Sticking with Programs That Do Not Work: The Role of Escalation of Commitment in Schools

Courtenay A. Barrett[1] · Dustin J. Sleesman[2] · Shelbie E. Spear[1] · Andryce Clinkscales[1] · Tazkira Amin[1]

Abstract
Schools are the most common site to implement evidence-based prevention programs and practices (EBPs) to improve behavioral and mental health outcomes among children and adolescents. Research has highlighted the critical role of school administrators in the adoption, implementation, and evaluation of such EBPs, focusing on the factors they should consider during the adoption decision and the behaviors needed for successful implementation. However, scholars have only recently begun to focus on the de-adoption or de-implementation of low-value programs and practices to make room for evidence-based alternatives. This study introduces escalation of commitment as a theoretical framework for understanding why school administrators may stick with ineffective programs and practices. Escalation of commitment is a robust decision-making bias in which people feel compelled to continue with a course of action even when performance indicators suggest it is not going well. Using grounded theory methodology, we conducted semi-structured interviews with 24 building- and district-level school administrators in the Midwestern United States. Results suggested that escalation of commitment occurs when administrators attribute the underlying causes of poor program performance not to the program itself but instead to issues related to implementation, leadership, or the limitations of the performance indicators themselves. We also identified a variety of psychological, organizational, and external determinants that accentuate administrators' continuance of ineffective prevention programs. Based on our findings, we highlight several contributions to theory and practice.

Keywords Escalation of commitment · Implementation science · Adoption · De-adoption · De-implementation · Strategic abandonment

Prevention science has emerged at the intersection of lifespan development, contextual factors, and preventative intervention trials to reduce the likelihood of poor long-term mental and behavioral health, economic, and physical outcomes in adulthood (Catalano et al., 2012; Fagan et al., 2019). Risk factors in childhood and adolescence, including bullying and victimization, academic failure, and substance abuse, may be experienced at schools (O'Connell et al., 2009). To help improve behavioral and mental health outcomes, evidence-based prevention programs and practices (EBPs) often focus on schools (Duong et al., 2021; Lyon et al., 2013) because they are an ideal context to target risk and protective factors, such as social networks, and they circumvent many of the barriers to community-based services (Bear et al., 2014).

Research has found that strong administrative leadership is necessary for successful EBPs in schools, so administrators play a critical role in their adoption, implementation, and evaluation (Aarons et al., 2014; Locke et al., 2019; Lyon et al., 2018). For example, administrators can serve as either an internal champion of an EBP or a gatekeeper who prevents its uptake. To optimize the adoption decision, administrators should consider a variety of factors, including program characteristics (e.g., feasibility and useability), the

✉ Courtenay A. Barrett
  morsicou@msu.edu

  Dustin J. Sleesman
  sleesman@udel.edu

  Shelbie E. Spear
  spearshe@msu.edu

  Andryce Clinkscales
  clinksc4@msu.edu

  Tazkira Amin
  amintazk@msu.edu

1 Michigan State University, East Lansing, USA

2 University of Delaware, Newark, USA

evidence base in support of its effectiveness, data indicating the extent to which it is needed, capacity for implementation, program fit with the local school context, and cost considerations (Fixsen et al., 2005; Metz et al., 2021). Then, school administrators should enact several behaviors to support implementation, such as becoming knowledgeable about the program, proactively developing an implementation plan, supporting implementers, and persevering throughout implementation (Lyon et al., 2018). District- and building-level administrators typically have unique roles in the adoption, implementation, and evaluation process. For example, district-level administrators often play a more central role in the adoption decision, whereas those at the building-level are generally more critical in the implementation process because they are closer to the implementers (e.g., teachers; Aarons et al., 2014). In sum, school administrators are a pivotal player in the delivery of school-based EBPs.

Although this body of literature has advanced our understanding of how EBPs can be successfully adopted and implemented in schools, researchers have only recently begun to recognize that *de-adoption* (or *de-implementation*) of previous practices or programs is often necessary to make room for more effective alternatives. The nascent work in this area has been conducted in healthcare settings, focusing on various therapeutic, diagnostic, and other healthcare-specific practices. Very few studies have been conducted in the education setting (with the exception of Nadeem & Ringle, 2016), and this study fills this gap in the literature. Moreover, recent reviews of the de-adoption literature have identified a need for a better theoretical understanding of the barriers and facilitators of de-adoption in specific contexts (Nilsen et al., 2020; Niven et al., 2016). The current study addresses this need by introducing escalation of commitment as a theoretical framework to explain why schools may continue to implement ineffective programs when administrators are aware that performance indicators are negative.

De-Adoption as the Key to Adoption? The Role of Escalation of Commitment

Escalation of commitment is a phenomenon in which people experience a strong desire to continue with a course of action (such as a decision or program) that has resulted in negative outcomes (Staw, 1976). Several defining characteristics must be present for escalation of commitment to occur. First, the decision-maker has invested resources, such as time or money, in pursuit of a course of action. Second, they become aware of negative performance indicators (feedback) suggesting that the course of action is not going well. Third, they face uncertainty about whether continuing will eventually result in success, creating a

dilemma of whether they should persist or stop the course of action (Brockner, 1992). Systematic reviews of the literature have indicated that escalation of commitment is a robust behavior that has been observed in a variety of contexts and fields, such as business, psychology, political science, international relations, and animal behavior (Sleesman et al., 2012, 2018), but the current research is the first to consider the role of escalation of commitment in education, prevention science, or implementation science.

Determinants of Escalation of Commitment

Research has identified a wide range of determinants that influence escalation of commitment. The most studied are psychological determinants, which focus on how cognitive and affective processes can bias people toward continuing with the failing course. For example, the more resources people have invested in the course (i.e., sunk costs), the more they want to press forward with it to avoid having to admit that their investments were wasteful (Arkes & Blumer, 1985). Other examples of psychological determinants include the extent to which the decision-maker's professional identity is at stake, they feel a sense of personal responsibility for the initial decision, or they are confident that success is imminent (Sleesman et al., 2012).

Other research has focused on escalation of commitment determinants from a multilevel perspective. For instance, the group or organization in which the decision-maker is embedded can profoundly shape behavior. Their colleagues can create an echo chamber of positivity by highlighting small signs of success or improvement while overlooking negative feedback (Van Oorschot et al., 2013). Similarly, an organizational culture that values success or consistency may discourage people from speaking up when problems emerge, especially when powerful individuals want the course of action to continue (Keil & Mähring, 2010). Beyond the organizational level, research has found that external determinants can play an important role as well. To illustrate, decision-makers may hold onto an ineffective course of action due to pressure from outside stakeholders or simply because other organizations are engaged in a similar course of action (Hsieh et al., 2015). Importantly, escalation of commitment scholars maintain that although many factors may influence continuance decisions, researchers must pay close attention to the particular context in which the decision is embedded (Sleesman et al., 2018). Doing so allows for a more nuanced and comprehensive understanding of escalation of commitment in that context, which is especially important if the aim is to develop recommendations for practitioners. In this spirit, our study presents an in-depth qualitative analysis of escalation of commitment in a school context.

Current Study

The research aim of this study was to develop an empirically based theoretical framework to understand how escalation of commitment inhibits schools from de-adopting ineffective programs. We utilized grounded theory methodology (Thornberg & Charmaz, 2014) because it allows for exploration and hypothesis generation, which is helpful when studying underdeveloped constructs like de-adoption. We conducted semi-structured interviews with 24 school administrators at the district and building levels, given their critical role in overseeing education programs (Lyon et al., 2018). The study advances knowledge in prevention and implementation science by identifying a theoretical framework for understanding the de-adoption of low-value practices, as a recent scoping review identified only five empirical studies examining de-adoption across any context (Nilsen et al., 2020). We do this by addressing the following research question: How does escalation of commitment inhibit the de-adoption of programs after school administrators review indicators that suggest the program is performing poorly?

Method

Grounded theory methodology was appropriate for the study because it emphasizes "individual and collective actions and social and social psychological processes, such as… organization changes, and establishing and maintaining workplace practices" (Thornberg & Charmaz, 2014, p. 154). This approach mirrors extant de-adoption literature conducted in education (Nadeem & Ringle, 2016) and public health (Pinto & Park, 2019). In addition, grounded theory includes a range of techniques that embrace an iterative and inductive data analytic process, using comparative methods to develop new conceptual categories that emerge from data. As such, it supports exploration of a relatively new area of research. Grounded theory also emphasizes actions and process as opposed to themes and structure, allowing researchers to identify sources of variation within the de-adoption process. This can reveal fascinating insights when examining constructs that unfold over time, like escalation of commitment.

Sampling and Participants

We used a stratified purposeful sampling strategy to allow for maximum variation across organizational structures (Palinkas et al., 2015). A total of 24 administrators in one Midwest state participated in the study at which point we met theoretical saturation. This satisfies the recommended sample size of 20–30 participants for qualitative studies using semi-structured interviews (Vasileiou et al., 2018). Of the 23 participants who completed the demographic survey,

one identified as Black/African American, and the rest (91.7%) were White. Slightly over half of the participants were female (*n* = 13; 56.5%). These characteristics reflect those in the state, with 87.4% of administrators identifying as White and female (54.0%; National Center for Education Statistics, 2020). The average age of participants was 47.22 years (SD = 6.74 years). On average, participants served 13.00 years in their district (SD = 8.87 years) and had previously worked in 2.74 other districts (SD = 1.67). Most participants had master's (*n* = 11; 47.8%) or professional degrees (e.g., educational specialist; *n* = 8; 34.8%), and four had doctoral degrees (17.4%).

Participants were approximately evenly distributed across elementary (*n* = 8; 33.3%), middle (*n* = 6; 25.0%), secondary (*n* = 6; 25.0%), and district-level (*n* = 4; 16.6%) settings. They represented schools or districts in all forms of geographic locations, including city (*n* = 1; 4.2%), suburban/town (*n* = 13; 56.5%), and rural (*n* = 10; 41.7%) settings based on NCES codes. This composition reflected suburban/town schools in the state (49.2%), although there was an underrepresentation of schools/districts in cities (22.6%) and an overrepresentation of schools/districts in rural locations (28.1%; National Center for Education Statistics, 2017).

Data Gathering and Analysis

We drafted the interview protocol based on theory and research on escalation of commitment to ensure that responses met the defining characteristics of the phenomenon, which we discussed earlier (Brockner, 1992). The protocol was also designed to encourage participants to provide transparent and candid responses rather than ask them to justify or defend their previous actions. We did this by asking participants to discuss why escalation occurs for administrators *in general* and to recall how a *colleague* had engaged in the behavior, before discussing their own behavior (as we detail later).

Next, we piloted the protocol with three administrators with over 75 cumulative years of experience at regional, district, and building levels–and we refined the protocol based on their feedback. The interviews comprised three standardized sections to elicit administrators' perspectives on the (1) factors they considered during the adoption decision, (2) types of indicators they would review to determine how well a program was performing, and (3) aspects that related to escalation of commitment. The first two sections provide context for the third section, which was the focus of the study. The final section was partitioned into three focused blocks involving asking participants to (a) discuss the reasons why an administrator might, generally speaking, choose to continue a program when indicators suggested the program was not performing well, (b) provide an example of a colleague who chose to continue a program

when the indicators suggested the program was not performing well, and (c) describe their own lived experiences as an administrator continuing programs when indicators suggested a program was not performing well.

Importantly, the escalation of commitment section was crafted to explicitly include the indicators they had mentioned in the second section, which ensured that it was their own, locally identified indicators that suggested the program was ineffective. This helped to reinforce our aim to contextualize escalation of commitment as much as possible. Participants were proactively instructed to "think aloud" during their responses by verbally communicating their reasoning for each answer (DeSimone & Le Floch, 2004). We also prompted for more information about the situational factors contributing to the decision-making process, what administrators were thinking or feeling throughout the process, and how they weighted or prioritized various factors.

We conducted the interviews between February and May 2021, and they lasted an average of 37.14 min (SD = 8.26 min; range = 28.06 to 52.42 min). They were recorded and transcribed using an online transcription service. Then, the third and fifth authors cleaned the transcripts to remove identifying information and check for accuracy. Next, transcripts were randomly assigned to each of the first four authors, who analyzed them with several approaches that occurred simultaneously and iteratively. First, these authors used both *memo writing* (a record of their thoughts throughout the analytical process) and *open coding* of the transcripts (brief notes often framed as gerunds to emphasize actions and processes). Afterwards, they created *focus codes*, which were developed based on the most significant or frequent open codes, allowing them to sift through the large amounts of data (Thornberg & Charmaz, 2014).

Then, the first four authors developed diagrams to illustrate the process and relationships among the focus codes. After each transcript had been analyzed by one member, the others independently reviewed all the open codes, focus codes, and diagrams across all of the transcripts. Lastly, they met to engage in constant comparison of the data to develop consensus about the new conceptual categories and the relationships among them and to identify the determinants that contributed to variability (moderation) within the identified relationships. Further, to ensure trustworthiness of the data, we employed a member checking process through which we presented participants with initial focus codes and transcripts of their own interviews (Birt et al., 2016), which occurred approximately 4–6 weeks after the first interview. All participants indicated they found the data gathering and analytic process to accurately and comprehensively reflect their administrative experiences in schools.

Results

We first summarize how participants responded to the initial questions about the factors they typically consider during the program adoption decision and the types of indicators they would review to determine how well a program was performing. First, participant responses regarding the adoption decision generally reflected what has been described previously in the literature (e.g., Metz et al., 2021). Administrators emphasized the collaborative or team-based nature of the decision, through which they often considered the research base of the program, stakeholder buy-in, program costs, their capacity to implement the program and available support for it, and program characteristics, including useability, feasibility, enrichment or remediation materials, technological capabilities for virtual delivery, and language accessibility. They also emphasized the importance of contextual fit (e.g., alignment with the school's mission, goals, or current programming) and external constraints, such as state-approved curricula.

Next, participants identified several indicators they would rely on to discern the performance of a program. Their responses largely aligned with the existing program evaluation and progress monitoring literature (e.g., Royse et al., 2016). Indicators included *quantitative* student academic, behavioral, or social-emotional data that were both formative (e.g., curriculum-based benchmarks and social-emotional screeners) and summative (e.g., state assessments). Administrators also highlighted the importance of *qualitative* feedback from teachers, students, parents, and community members; observations or classroom walkthroughs; and qualitative observations of student growth (e.g., fewer externalizing behaviors in the hallways). Most administrators stated they prioritized quantitative student data and qualitative teacher input more than the other types of indicators.

We now describe the focus of our research, namely, the determinants that participants identified and the ways in which they contribute to escalation of commitment to school programs. A summary of focus codes, their definitions, and examples from this section of the interview can be found in the Online Supplementary Materials. Figure 1 displays the escalation of commitment theoretical framework that emerged from the data. First, administrators noted that a formal or informal event triggers a review of the indicators regarding the program's performance. Formal events included legislation that mandated program changes, regularly scheduled data review days, internal curriculum review cycles, and other periodic meetings (e.g., board or staff meetings). Informal events included receiving feedback from teachers, students, parents, or community members outside of a formal setting. After identifying these triggering events, administrators
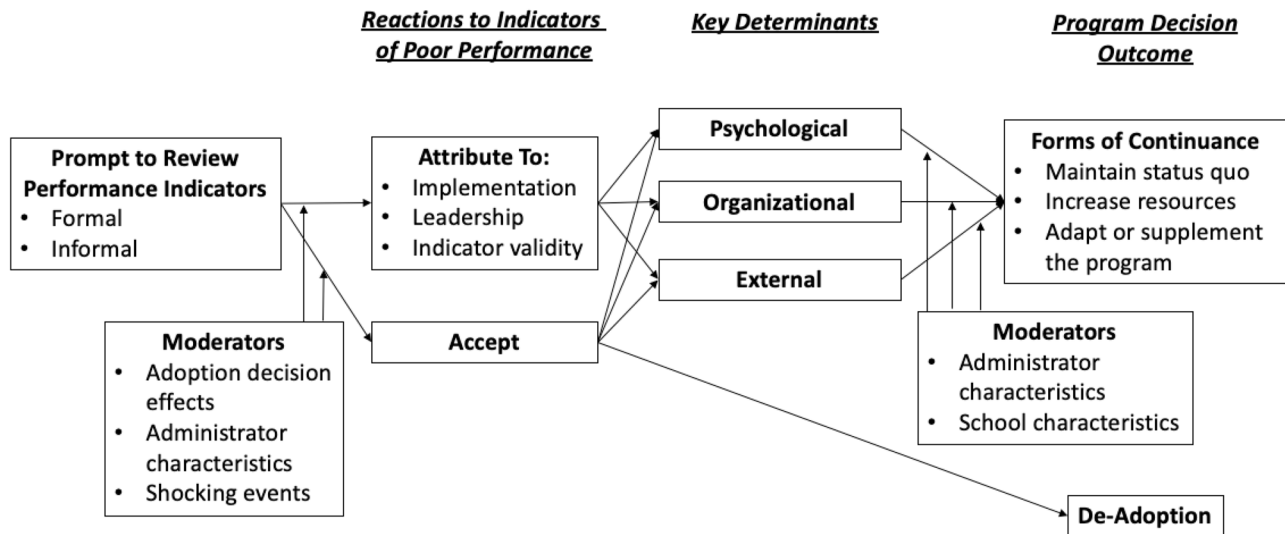
Fig. 1 Theoretical framework for escalation of commitment in schools

described various ways in which they reacted to them, mainly to (a) attribute the poor performance to other factors or in some cases and (b) accept them.

Reactions to Indicators of Poor Performance

Attribute to Other Factors Administrators reported attributing indicators of poor performance to factors other than program choice. First, administrators noted that intervention fidelity can obscure assessments about the performance of a program. They highlighted that sometimes implementers do not receive enough training, insufficient time is allocated for implementation, or there is a lack of buy-in for the program. These issues make it difficult for administrators to disentangle whether the poor performance indicators reflect a low-value program or merely an implementation problem.

Second, administrators noted that a lack of leadership could be a cause for program failure, rather than the program itself being low-value or a poor fit with the local context. In fact, some district-level administrators noted that they might consider replacing the building-level administrator rather than the program, under the assumption that a weak leader would be unable to support *any* program successfully. Other administrators emphasized the need to provide leaders with more training and coaching to support the implementation process.

Third, administrators reported that they might question the validity of the indicators themselves, even though they were locally defined. They described two main concerns: (a) ambiguity about measurement or the analytical process and (b) timing of the review of indicators. In terms of the former, administrators highlighted several situations in which

measurement or analysis issues clouded their interpretation of performance indicators. They noted having a lack of relevant measures that were closely aligned with the wide range of programs they implemented and the challenges of measuring long-term outcomes, such as student success in post-secondary education. They also described the problem of not having clear goals with which to compare progress. Thus, while there may have been some modest improvement of student skills, the extent to which it could have been achieved through maturation alone may be unclear. Administrators also described the difficulty of measuring success when programs work for some students but not others. Vertical alignment and programming coherence across grade levels and student groups (e.g., special education, gifted and talented) were common factors in the adoption decision, but this also created ambiguity when programs were effective at some grade levels but not others due to the intersection of child development trajectories and program characteristics. They also reported that program quality varied across supplemental components (e.g., remediation, enrichment, or language translations), potentially leading to variation in program effectiveness across different student populations.

Regarding the timing of indicators, there were often unclear guidelines for how long it should take implementers to master the learning curve associated with a new program or practice and that observable (and measurable) improvements in student data may not materialize for several years post-adoption. This concern aligns with variations noted in the literature, ranging from one to three years (Durlak & Dupre, 2008). Essentially, the events triggering a review of indicators were either not aligned with the adoption and implementation process or were unanticipated. Several

administrators noted the "implementation dip" (Fullan, 2001) when implementation skills and confidence temporarily deteriorate after learning a new program or practice. They were unsure if some poor performance indicators, such as qualitative teacher feedback or observations, were sometimes the result of this dip as opposed to reflecting a low-value program. In cases where administrators attributed poor performance indicators to the normative change process, they noted the importance of resilience and perseverance. Ironically, these traits are often viewed as necessary for educational leaders, especially for program implementation (Lyon et al., 2018)–and yet such perseverance is shown when they persist with ineffective programs without a clear way of differentiating among underlying causes of poor performance. Nearly all the participants mentioned the significant amount of subjectivity in determining the extent to which programs were effective.

Accept Some participants stated that indicators may be accepted at face value, with two of them noting that they would, hypothetically speaking, not continue with an ineffective program. For example, one elementary school principal said, "Well, they'd be a bad administrator if they're continuing something that's not working." Put another way, the only time administrators reported they would de-adopt a program is if they clearly attributed the poor indicators to a low-quality program or poor program choice.

Variation Within Reactions (Moderators) Results revealed three sources of variability that influenced how administrators reacted to the indicators: (a) adoption decision effects, (b) administrator characteristics, and (c) shocking events. First, administrators described how social influences during the program adoption decision might carry over into how the indicators were interpreted, such that buy-in may have changed over time or a fraction of implementers may have undermined the program, which could affect intervention fidelity. Second, participants noted how their attributions about the performance indicators may be influenced by their beliefs, values, and attitudes toward data, research, and other factors. For instance, administrators who de-prioritized the research base of a program during the adoption decision were less likely to attribute poor indicators to program quality. Lastly, administrators described several illustrations of "shocking" events (Morgeson et al., 2015) that would cause them to attribute failing indicators to factors other than program quality or fit. Examples included the COVID-19 global pandemic, school shootings, and other unexpected traumas occurring within the school (e.g., fatal accidents or a homicide) that would prevent students from benefitting from a program. In these examples, poor performance indicators were interpreted as a lack of student readiness to learn or teacher readiness to implement.

Key Determinants

Administrators described three categories of determinants that could affect their decision to continue a program after their interpretation of the performance indicators: (a) psychological, (b) organizational, and (c) external factors.

Psychological Administrators discussed several psychologically oriented pressures that influenced the continuance decision, including fear of innovation fatigue, change, or harming their own reputation (Bolino et al., 2016). They also described personal benefits for the administrator if the program was sustained, feelings of personal ownership over the decision to adopt the program, and hope for improvement of the indicators over time. For example, one elementary school principal said, "It really was a self-preservation decision made more on politics versus what was best for kids." Administrators also reported how leadership style, professional identity, and personal notions of what it meant to be a "good leader" can compel them to persist with programs. For example, one high school principal said, "[Good leaders] avoid these situations, because they've learned to do enough processing before and strategizing that it's just their idea, and that they've done the work around it." Finally, administrators noted that sunk costs, or resources that have been expended in the past and cannot be recovered (Arkes & Blumer, 1985), can influence their decision. For example, one elementary school principal said, "That was the reason I recall as to why [the underperforming program] was still being used. They had spent a lot of money not that long ago and weren't prepared to just eat the cost and start all over again." These resources included financial costs associated with purchasing the program; time spent garnering buy-in; resources needed for training, professional development, and coaching; and grant dollars that had been expended. While sunk costs leave fewer resources to pursue better alternatives, they also have a psychological effect on decision-makers to press forward to avoid feeling like their investments were made in vain (Sleesman et al., 2012).

Organizational Administrators also mentioned several pressures at an organizational level, largely focused on their building or school district, which comprised a wide variety of stakeholders, such as teachers, students, and other school buildings. They described being influenced by internal stakeholder support for the program or fear of change, as well as interdependencies with other schools in their district. For example, when prevention programs were adopted at the district level across multiple buildings, collective decision-making prevented de-adoption if the program was effective (or perceived to be) in some buildings or grade levels but not others. This was particularly challenging when districts served buildings that were very diverse, with some buildings having

significantly more resources, different student demographics, or different organizational structures or climates than others.

Shifting priorities also led administrators to continue programs that they knew were ineffective. For example, if a poorly performing reading program had been adopted two years prior and a school planned to adopt a new social-emotional learning program that year, they were unlikely to abandon that reading program for a new one, as they would not be able to garner sufficient buy-in or resources to initiate two new programs simultaneously. Lastly, administrators described institutionalization as another key organizational pressure that influenced their decision to continue a program. For instance, a program may have become a tradition over time such that the identity of the school was tied to its existence. This made it seem impossible to de-adopt the program, no matter how ineffective the performance indicators suggested the program was. This appeared to be particularly true for programs that were developed internally or when key stakeholders had a vested interest in the program.

External Administrators also discussed pressures from sources external to the school district, including community members, the state department of education, and other stakeholders not directly tied to the district. Social influences in the community sometimes created a sense of institutionalization when the identity of the community became closely tied to the program regardless of how the program was performing. For example, one middle school principal said, "[The program] was funded off a parent…It was a cherished district presentation…But, it was not impactful. [The students] just could not really grasp it. But for many years as a district, we continued that program just because the decision makers were still so moved by the moment of time [when the adoption decision was made]."

Interestingly, administrators reported they were sometimes *required* to continue with an underperforming program for reasons such as state mandated programs, practices, policies, or assessments; multiyear contractual obligations with the vendor that required them to offer a program for a predetermined amount of time; and teacher contractual obligations that required them to continue certain programs or practices until the contract could be renegotiated. For example, one middle school principal said, "You might be locked into it, it might have been a two-year commitment or something, or you might be contractually obligated." Administrators also highlighted the fact that sometimes there are no alternative programs available from vendors or other program suppliers. In these cases, they acknowledged that the indicators signaled an ineffective program, but they were stuck with it because they were not aware of other options.

Variation Within Determinants (Moderators) Participants revealed several administrator and school characteristics that affected how the various determinants shaped their decision about continuing a program. For example, they noted how their years of experience or closeness to retirement would affect their susceptibility to psychological and organizational determinants. Their rationale was that administrators who were closer to retirement had the capability to take more risks, push against the status quo, or dismiss institutionalization pressures. Alternatively, other administrators noted how, earlier in their own careers, they were less susceptible to social influences (e.g., stakeholder beliefs) but more susceptible to psychological influences (e.g., reputation concerns). Regarding school characteristics, participants illustrated how school demographic composition influenced the kinds of external resources available to them, mitigating the influence of sunk costs in particular. For example, administrators from districts that did not qualify for external funding (e.g., Title I) but were also situated in less affluent communities found themselves particularly susceptible to honoring their previous expenditures.

Program Decision Outcome

The last aspect of the theoretical framework that emerged from our interviews was the administrator's decision about the program. Although they were not specifically asked about the various ways that such continuance may occur, administrators referred to three different possibilities: maintain the status quo, increase resources allocated to the program, or adapt or supplement the program. Administrators noted they may boost program resources to increase its chances of success, particularly when they attributed the cause of the poor indicators to low intervention fidelity or poor leadership. In these cases, resources included additional training, professional development, and coaching for leaders and implementers, more time allocated for implementation, or more materials (e.g., workbooks or rewards). Lastly, administrators reported continuing with an underperforming program vis-à-vis adapting various facets of it, implementing it at another grade-level or student population, or purchasing an additional program to supplement the existing one. These strategies were different than increasing resources allocated to the program because they involved changing the nature of the program itself. However, administrators noted that it was unclear when program adaptations yielded a completely new program. Regardless, their intention was to make changes in the hope of turning the underperforming program around.

Discussion

Overall, this study advanced scientific knowledge by investigating how escalation of commitment can serve as a theoretical framework for understanding the de-adoption

of low-value practices. Implementation science has high-lighted the critical role that administrators play throughout the EBP adoption and implementation processes in schools (Aarons et al., 2014; Lyon et al., 2018). Our results extend this literature by emphasizing how administrators can subsequently influence de-adoption, as leadership appeared to permeate multiple determinants of the escalation decision. Administrators emphasized the various pressures, and sometimes *requirements*, to continue programs even in the face of negative feedback. The de-adoption of low-value practices and the adoption of EBPs require collaboration across the research, policy, and practice sectors to create conditions in which administrators do not feel they need to choose between their own interests (e.g., "self-preservation" as noted by one elementary school principal) and what is most likely to improve student outcomes.

Further, the evidence base supporting a program was only considered during the adoption phase, but not during the escalation of commitment decision. For example, administrators did not mention how the rigor of the evidence base might help them differentiate between implementation issues and program quality, how alignment between the samples and context of the original research studies and their local setting might influence program fit, or how the original research studies might provide guidance for how to select or interpret indicators. Further, it was not always clear how administrators conceptualized "research." Academics and practitioners use the term differently (Mills et al., 2020), with academics referring to rigorous research designs (e.g., randomized controlled trials) and practitioners emphasizing expert testimonials, implementer feedback, and stakeholder buy-in (Honig & Coburn, 2008). Researchers, policy-makers, and practitioners should also consider how each of the factors considered during the adoption decision is weighted. If the evidence base of the program is not prioritized during the adoption decision, student outcomes are unlikely to improve even if low-value practices are de-adopted. The potential benefit of de-adopting low-value practices *assumes* that the next adoption decision will yield successful implementation of a more effective EBP.

Importantly, all of the participants could provide multiple examples of themselves and other administrators continuing programs even after locally identified indicators suggested the program was not going well. Nearly all of the administrators highlighted the subjectivity involved when reviewing and interpreting indicators, and escalation of commitment tends to be especially strong when indicators are ambiguous (Ross & Staw, 1993). Indeed, participants noted it was unlikely that every source and type of indicator in schools would unequivocally suggest program failure, highlighting why it may be so challenging for schools to abandon low-value programs.

Finally, responses to questions about the adoption and escalation of commitment decisions differed in two critical ways. First, administrators were able to articulate an explicit process for how to make the adoption decision and the types of indicators they would review. Some administrators even noted they had received professional development for these activities from state or regional technical assistance centers. By contrast, *how* to interpret program performance indicators and respond accordingly was not an explicit or predetermined process. Rather, it was implicit, with multiple administrators responding throughout the interviews with intrigue, noting that "these were good questions" and that the interview was helping them think through the decision in real time. Second, the adoption and continuance decisions differed in that all administrators emphasized that the adoption decision was a team effort, yet none of them mentioned that the continuance decision was necessarily team-based or collaborative. Together, results have implications for prevention science, as it focuses on disseminating and implementing preventative EBPs to improve long-term outcomes.

Implications for Prevention Science

Results highlight the importance of training school administrators on the de-adoption of low-value practices and the adoption of EBPs, as well as the types of strategies that may be useful in supporting their implementation. Training should explicitly address how factors are weighed, prioritizing the rigor of the evidence base supporting the program. A common de-biasing strategy in the field of escalation of commitment includes setting clearly defined benchmarks ahead of time so decision-makers can make better continuance decisions later on (Keil & Mähring, 2010). Resources should be allocated to training administrators using evidence-based approaches, such as the Leadership and Organizational Change for Implementation protocol (Aarons et al., 2017), which could help administrators differentiate between poor intervention fidelity and poor program performance. Administrators should also collaborate with colleagues when making the continuance decision, as this can help them to sift through the complexity from multiple perspectives. However, it is important to note that group decision-making can sometimes *exacerbate* escalation of commitment (Sleesman et al., 2018), so it is critical that each team member is willing to speak up when problems surface and not be afraid to let go of low-value programs.

Limitations

Despite the merits of this study, there are several limitations worth noting. First, the sample only included administrators from one state in the Midwestern United States,

with an underrepresentation of administrators of color and in city settings and an overrepresentation in rural settings. This may have influenced the results as administrator and school/district characteristics may moderate the relationship between each of the determinants and the program continuance decision. Second, the interviews may have been influenced by social desirability. Although we made significant attempts to encourage honesty by continuously reminding participants that we were interested in their opinions and experiences rather than evaluating them, it may have been difficult for participants to be candid when discussing decision-making, research evidence, and program failure with external researchers.

Suggestions for Future Research

Future research may wish to replicate or expand on the current study with a more diverse sample to understand how different determinants or processes may drive decision-making in a broader array of schools (e.g., urban settings, regions of the country, and available resources). Further, future research should employ quantitative methods to build upon these qualitative results and empirically test the relationships proposed in the model using a nationally representative sample that mirrors the demographics of school administrators and school characteristics. We highlighted potential moderators, such as school resources (e.g., Title I funding), student demographics, district size and location, years of experience, leadership style and professional identity, and orientations towards research and data, which should all be considered in future research. Scholars should also explore effective approaches or strategies for training school administrators throughout the adoption and implementation process, given that results suggested that intervention fidelity may be one cause of poor performance indicators.

Conclusion

This was the first study to examine escalation of commitment in a school context, and it proved to develop a valuable theoretical framework for understanding the determinants that can prevent the de-adoption of low-value practices to make room for the adoption of evidence-based alternatives. Results highlighted determinants across multiple levels, and they shined a light on the importance of leadership throughout the entire decision process. The study offers a foundation for future research to test the proposed model, and it provides an empirical rationale for allocating resources to administrator training to support decision-making and the implementation of EBPs.

Declarations

Ethics Approval This study was determined to be exempt by the Michigan State University Human Research Protection Program, Office of Regulatory Affairs.

Consent to Participate Informed consent was obtained from all participants in the study.

Conflict of Interest The authors declare that they have no conflict of interest.

References

Aarons, G. A., Ehrhart, M. G., Farahnak, L. R., & Sklar, M. (2014). Aligning leadership across systems and organizations to develop a strategic climate for evidence-based practice implementation. *Annual Review of Public Health, 35*, 255–274.

Aarons, G. A., Ehrhart, M. G., Moullin, J. C., Torres, E. L., & Green, A. E. (2017). Testing the leadership and organizational change for implementation (LOCI) intervention in substance abuse treatment: a cluster randomized trial study protocol. *Implementation Science, 12*(1), 29. PMCID: PMC5335741.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*, 124–140.

Bear, L., Finer, R., Guo, S., & Lau, A. S. (2014). Building the gateway to success: an appraisal of progress in reaching underserved families and reducing racial disparities in school-based mental health. *Psychological Services, 11*(4), 388–397. https://doi.org/10.1037/a0037969

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member checking: a tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research, 26*(13), 1802–1811. https://doi.org/10.1177/1049732316654870

Bolino, M., Long, D., & Turnley, W. (2016). Impression management in organizations: critical questions, answers, and areas for future research. *Annual Review of Organizational Psychology and Organizational Behavior, 3*, 377–406.

Brockner, J. (1992). The escalation of commitment to a failing course of action: toward theoretical progress. *The Academy of Management Review, 17*(1), 39–61. https://doi.org/10.2307/258647

Catalano, R. F., Fagan, A. A., Gavin, L. E., Greenberg, M. T., Irwin, J. C. E., Ross, D. A., & Shek, D. T. (2012). Worldwide application of prevention science in adolescent health. *The Lancet, 379*(9826), 1653–1664. https://doi.org/10.1016/S0140-6736(12)60238-4

DeSimone, L. M., & Le Floch, K. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis, 26*, 1–22.

Duong, M. T., Bruns, E. J., Lee, K., Cox, S., Coifman, J., Mayworm, A., & Lyon, A. R. (2021). Rates of mental health service utilization by children and adolescents in schools and other common service settings: a systematic review and meta-analysis. *Administration and Policy in Mental Health, 48*(3), 420–439. https://doi.org/10.1007/s10488-020-01080-9

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3–4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Fagan, A. A., Bumbarger, B. K., Barth, R. P., Bradshaw, C. P., Cooper, B. R., Supplee, L. H., & Walker, D. K. (2019). Scaling up evidence-based interventions in US public systems to prevent behavioral health problems: challenges and opportunities. *Prevention Science, 20*, 1147–1168.

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: a synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network. FMHI Publication #231.

Fullan, M. (2001). *Leading in a culture of change*. San Francisco, CA: Jossey-Bass.

Honig, M. I., & Coburn, C. (2008). Evidence-based decision making in school district central offices toward a policy and research agenda. *Educational Policy, 22*, 578–608. https://doi.org/10.1177/0895904807307067

Hsieh, K.-Y., Tsai, W., & Chen, M.-J. (2015). If they can do it, why not us? Competitors as reference points for justifying escalation of commitment. *Academy of Management Journal, 58*(1), 38–58. https://doi.org/10.5465/amj.2011.0869

Keil, M., & Mähring, M. (2010). Is your project turning into a black hole? *California Management Review, 53*, 6–31.

Locke, J., Lee, K., Cook, C. R., Frederick, L., Vázquez-Colón, C., Ehrhart, M. G., Aarons, G. A., Davis, C., & Lyon, A. R. (2019). Understanding the organizational implementation context of schools: a qualitative study of school district administrators, principals, and teachers. *School Mental Health, 11*(3), 379–399. https://doi.org/10.1007/s12310-018-9292-1

Lyon, A. R., Cook, C. R., Brown, E. C., Locke, J., Davis, C., Ehrhart, M., & Aarons, G. A. (2018). Assessing organizational implementation context in the education sector: confirmatory factor analysis of measures of implementation leadership, climate, and citizenship. *Implementation Science, 13*, 5.

Lyon, A. R., Ludwig, K. A., Vander Stoep, A., Gudmundsen, G., & McCauley, E. (2013). Patterns and predictors of mental healthcare utilization in schools and other service sectors among adolescents at risk for depression. *School Mental Health, 5*, 155–165.

Metz, A., Albers, B., Burke, K., Bartley, L., Louison, L., Ward, C., & Farley, A. (2021). Implementation practice in human service systems: understanding the principles and competencies of professionals who support implementation. *Human Service Organizations: Management, Leadership & Governance, 45*, 238–259. https://doi.org/10.1080/23303131.2021.1895401

Mills, K. J., Lawlor, J. A., Neal, J. W., Neal, Z. P., & McAlindon, K. (2020). What is research? Educators' conceptions and alignment with United States federal policies. *Evidence & Policy, 16*(3), 337–358. https://doi.org/10.1332/174426419X15468576296175*

Morgeson, F. P., Mitchell, T. R., & Liu, D. (2015). Event system theory: an event-oriented approach to the organizational sciences. *Academy of Management Review, 40*(4), 515–537.

Nadeem, E., & Ringle, V. (2016). De-adoption of an evidence-based trauma intervention in schools: a retrospective report from an urban school district. *School Mental Health, 8*(1), 132–143. https://doi.org/10.1007/s12310-016-9179-y

National Center for Education Statistics. (2017). *Number of city, suburban, town, and rural regular, operating public elementary and secondary schools with student membership and percentage distribution of students in membership, by state or jurisdiction: school year 2015–16*. Retrieved November 6, 2022, from https://nces.ed.gov/pubs2018/2018052/tables/table_04.asp

National Center for Education Statistics. (2020). *Characteristics of public school principals*. Retrieved October 12, 2022, from https://nces.ed.gov/programs/coe/indicator/cls

Nilsen, P., Ingvarsson, S., Hasson, H., von Thiele Schwarz, U., & Augustsson, H. (2020). Theories, models, and frameworks for de-implementation of low-value care: a scoping review of the literature. *Implementation Research and Practice, 1*, 1–15.

Niven, D. J., Leigh, J. P., & Stelfox, H. T. (2016). Ethical considerations in the de adoption of ineffective or harmful aspects of healthcare. *Healthcare Management Forum, 29*, 214–217. https://doi.org/10.1177/0840470416646632

O'Connell, M. E., Boat, T., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: progress and possibilities*. Washington, DC: The National Academies Press. Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults: Research Advances and Promising Interventions.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health, 42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Pinto, R. M., & Park, S. E. (2019). De-implementation of evidence-based interventions: implications for organizational and managerial research. *Human Service Organizations: Management, Leadership & Governance, 43*, 336–343.

Ross, J., & Staw, B. M. (1993). Organizational escalation and exit: lessons from the Shoreham nuclear power plant. *Academy of Management Journal, 36*, 701–732.

Royse, D., Thyer, B. A., & Padgett, D. K. (2016). *Program evaluation: an introduction to an evidence-based approach*. Boston, MA: Cengage Learning.

Sleesman, D. J., Conlon, D. E., McNamara, G., & Miles, J. E. (2012). Cleaning up the big muddy: a meta-analytic review of the determinants of escalation of commitment. *Academy of Management Journal, 55*, 541–562.

Sleesman, D. J., Lennard, A. C., McNamara, G., & Conlon, D. E. (2018). Putting escalation of commitment in context: a multilevel review and analysis. *Academy of Management Annals, 12*, 178–207.

Staw, B. M. (1976). Knee-deep in the big muddy: a study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16*, 27–44.

Thornberg, R., & Charmaz, K. (2014). Grounded theory and theoretical coding. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 153–169). SAGE Publications Ltd. https://doi.org/10.4135/9781446282243.n11

Van Oorschot, K. E., Akkermans, H., Sengupta, K., & Van Wassenhove, L. N. (2013). Anatomy of a decision trap in complex new product development projects. *Academy of Management Journal, 56*(1), 285–307. https://doi.org/10.5465/amj.2010.0742

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterizing and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology, 18*(1), 148.