation successes and challenges?

There are several common successes across the impact and descriptive evaluations with favorable findings. First, most of these evaluations analyzed large sample sizes—on average, about 775 for impact studies and 900 for descriptive studies. Second, with a few exceptions, attrition, or the amount of sample lost from the start to the end of the evaluation, was also low, averaging about 30 percent for impact evaluations and 35 percent for descriptive ones.[9] Third, evaluations that conducted an implementation study frequently noted two important aspects of program delivery—flexibility in the mode the content was offered in (such as a virtual option) and facilitators that delivered engaging and tailored content. These successes helped grant recipients recruit, enroll, and retain large numbers of participants in the local evaluations. These successes helped the evaluation have enough statistical power to find favorable outcomes that might be generalizable to similar populations of adults and youth.

Notably, nearly all impact evaluations with favorable findings reported program implementation strengths such as high dosage, fidelity, and participant satisfaction with services. Most participants received 80 to 95 percent of the intended content and said having child care options or flexibility in schedule or delivery of content kept them coming back.

Most descriptive evaluations showed at least one favorable finding, which reflects in part the way descriptive evaluations are executed. Descriptive evaluations had flexibility in the design of their studies. About half of descriptive studies with favorable findings focused on only those participants who completed services. Thus, it follows that many of the outcomes tested by these evaluations would be favorable because they reflect the outcomes of participants most engaged in services. In contrast, impact evaluations by design had to include all participants offered services, and they tested outcomes collected months past the end of the program to isolate longer-term program effects. Despite this difference, descriptive evaluations play a critical role by generating evidence about which program services are associated with participant outcomes, which paves the way for impact evaluations of similar services to isolate the effects, or for more descriptive studies of lesser-tested measures in this cohort such as parenting outcomes.

Evaluation challenges

The impact and descriptive evaluations that did not report favorable findings typically faced high rates of sample loss over the course of the evaluation. Many had smaller sample sizes to analyze than they originally planned and high levels of attrition across their tested outcomes—for example, more than 50 percent attrition for impact studies without favorable findings. These factors may have resulted in the evaluations lacking the statistical power to find a favorable outcome, despite the program providing thoughtful, tailored content to participants (Box 9).

Of the eight impact evaluations that did not show favorable findings, seven were enhancement tests, studying the impact of additional content or testing two similar curricula head-to-head.[10] Conducting an enhancement study can pose a risk for evaluations seeking to find an effect on outcomes, because these studies often have a low contrast between the study groups—that is, there is little distinguishing the services offered to the two (or more) groups. With low contrast, both groups often receive similar services, supports, or exposure to the program, so the evaluation is unlikely to be able to detect differences between groups. This problem can be worsened by high attrition, which further reduces the statistical power needed to detect an effect. The combination of a low contrast and high attrition characterized many

[9] The average attrition for descriptive evaluations does not include the evaluation that experienced very high attrition, in the range of 60 to 70 percent. Including this evaluation in the calculation raises average attrition to about 40 percent.

[10] The remaining one of these eight was a test of delivery mode (virtual, on-demand format versus in person) but had very high attrition and could not demonstrate baseline equivalence. Therefore, the study findings were descriptive.

Spotlight: Key findings for practice

Box 9. BCFS Health and Human Services learned how flexible services and added resources may improve program retention for youth transitioning out of foster care

BCFS Health and Human Services conducted a descriptive evaluation of connectED—Phase 1 (Love Notes curriculum), a healthy relationship program serving transition-age youth (ages 16–23) with experience in foster care. Although the evaluation did not find significant short-term changes in outcomes reflecting relationship knowledge or attitudes, it did yield valuable insights into how to effectively engage and support this population. Youth who participated said they were motivated by free programming, stipends to offset attendance costs, flexible scheduling, and opportunities for personal growth, peer connection, and learning practical relationship skills. The evaluation also highlighted common challenges to sustained participation, including housing and food instability, transportation and child care needs, mental health concerns, and competing work or school demands. The program evaluation pointed to the importance of trauma-informed approaches, trusted and consistent staff, flexible content delivery options, and incentives that help meet basic needs to support engagement and retention among youth transitioning out of foster care.

of the enhancement tests with no favorable findings during this cohort. The evaluations tested small, low-contrast additions such as short booster sessions, two 30-minute case management sessions, or virtual prompts to reinforce content, on top of offering the same evidence-based curricula to both groups. Other enhancement tests examined the differences between similar curricula tested head-to-head, where it was unclear how the content of the curricula differed between groups. However, the enhancement tests with favorable findings examined the addition of well-integrated, substantial content into HMRE programs with prior evidence of effectiveness and kept attrition low. Given these differences, even well-implemented interventions can appear to have small or no observable effects on outcomes if there is no clear contrast between the groups.

Future programs interested in designing enhancement tests should work closely with their evaluators to consider if the added service or component is substantial and sufficiently different enough from the program's core services to detect an impact. Similarly, evaluators can work closely with program staff to emphasize program retention (to keep attrition low) and ensure participants take up the added services to maintain a clear distinction between the study groups and the services they receive.

Considerations for HMRE programs from evaluation findings

- **Build enhancements on programs that already show promise.** Several evaluations tested enhancements by adding services, such as economic or employment supports, to relationship education models that had prior evidence of effectiveness. Programs that layered or integrated financial planning or career coaching onto established relationship education saw modest added benefits, suggesting enhancements may be most useful when they strengthen a solid core program.
- **Design for engagement first—delivery mode matters less than how services are delivered.** Programs that sustained participation across virtual and in-person formats emphasized skilled facilitation, interactive sessions, and flexibility in attending workshops. Programs used practices such as framing sessions as meaningful time together, using discussion and activities rather than lectures, and maintaining consistent staff contact to promote engagement. However, some programs reported higher attendance or completion in virtual formats, indicating that virtual delivery may be worth exploring to expand reach, reduce barriers, and potentially lower costs.

continued...

- **Emphasize skills that can be practiced and reinforced over time.** The most consistent findings across evaluations were improvements in applied skills. Youth programs showed longer-term gains in financial behaviors when participants had opportunities to practice money management through simulations, coaching, or work-based experiences. Similarly, mindfulness, self-regulation, and emotional management were skills practiced across classroom activities, coaching, and real-world settings and showed improvement for youth and adults.
- **Adapt programs to participant needs and context.** Evaluations showed that programs were more engaging when content and delivery reflected participants' experiences. Examples are tailored relationship education for specific communities, or co-parenting supports designed for stepfamilies. These adaptations did not require entirely new models but rather thoughtful adjustments to curriculum, facilitation, and supports made in collaboration with curriculum developers and ACF partners.
- **Use evaluation findings to find the areas where more learning is needed.** Although positive findings on relationship skills outcomes were common, fewer evaluations found strong or consistent evidence related to parenting outcomes or longer-term economic stability changes. These gaps point to opportunities for programs to refine how they address parenting and financial skills, strengthen their measurement, or test new approaches in areas where evidence is still emerging.

Conclusion

Across the 2020 cohort of HMRE local evaluations, the strongest and most consistent evidence was found in the domain of relationship quality, in which programs frequently measured outcomes such as relationship satisfaction, communication, and conflict and often observed favorable results. Fewer evaluations focused on parenting outcomes, mental health, and economic stability, and the evidence in these domains was consequently more limited. Taken together, these findings suggest that HMRE programs are well positioned to strengthen relationship skills, and they highlight opportunities to fully examine how relationship education connects to longer-term parenting, mental health, and economic stability outcomes.

The evaluation findings also point to practical steps programs can take to strengthen both HMRE services and future learning. Programs that offered flexible delivery options, engaging facilitation, and supports such as child care or scheduling accommodations reported success at keeping participants engaged. These factors point to important considerations program and evaluation teams should take into account to improve both service delivery and evaluation quality. To build stronger evidence, future evaluations may benefit from strategies that keep participants engaged and focus on data collection methods to reduce attrition in the evaluation sample. In addition, collaboration between the program and evaluator is essential to ensure sufficient program recruitment and enrollment of participants, especially for designs like equivalent effects tests that require larger sample sizes. It is particularly important for programs that want to evaluate enhancements to their services to work closely with the evaluation team to ensure the new content is offered in a substantial dose and well integrated with core programming to maximize the chances of detecting impacts. Together, these lessons highlight how thoughtful program design paired with intentional evaluation can support continuous improvement for programs and stronger outcomes for families.

Suggested citation: Angela Valdovinos D'Angelo, Scott B. Richman, and Avery Hennigar (2026). "Local Evaluation Findings from the 2020 Cohort of Healthy Marriage and Relationship Education Grant Recipients." OPRE-2026-018. Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services.

Acknowledgements: Many people contributed to this brief. First, we acknowledge staff in the Administration for Children and Families at the U.S. Department of Health and Human Services. We are particularly grateful for the direction and feedback from Kathleen McCoy and Rebecca Hjelm. We extend our gratitude to Quinn Moore for reviewing drafts of the brief, and Hannah McInerney and Anna Marston for their assistance with analysis of the grant recipient final reports. We also extend our appreciation to Effie Metropoulos for editing and Yonne Marki-Korosec for designing the graphics in this brief.

*This product was finalized and published after March 2026, reflecting ACF's updated approach in which some research portfolios were realigned to be embedded in the ACF offices that manage the programs. Consistent with this approach, this product was reviewed and approved by leadership in the ACF Office of Family Assistance.

The views expressed in this publication do not necessarily reflect the views or policies of the Administration for Children and Families or the U.S. Department of Health and Human Services.

Appendix A.

Table A.1. Characteristics of HMRE programs with evaluations, by focal population

	Adult	Youth
Workshop dosage		
12 to 14 hours	50%	63%
15 to 19 hours	18%	25%
20 hours or more	32%	13%
Service delivery format		
In person	18%	56%
Virtual	14%	0%
Hybrid	68%	44%
Curriculum		
Most commonly used[a]	PREP: 27% Within My Reach: 18% Money Smart: 14%	Love Notes: 53% Mind Matters: 40% Money Habitudes: 33%
Single curriculum	73%	44%
Multiple curricula	27%	56%
Program setting[b]		
Community setting	82%	44%
Schools	0%	81%
Virtual	18%	13%
Other	14%	13%
Region		
Midwest	9%	19%
Northeast	13%	6%
South	65%	56%
West	13%	19%

Note: Percentages are calculated within the column for each variable and sum to 100 percent except when noted otherwise. One grant recipient, Anthem, was excluded from the calculations for workshop dosage, service delivery format, curriculum, and program setting, because the primary focus of the local evaluation was to implement and test a continuous quality improvement process to support the delivery of its program.

[a] HMRE-evaluated programs could consist of one or more curricula.

[b] Percentages do not sum to 100 percent because HMRE-evaluated programs could have been implemented in multiple settings.