



RESEARCH ARTICLE

Implementation and Policy Barriers to Evaluating Innovative Sexuality Education Programs for Early Adolescents: Implications for Evidence Standards

Kelly L. Wilson, PhD, MCHES¹, Whitney R. Garney, PhD, MPH, FAHA², Brittany L. Rosen, PhD, Med³, Christi H. Esquivel, PhD, CHES⁴, Caitlin B. Holden, PhD, CHES⁴, Sydney Hineline, MPH⁵

¹Professor, Texas A&M University, College of Nursing

²Associate Professor, Texas A&M University, School of Public Health; Director of Center for Community Health and Aging

³Associate Professor of Pediatrics, Cincinnati Children's Hospital Medical Center, Division of Adolescent and Young Adult Medicine; University of Cincinnati College of Medicine; Director of Evaluation for the Center for Clinical and Translational Science and Training; Director of Dissemination and Implementation for the Center for Clinical and Translational Science and Training

⁴Assistant Research Scientist, Texas A&M University, School of Public Health

⁵Clinical Research Coordinator IV, Cincinnati Children's Hospital Medical Center, Division of Adolescent and Young Adult Medicine



OPEN ACCESS

PUBLISHED

28 February 2026

CITATION

Wilson, K.L., et al., 2026. Implementation and Policy Barriers to Evaluating Innovative Sexuality Education Programs for Early Adolescents: Implications for Evidence Standards. *Medical Research Archives*, [online] 14(2).

COPYRIGHT

© 2026 European Society of Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ISSN

2375-1924

ABSTRACT

Background: Early adolescence is a critical period for sexual and reproductive health education, yet few rigorously evaluated sexuality education programs are designed for middle school youth. Innovative approaches such as game based, technology enabled, and youth co developed interventions offer developmentally appropriate strategies for engaging early adolescents. However, these programs face substantial barriers to evaluation that are shaped by policy environments, ethical requirements, and systems level evidence standards developed primarily for older adolescents and young adults.

Objective: This manuscript examines implementation and policy related barriers that constrain the evaluation of innovative sexuality education programs for middle school youth. It specifically considers how current evidence standards, including those used by the Teen Pregnancy Prevention Evidence Review, limit the assessment, progression, and scalability of early-stage innovations.

Methods: This manuscript draws on implementation and evaluation insights generated through the development of a developmentally appropriate survey to assess a middle school sexual health program. The perspectives presented are informed by systematic survey development activities grounded in prior adolescent pregnancy prevention evaluations, psychometric testing, and youth centered refinement processes. These experiences are used to illustrate broader structural and policy related challenges associated with measuring outcomes for early adolescent sexuality education programs within existing federal evidence frameworks.

Results: Four interrelated evaluation barriers were identified. These included insufficient funding for developmentally appropriate evaluation, limited readiness of early-stage innovations for experimental designs prioritized by federal evidence reviews, gaps in evaluation expertise among innovation focused teams, and misalignment between middle school program goals and outcome requirements emphasizing sexual behavior change. These barriers were amplified by ethical and legal constraints associated with research involving minors and by state and local policy environments governing sexuality education.

Conclusions: Current evaluation frameworks inadequately support innovative sexuality education programs for early adolescents. Advancing the evidence base requires expanded evaluation approaches that recognize developmental appropriateness, upstream outcomes, and implementation context. Aligning evidence standards with the realities of early adolescent development is essential to sup and reposition evidence gaps as products of system design rather than program failure.

Keywords: early adolescence, middle school youth, sexuality education, sexual and reproductive health, program evaluation, implementation science, teen pregnancy prevention

Introduction

Early adolescence represents a foundational developmental period for sexual and reproductive health (SRH), yet early adolescents remain underrepresented in the evidence base for sexuality education programs.¹⁻³ National policy emphasizes preventing teen pregnancy and promoting healthy adolescent development, but most evidence-based interventions target older adolescents and rely on behavioral outcomes such as sexual initiation, contraceptive use, or pregnancy, that are rare or developmentally inappropriate for younger youth.⁴⁻⁶ Consequently, few programs designed for middle school-aged youth meet federal evidence standards or appear in registries that guide funding and program adoption.⁷⁻¹¹ Evidence from systematic reviews and meta-analyses indicates that school-based teen pregnancy prevention programs show modest effectiveness overall, but data for middle school populations remain limited.^{12,13}

Innovative approaches have emerged to address the limitations of traditional, curriculum-based interventions.¹⁴⁻¹⁷ For example, middle school programs demonstrate that fidelity and implementation in real-world settings are critical for success.¹⁸ These include game-based learning, digital platforms, and youth co-designed programs that emphasize engagement, skill building, and protective factors rather than sexual behavior change alone.¹⁴⁻¹⁷ Such approaches align with early adolescents' cognitive and social development and support public health goals related to health literacy, communication, and decision making.^{15,16} Despite their promise, these programs face significant challenges when evaluated under prevailing evidence standards.¹⁷

From an implementation science perspective, evaluation is shaped not only by program design but also by contextual factors, policy environments, and system level expectations.^{19,20} Federal evidence frameworks, including the US federal evidence standards called Teen Pregnancy Prevention Evidence Review (TPPER), prioritize randomized controlled trials and behavioral outcomes.⁷⁻¹¹ Systems have been

developed to synthesize and translate evidence for teen pregnancy, HIV, and STI prevention, yet these structures often favor older adolescents and established interventions.²¹ While appropriate for older adolescents, interventions, they may be poorly suited to early stage innovations targeting middle school youth, where sexual behaviors are infrequent and ethical considerations limit data collection.^{5,11}

Research involving minors introduces additional barriers.²² Federal regulations require youth assent, and may require parental consent depending on the research being conducted, increasing recruitment burden and contributing to potential selection bias.^{22,23} In sexuality education, parental concerns, community norms, and varying state level policies further constrain program delivery and evaluation design.^{8,23} These challenges are amplified in under-resourced communities, where staffing, scheduling, and technology limitations complicate implementation and data collection.^{23,13}

Collectively, these factors create a misalignment between the developmental goals of innovative middle school sexuality education programs and the systems used to evaluate and validate them.^{17,19} Understanding these barriers is critical for advancing evidence-building approaches that support SRH during early adolescence. Without such understanding, middle school sexuality education remains systematically disadvantaged within evidence-building systems designed for older populations.

The purpose of this manuscript is to identify and analyze structural barriers to evaluating sexuality education programs for middle school youth, with particular attention to how federal evidence standards, policy environments, and implementation contexts constrain developmentally appropriate evaluation. Drawing on implementation science and a survey development case example, the paper advances recommendations for modernizing evidence-building frameworks to better support early adolescent sexual and reproductive health.

By focusing on system constraints rather than presenting new trial data, this manuscript illustrates how funding structures, innovation readiness, evaluation capacity, and evidence standards interact to limit the assessment and scalability of early-stage programs designed for middle school youth, and highlights structural considerations that shape program development, evaluation and policy adoption. Figure 1 illustrates this structural misalignment, depicting how early adolescent

developmental considerations, program design and innovation, implementation context, evaluation capacity, and federal evidence standards interact to constrain developmentally appropriate evaluation and limit scalability of middle school sexuality education programs. Rather than representing a linear evaluation pathway, the model emphasizes feedback loops through which policy constraints, evidence standards and resource burdens reinforce systemic barriers.

Figure 1. Conceptual Model of Structural Misalignment in Evaluating Middle School Sexuality Education

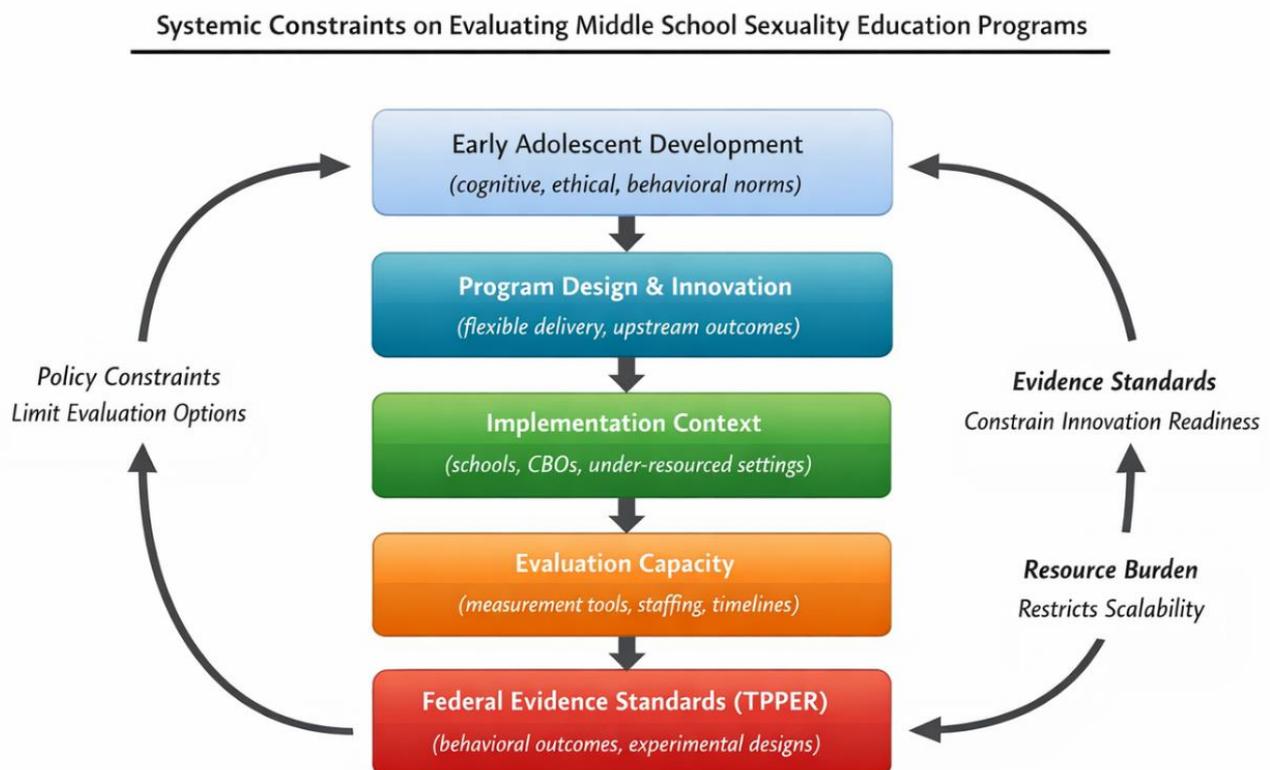


Figure 1. Conceptual model illustrating structural misalignment between early adolescent development, program design, implementation context, and federal evidence standards in the evaluation of middle school sexuality education programs. The model demonstrates how policy environments, evaluation capacity, and evidence review criteria interact to constrain developmentally appropriate measurement, limit innovation readiness, and impede progression along evidence-building pathways.

Perspectives/Case Example: Survey Development as a Case Example Demonstrating Systemic Evaluation Barriers

As an illustrative case within this perspective, evaluation of sexuality education programs for this population is fundamentally constrained by a

mismatch between developmental appropriateness and prevailing evidence standards.^{5,11,12} The development of a sexual health evaluation survey for middle school youth provides a case example of how systemic barriers in policy, funding, and evidence frameworks shape evaluation practices, rather than reflecting deficiencies in program design. Our case study centers on a prospective, two-arm, multi-site

clustered randomized controlled trial using survey data collection for the evaluation. Aligned with developmental appropriateness, we compiled a survey using established scales to measure youths' intention to delay sex, intention to use condoms and contraceptives during sex, communication skills, decision-making skills, skills to access sexual health information and resources, and perception of connection to trusted adults.

Developing a survey for middle school participants required a deliberate departure from commonly used pregnancy prevention outcome measures. Items assessing sexual activity, condom use, and contraceptive use are central to evidence standards such as the Teen Pregnancy Prevention Evidence Review. They were deemed developmentally inappropriate, ethically sensitive, and conceptually inaccessible for early adolescents by our school and community based implementation partners and reinforced in the literature.^{5,11-13} Even when drawn from validated instruments, terminology, assumptions, and response formats exceeded early adolescents' cognitive and experiential readiness.^{12,13} Consequently, survey sections focused on sexual behavior were removed as part of an iterative development process aligned with the evaluation design that included cognitive interviews, content expert review, youth stakeholder review, and pilot administration. Sections were removed not because they lacked theoretical relevance, but to preserve data validity and participant comprehension.¹²

This case exemplifies a system-level barrier: federal evidence frameworks prioritize behavioral outcomes that are neither expected nor ethically appropriate for middle school youth. Programs must rely on upstream indicators such as intentions, communication skills, decision-making, access to information, and perceptions of adult connectedness.²⁴ While these constructs are supported by developmental and prevention science, they are often treated as secondary outcomes within federal evidence review criteria, limiting programs' ability to demonstrate recognized effectiveness.¹⁹ Accordingly, the selected

outcomes were pre-specified as appropriate given the developmental capacities and meaningful outcomes for middle school youth as well as consent limitations related to measures that could be ethically and feasibly collected.

Implementation context further shapes feasible evaluation approaches. By conducting cognitive interview with youth aged 11-14 years old to assess their comprehension and readability of the survey, we identified even seemingly neutral survey elements, such as Likert scale response options or generic references to "someone," introduced comprehension challenges. Addressing these issues required substantial simplification of language and reduction to a fourth grade reading level, reinforcing the need for developmentally responsive measurement strategies. These adaptations are rarely accounted for in guidance tied to federal funding, which assumes the transferability of instruments validated with older adolescents.

From an implementation science perspective, these challenges reflect limited innovation readiness within evaluation systems rather than within programs themselves. Early stage and sexuality education innovations are often designed for flexible delivery across schools, community-based organizations, or after school settings, yet evaluation expectations favor tightly controlled designs and standardized measures. This tension constrains learning during early phases of implementation, when formative and developmental evaluation approaches would be more appropriate.

Furthermore, policy environments governing sexuality education (i.e., parental consent, community norms, and state level restrictions) directly influence permissible evaluation measures. These restrictions pose even greater challenges in multi-site evaluations conducted across state lines, where continually shifting state legislation and institutional policies can be difficult to monitor and integrate throughout a multi-year evaluation. These factors compound barriers to data collection and reinforce reliance on

selecting conservative outcomes (such as, knowledge, attitudes, and intentions), further distancing middle school programs from evidence thresholds required for broader dissemination. Our evaluation design allowed each site to adapt implementation and data-collection procedures, within approved parameters, to meet local policy context, while using conservative, low-burden measures that could be used in all settings.

Table 1 summarized systematic barriers to evaluating middle school sexuality education programs and recommended adaptations aligned with a modernized TPPER framework. This case example demonstrates how evaluation barriers are embedded

within systemic structures of funding, policy, and evidence review. Addressing these barriers require expanded evidence standards that recognize developmentally appropriate outcomes, value implementation and formative learning, and support innovation pathways tailored to early adolescents. Without such reforms, promising sexuality education programs will continue to face structural obstacles to real-world evaluation, scaling, and integration in evidence-based program portfolios. Through our case study, we demonstrate that our evaluation design reflects the systemic constraints influencing which designs are available for middle school sexuality education.

Table 1. Systemic Barriers to Evaluating Middle School Sexuality Education and Implications for Teen Pregnancy Prevention Evidence Review Modernization

Systemic Barrier	Recommended Adaptation	Alignment with TPPER Modernized
Developmentally inappropriate behavioral outcomes	Use upstream protective factors (skills, literacy, connectedness)	Intermediate outcomes recognized
Limited RCT feasibility	Formative and implementation-focused designs	Broader design inclusion
Policy and consent constraints	Context-aware outcome interpretation	Equity-oriented evidence review

a. TPPER = Teen Pregnancy Evidence Review.

b. Adapted from prior research and implementation science perspectives.^{2,5,19,24,28}

Discussion: Key Implementation and Evaluation Barriers to Assessing Sexuality Education Programs for Middle School Youth

The following discussion synthesizes these observations to identify key implementation and evaluation barriers that operate across program contexts, rather than issues unique to any single intervention.

BARRIER 1: ETHICAL, LEGAL, AND RESOURCE CONSTRAINTS SHAPE FEASIBILITY AND ADOPTION
 Evaluation of sexuality education programs for middle school youth is fundamentally shaped by ethical and legal requirements governing research with minors, which directly affect feasibility, adoption, and sustainability of evaluation efforts.^{22,23} Federal regulations require parental or guardian consents and youth assent for participants under 18, protections that are ethically essential but administratively and operationally demanding. Institutional review board

processes, consent documentation, and compliance monitoring extend evaluation timelines and require specialized expertise that many intervention and design teams lack.

These ethical and legal requirements substantially increase resource demands. Evaluations frequently require intensive recruitment strategies, repeated parental outreach, and ongoing follow-up to maintain adequate participation and retention. Additionally, effective follow-up strategies for middle school youth outside of school settings require proactively collecting multiple forms of contact information, since youth at this age may not reliably respond to texts, emails or physical mail.

Despite these investments, evaluations of middle school programs are unlikely to yield meaningful data on downstream behavioral outcomes such as sexual initiation, contraceptive use or pregnancy.⁴⁻⁶ Innovative approaches in rural contexts emphasize leveraging community resiliency and local resources to support prevention efforts, which may not be captured by traditional metrics.²⁵ This creates a structural disincentive for funders and limits the resources available to assess developmentally appropriate outcomes, including knowledge, healthy relationship attitudes, communication skills, self-efficacy, and decision-making skills. As a result, under-resourced projects often lack the capacity to measure and assess meaningful changes in developmentally appropriate protective factors, including knowledge, communication skills, decision-making skills, self-efficacy, and healthy relationship attitudes.

In under-resourced projects these challenges are compounded by economic instability, caregiver time constraints, and competing family priorities, such as work obligations, childcare responsibilities, household management and other pressing demands, further impeding recruitment and retention. From an implementation science perspective, these conditions reflect misalignment between contextual realities and evaluation expectations, where ethical protections and real-world constraints undermine

the feasibility of producing the type of evidence required for broader adoption and scale.

BARRIER 2: LIMITED FEASIBILITY OF EXPERIMENTAL DESIGNS UNDERMINES FIDELITY AND INTERNAL VALIDITY

From an implementation science perspective these challenges reflect a misalignment between stage of innovation and evaluation expectations. As illustrated in Figure 2, evidence-building for middle school sexuality education is inherently iterative and developmental, yet current federal evidence standards disproportionately emphasize later-stage effectiveness testing. The real-world settings in which middle school sexuality education programs are delivered (i.e., schools, community-based organizations, and after-school programs), present significant barriers to rigorous experimental testing.^{5,11} Challenges obtaining parental consent for sensitive topics such as sexual and reproductive health are common, leading to reduced participation rates and increased risk of selection bias, which threatens internal validity and limits generalizability.

Innovative programs for early adolescents are often intentionally designed for flexibility and contextual adaptation, supporting reach and acceptability across diverse settings. However, this flexibility complicates implementation consistency and limits the ability to maintain fidelity across sites, particularly in under-resourced environments characterized by staffing shortages, competing curricular demands, limited physical space, and scheduling liability. These conditions make randomized controlled trials unrealistic and difficult to implement and sustain, despite their central role in federal evidence frameworks.^{5,6}

Figure 2. Evaluation and Evidence-Building Lifecycle for Middle School Sexuality Education Programs

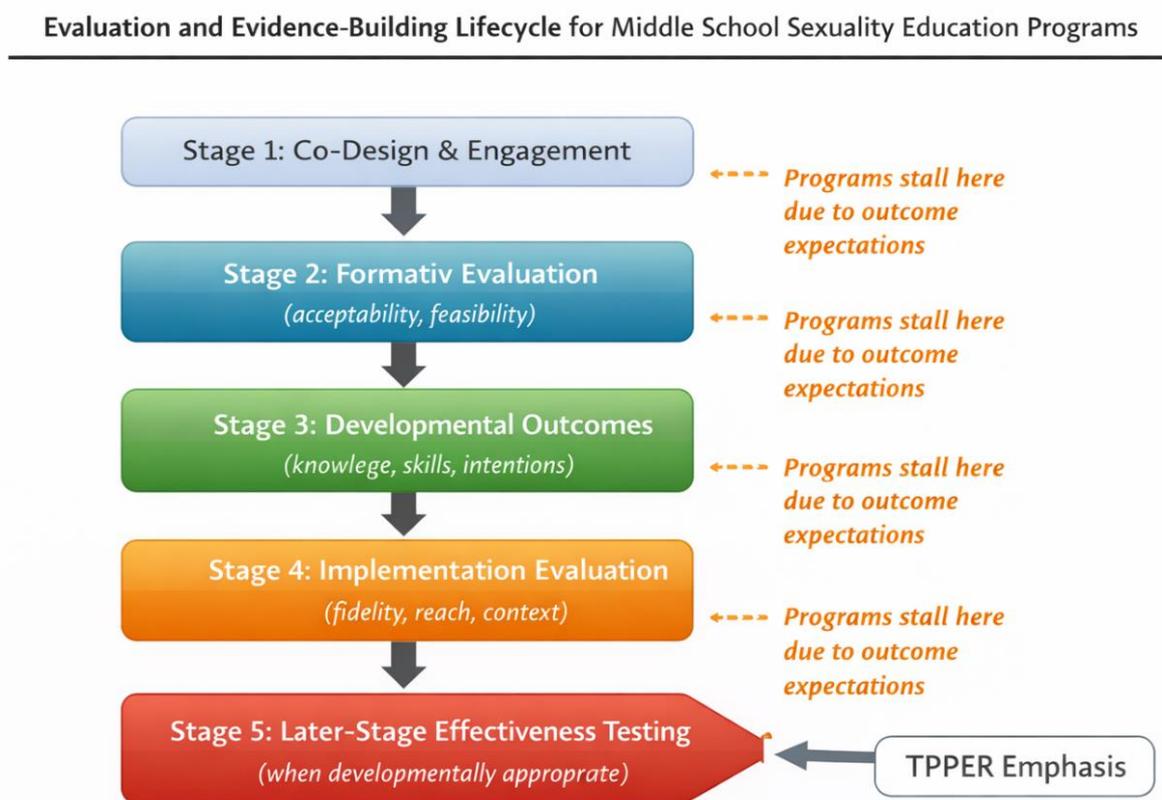


Figure 2. Developmentally appropriate evidence-building lifecycle for middle school sexuality education programs. Early stages emphasize formative, developmental, and implementation-focused evaluation aligned with early adolescence, while later-stage experimental testing becomes appropriate as programs mature. Current federal evidence standards disproportionately prioritize later-stage outcomes, limiting progression for early-stage innovations.

Moreover, many early-stage innovations are not yet sufficiently mature for summative experimental testing. When program goals emphasize upstream protective factors rather than immediate behavioral change, traditional experimental designs may undervalue meaningful program effects. From an implementation science standpoint, this reflects a mismatch between stage of innovation and evaluation approach, where premature demands for high-level evidence restricts learning, delay refinement and impede progression along evidence-building pathways.¹⁹ These challenges are situated within a developmentally appropriate evidence-building lifecycle for middle school sexuality education programs, highlighting how early-stage innovations frequently stall when later-stage outcome expectations are applied prematurely.

BARRIER 3: MISALIGNMENT BETWEEN EVIDENCE STANDARDS, EVALUATION CAPACITY, AND DEVELOPMENTAL APPROPRIATENESS LIMITS SCALABILITY

A third barrier arises from misalignment between federal evidence standards, the evaluation capacity of innovation teams, and the developmental realities of early adolescence, directly affecting scalability and system readiness.⁷⁻¹¹ Teams developing sexuality education programs for middle school youth often emphasize youth engagement, co-design, cultural relevance, and trauma-informed approaches. While these strategies enhance acceptability and fit, they are not always accompanied by access to evaluation expertise in developmental, formative, or implementation-focused methodologies.

Federal evidence-based standards, including those used by the Teen Pregnancy Prevention Evidence Review, prioritize behavioral outcomes and specific experimental designs that are often inappropriate or negligible for early adolescents.^{4-6,11} At the same time, state and local policies governing sexuality education may restrict content or prohibit collection of sexual behavior data altogether, further narrowing permissible outcomes. Consequently, evaluations must often rely on proxy measures, such as intentions, attitudes, communication skills or perceived support, that are well supported by developmental and prevention sciences but are frequently considered secondary or insufficient for meeting federal evidence review criteria.

Data collection challenges further constrain evaluation capacity. While electronic and virtual data collection methods offer efficiency, disparities in access to devices, internet connectivity and digital literacy, particularly in under-resourced communities, introduce threats to data completeness, quality, and consistency in follow up participation. Collectively, these constraints highlight limitations in system-level innovation readiness,^{12,14,26} including the challenges of implementing multi-level ecological approaches in diverse school and community settings where evidence frameworks have not evolved to accommodate ethical protections, developmental appropriateness, and implementation context.²⁷

Recommendations and Implications for the Future of Sexuality Education Programs

Consistent with the aims of this manuscript, these recommendations focus on system-level reforms rather than program-level modifications. Taken together, the barriers identified in this manuscript reflect a dynamic interaction between early adolescent developmental stages, real-world implementation contexts, and policy and evidence environments. Limitations in the evidence base for middle school sexuality education are not indicative of weak or ineffective programs, but rather of structural

misalignment between how programs are designed to work and how evidence is generated and assessed, as documented in the US syntheses of teen pregnancy prevention initiatives (Figures 1 and 2).²⁵

Together, Figures 1 and 2 underscore that limitations in the evidence base for middle school sexuality education stem from systemic constraints and misaligned evidence pathways, rather than from deficiencies in program theory or implementation. Addressing these challenges requires coordinated shifts across policy, funding, and evaluation frameworks to ensure that innovation in early adolescent sexual and reproductive health is both assessable and scalable.^{13-15,28}

First, evidence standards should more explicitly recognize developmentally appropriate outcomes as valid indicators of program effectiveness for early adolescents. For middle school youth, upstream and protective factors, such as health knowledge, communication and refusal skills, decision-making capacity, self-efficacy, health literacy, and perceptions of trusted adult support, are theoretically grounded and developmentally salient. These outcomes align with prevention science and are increasingly acknowledged within the Teen Pregnancy Prevention Evidence Review (TPPER) as intermediate indicators of impact. Establishing clearer age-specific pathways within evidence reviews would allow programs targeting early adolescents to demonstrate meaningful contributions to public health goals without reliance on behavioral outcomes that are ethically or developmentally inappropriate to measure.

Second, funding mechanisms must better align with the ethical and logistical realities of conducting research with minors. Evaluation of middle school sexuality education programs requires greater investment in parental engagement, consent processes, recruitment and retention, activities that directly affect feasibility and data quality.²⁸ Funders should support formative, developmental, and implementation-focused evaluations during early stages of innovation, recognizing that such approaches

generate critical evidence about feasibility, acceptability, and mechanisms of change before large-scale effectiveness testing is appropriate.

Third, evidence-building systems should more fully integrate implementation science outcomes into evaluation expectations. Programs for early adolescents are often designed for more flexible delivery across schools, community-based organizations, and informal learning environments. Evaluation designs that assess feasibility, fidelity, acceptability, reach and implementation quality,²⁸ alongside developmentally appropriate outcomes, provide essential insight into how programs function within varied contexts. Under TPPER, these data can meaningfully contribute to evidence tiers when they are explicitly recognized as part of the innovation lifecycle rather than treated as insufficient substitutes for behavioral outcomes.

Fourth, greater attention to evaluation capacity-building is needed. Innovation-focused program teams may lack access to expertise in developmentally appropriate measurement or alternative study designs. Technical assistance models that pair program developers with evaluation specialists experienced in early adolescent research, and that emphasize mixed-methods and ethical data collection, can strengthen alignment between innovation and evidence expectations.

Finally, policy environments governing sexuality education should be treated as integral contextual determinants of both implementation and evaluation feasibility. State and local restrictions on program content, parental consent requirements, and limits on data collection shape what outcomes can be measured and how studies are conducted. Evidence reviews and funding decisions should explicitly account for these constraints to avoid penalizing programs for factors beyond their control. Greater coordination between policy development and evidence standards would support more equitable pathways to dissemination. Together these recommendations offer a framework for aligning

evaluation rigor with developmental science, rather than treating the two as competing priorities.

Implications for Teen Pregnancy Prevention Evidence Review Modernization and Evidence Building

The findings presented in this manuscript underscore a central conclusion: a modernized Teen Pregnancy Prevention Evidence Review (TPPER) is necessary to support effective evidence-building for sexuality education programs designed for early adolescents. This manuscript is intended to inform ongoing modernization efforts by clarifying how early adolescent programs interact with existing evidence criteria in practice. While recent modernization efforts represent important progress, through expanded evidence tiers, increased transparency and acknowledgement of diverse study designs, persistent misalignment remains between TPPER criteria and the developmental, contextual, and ethical realities of middle school sexuality education. Ethical considerations are integral to implementation research and must be embedded throughout evidence-building processes, particularly when working with minors and marginalized populations.²⁹⁻³⁰

Current TPPER frameworks continue to prioritize behavioral outcomes related to sexual activity and pregnancy prevention as primary indicators of effectiveness. Although appropriate for programs serving older adolescents, this emphasis is poorly aligned with early adolescence, a developmental stage in which such behaviors are uncommon and often inappropriate or infeasible to measure. For middle school youth, developmentally appropriate outcomes, including health literacy, communication and decision-making skills, self-efficacy, boundary-setting and trusted adult connectedness, represent theoretically grounded and useful indicators of program impact. A modernized TPPER must explicitly recognize the outcomes as valid and sufficient indicators of effectiveness for early adolescent populations.

In addition, the evidence pathways within TPPER must better reflect the innovation lifecycle. Early-stage sexuality education programs targeting middle school youth frequently generate rigorous evidence related to feasibility, acceptability, fidelity and intermediate outcomes, yet remain stalled within evidence tiers because behavioral outcomes cannot be measured. This manuscript argues there is a need for TPPER to formally incorporate formative, developmental and implementation-focused evaluation as meaningful contributors to evidence ratings, particularly for programs at early stages of development. Without such pathways, promising innovations remain structurally excluded from progression toward scaling.

Clear, age-specific evidence expectations are also essential. Applying uniform standards across heterogeneous developmental stages disadvantages middle school programs and obscures meaningful evidence of impact. A modernized TPPER should establish differentiated criteria for early adolescents versus older youth, providing clearer guidance to program developers, evaluators, and funders while maintaining scientific rigor appropriate to developmental context.

Finally, this manuscript highlights the necessity of accounting for policy and contextual constraints within evidence review processes. State and local restrictions on sexuality education content, parental consent requirements, and limitations on data collection directly shape evaluation feasibility and outcome selection. A modernized TPPER must explicitly consider these constraints when interpreting evidence to avoid penalizing programs for external factors beyond their control and to prevent the systematic exclusion of programs serving under-resourced or policy-restricted settings.

Taken together, these findings indicate that continued modernization of TPPER is not simply beneficial, but essential for advancing developmentally appropriate sexuality education for all youth. By aligning evidence standards with early adolescent development, real-

world implementation contexts, and ethical research practices, a modernized TPPER can strengthen the evidence base, support responsible innovation and advance public health goals by investing in prevention before risk behaviors emerge.

Conclusion

Sexuality education for early adolescents is a public health priority that remains insufficiently supported by existing evaluation and evidence-building systems. As this manuscript argues, the primary challenges facing innovative middle school sexual and reproductive health programs stem from structural misalignment, not from weak theory or poor program design, but from evidence standards and policy environments that have not kept pace with developmental science or implementation realities. This manuscript does not evaluate the effectiveness of a specific intervention; rather, it clarifies why existing systems struggle to capture effectiveness for middle school sexuality education programs at all.

Aligning evaluation frameworks with early adolescent development, real world contexts, and ethical constraints is essential to advancing equity in sexual and reproductive health. Expanding evidence pathways to include developmentally appropriate outcomes and implementation-focused learning would strengthen, not dilute, scientific rigor by ensuring that evidence reflects how and for whom programs work. Without such alignment, the field risks continuing to value innovation in principle while constraining it in practice, delaying opportunities to support young people during a critical formative stage of development. By reframing evaluation challenges as structural rather than pragmatic, this manuscript provides a foundation for more equitable and developmentally responsive evidence building in early adolescent sexual and reproductive health.

References:

1. Esquivel CH, Flores SA, Garcia K, Garney WR, Wilson K. Partnering with youth to enhance healthcare access and experience: lessons learned from a teen advisory group. *Health Expect.* 2025;28(2):e70220. doi:10.1111/hex.70220
2. Wilson KL, Axelson S, Garney WR, et al. Addressing evaluation barriers with early innovation development for adolescent-focused sexual and reproductive health interventions. *Prev Sci.* 2023; 24(Suppl 2):222-228. doi:10.1007/s11121-023-01578-2
3. Wilson K, Axelson S, Garney WR, et al. Healthcare autonomy among adolescents and young adults in the United States: a scoping review. *J Adolesc Health.* 2023;72(3 Suppl):S70-S71.
4. Garcia KM, Esquivel CH, Garney WR, et al. A program evaluation of the Innovative Teen Pregnancy Prevention Programs (iTP3) Project. *Arch Public Health.* 2022;80:24. doi:10.1186/s13690-021-00723-z
5. Garney WR, Wilson KL, Garcia KM, et al. Supporting and enabling the process of innovation in public health: the framework for public health innovation. *Int J Environ Res Public Health.* 2022; 19(16):10099. doi:10.3390/ijerph191610099
6. Rosen B, Hineline S, Wilson K, Holden C, Garney W, Meisman A. Survey development to evaluate a sexual health program for middle school youth. *J Adolesc Health.* 2023;76(3 Suppl):S105-S106.
7. Hineline SK, Wilson K, Garney WR, et al. Collecting evaluation data remotely for middle school-aged children: lessons learned. *J Adolesc Health.* 2022;74(3 Suppl):S97.
8. Esquivel CH, Wilson KL, Garney WR, Exum Smith CE, McNeill EB, McMaughan DJ. A case study evaluating youth acceptability of Using The Connect—a sexuality education game-based learning program. *Am J Sex Educ.* 2022;17(1):57-83. doi:10.1080/15546128.2021.1971128
9. Esquivel CH, Wilson KL, Garney WR, Smith CE, McNeill EB, McMaughan DJ. Using human-centered design to develop an innovative teen pregnancy prevention program: lessons learned from a case study. *Creat Educ.* 2022;13(4):1439-1457. doi:10.4236/ce.2022.134088
10. Finer LB, Philbin JM. Sexual initiation, contraceptive use, and pregnancy among young adolescents. *Pediatrics.* 2013;131(5):886-891. doi:10.1542/peds.2012-3495
11. Grossman JM, Frye A, Charmaraman L, Erkut S. Family homework and school-based sex education: delaying early adolescents' sexual behavior. *J Sch Health.* 2013;83(11):810-817.
12. Lindberg LD, Maddow-Zimet I, Marcell AV. Prevalence of sexual initiation before age 13 years among male adolescents and young adults in the United States. *JAMA Pediatr.* 2019;173(6):553. doi:10.1001/jamapediatrics.2019.0458
13. Lindberg LD, Firestein L, Beavin C. Trends in U.S. adolescent sexual behavior and contraceptive use, 2006–2019. *Contracept X.* 2021;3:100064. doi:10.1016/j.conx.2021.100064
14. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health.* 2011;38(2):65-76. doi:10.1007/s10488-010-0319-7
15. Koh H. The teen pregnancy prevention program: an evidence-based public health program model. *J Adolesc Health.* 2014;54(3 Suppl):S1-S2.
16. Brindis CD. Advancing the field of teenage pregnancy prevention through community-wide pregnancy prevention initiatives. *J Adolesc Health.* 2017;60(3 Suppl):S1-S2.
17. Brindis CD, Decker MJ, Gutmann-Gonzalez A, Berglas NF. Perspectives on adolescent pregnancy prevention strategies in the United States: looking back, looking forward. *Adolesc Health Med Ther.* 2020;11:135-145.
18. Juras R, Tanner-Smith E, Kelsey M, Lipsey M, Layzer J. Adolescent pregnancy prevention: meta-analysis of federally funded program evaluations. *Am J Public Health.* 2019;109(4):e1-e8. doi:10.2105/AJPH.2018.304925

19. Marseille E, Mirzazadeh A, Biggs MA, et al. Effectiveness of school-based teen pregnancy prevention programs in the USA: a systematic review and meta-analysis. *Prev Sci*. 2018;19(4):468-489.
20. Rohrbach LA, Donatello RA, Moulton BD, Afifi AA, Meyer KI, De Rosa CJ. Effectiveness evaluation of *It's Your Game: Keep It Real*, a middle school HIV/sexually transmitted infection/pregnancy prevention program. *J Adolesc Health*. 2019;64(3):382-389.
21. Vasilenko SA, Odejimi OA, Glassman JR, et al. Who benefits from school-based teen pregnancy prevention programs? Examining multidimensional moderators of program effectiveness across four studies. *Prev Sci*. 2023;24(8):1535-1546.
22. Lewis KM, Lesesne CA, Zahniser SC, et al. Developing a prevention synthesis and translation system to promote science-based approaches to teen pregnancy, HIV and STI prevention. *Am J Community Psychol*. 2012;50(3-4):553-571.
23. Ott MA, Hunt AL, Katz AJ, Zaban LS. Tapping into community resiliency in rural adolescent pregnancy prevention: an implementation sciences approach. *Behav Med*. 2020;46(3-4):340-352.
24. Feldman Farb A, Margolis AL. The teen pregnancy prevention program (2010–2015): synthesis of impact findings. *Am J Public Health*. 2016;106(Suppl 1):S9-S15.
25. Garney W, Wilson K, Nelon J, Muraleetharan D, McLeroy K, Baletka DM. Ecological approaches to teen pregnancy prevention: an examination of evidence-based interventions. *Health Promot Pract*. 2019;20(4):494-501.
26. Coyle KK, Glassman JR. Exploring alternative outcome measures to improve pregnancy prevention programming in younger adolescents. *Am J Public Health*. 2016;106(Suppl 1):S20-S22. doi:10.2105/AJPH.2016.303383
27. Beidas RS, Dorsey S, Lewis CC, et al. Promises and pitfalls in implementation science from the perspective of U.S.-based researchers: learning from a pre-mortem. *Implement Sci*. 2022;17:55. doi:10.1186/s13012-022-01226-3
28. Juras R, Kelsey M. Meta-analysis of federally funded teen pregnancy prevention programs: final report. *Am J Public Health*. 2019;109(4):e1-e8. doi:10.2105/AJPH.2018.304925
29. Luyckx VA, Byskov J, Chandramohan D, et al. Highlighting the ethics of implementation research. *Lancet Glob Health*. 2019;7(9):e1170-e1171. doi:10.1016/S2214-109X(19)30286-5
30. Theobald S, Brandes N, Gyapong M, et al. Implementation research: new imperatives and opportunities in global health. *Lancet*. 2018;392(10160):2214-2228. doi:10.1016/S0140-6736(18)32205-0