aperio

Chow, J. C., & Hampton, L. H. (2025). Implementation of an adaptive intervention design to improve pre-service speech-language pathologists' knowledge of behavior support practices. *Research in Special Education*, *2*, 1–17. https://doi.org/10.25894/rise.2834

RiSE
RESEARCH IN SPECIAL EDUCATION

ALETHEIA SOCIETY

EMPIRICAL ARTICLE

Implementation of an Adaptive Intervention Design to Improve Pre-Service Speech-Language Pathologists' Knowledge of Behavior Support Practices

Jason C. Chow*, Vanderbilt University, US, jason.chow@vanderbilt.edu

Lauren H. Hampton, University of Texas at Austin, US

*Corresponding author

ABSTRACT

Purpose: The purpose of this registered report was to improve pre-service speech-language pathologists' (SLPs) behavior management knowledge and self-efficacy through adaptive intervention options.

Method: Using a sequential multiple-assignment randomized trial (SMART) design, we aimed to identify the best pathways to knowledge that considers an individual's response to instruction. We initially randomly assigned participants to either read a practitioner article or complete an online video module on two evidence-based behavior support practices.

Results: Through a series of re-randomizations and intervention components, results suggested that overall knowledge improved across participants regardless of intervention sequence. We intended to explore moderators of intervention efficacy (content engagement, self-efficacy) and assess social validity; however, our final sample was too small for a meaningful analysis. We reported our preliminary findings and added a post-hoc survey and conducted a focus group to explore issues of implementation.

Submitted: 10 April 2025 **Accepted:** 12 October 2025 **Published:** 01 December 2025

Conclusion: Implementation of the SMART design was a challenge to full participation. We couch the discussion of our results in the context of the design, implementation, and areas of learning and promise, and directions for future research.

Keywords: adaptive intervention; behavior; speech-language pathology; pre-service

Many school-based professionals do not receive adequate behavior management training (Flower et al., 2017; Oliver & Reschly, 2010), and it is becoming evident that speech-language pathologists (SLPs) rarely receive any dedicated content in behavior management. In a survey study of 223 school SLPs (Chow & Wallace, 2021), participants reported a minimal focus of their undergraduate and graduate programs on behavior management, with over 90% of participants reporting no courses in this area. In contrast, 95% of participants worked with children who exhibited challenging behaviors, suggesting a mismatch between pre-service and in-service experiences related to behavior. This is a particular point of concern, as we know 81% of children with behavior disorders have clinically meaningful and unidentified language deficits (Hollo et al., 2014), and these deficits emerge across subdomains of language (Hollo et al., 2020). Further, children with language delay demonstrate more problem behaviors and are twice as likely to demonstrate behavior problems later on than their typical peers (Curtis et al., 2018; Yew & O'Kearney, 2013). As such, an explicit and early focus on behavior management for SLPs is essential, because school-age children with behavior problems are likely to need language intervention (Chow, 2018; Chow & Wehby, 2019). Because children with language delay who exhibit problem behavior are less likely to engage in instruction and social interactions (Qi et al., 2006), limiting learning opportunities in the classroom, equipping our workforce with the tools to effectively enhance the quality of early learning experiences is essential.

NEED FOR EFFECTIVE BEHAVIOR SUPPORTS

Behavior management is broadly defined as the skills and strategies educators use to effectively address and support a student's behavior and interactions to maximize time for instruction and engagement. Effective behavior management sets the stage for high-quality instruction which can lead to increased engagement and learning opportunities. In the context of language intervention, practitioners should focus on first establishing an ideal setting to situate high-quality intervention (Chow et al., 2020; Cunningham et al., 2022). Because behavior management is a top concern of school-based professionals (Greenberg et al., 2014), we must prioritize equipping SLPs with effective tools to increase the likelihood that their intervention is most effective. The need for and importance of proactive behavior support settings is well documented and is understood to be an important element for effective intervention and instruction (Hemmeter et al., 2021; McLeod et al., 2017). As a field, systems need to effectively support SLPs to deliver intervention to children aligned with their individualized needs to ensure that they are receiving the support required to promote learning and social development. If SLPs are frequently experiencing challenging behavior but not equipped with the strategies to manage behavior effectively, it is likely that children may not be receiving the extent of the services required for them to demonstrate meaningful progress. This is particularly important, because SLPs are essential providers of special education services under the Individuals with Disabilities Education Act (IDEA, 2004).

THEORY OF CHANGE

Our theory of change for the adult learning that will occur via this intervention argues that learners (in this case, SLPs) may require different levels and modalities of instruction to be able to effectively integrate new knowledge. Instructional approaches vary, including activities like independent reading to learn, multimedia instruction via a virtual environment,

and live, face-to-face instruction. All these approaches are commonly used in education and adult learning and follow calls for the importance of demonstrating effectiveness in a range of contexts (Wayne et al., 2008), and we argue that in order to maximize outcomes and to provide instruction that accounts for and recognizes individual differences within learners, adapting the instructional materials and delivery to meet individual adult learning needs may improve overall outcomes. In this study, the goal is to empirically test and compare the relative effects of the components and their sequence to inform how typical modes of adult learning and instruction operate. To do this, we have selected a recent topic important to practitioners in the field of speech-language pathology.

We situate the need for behavior management during SLP service delivery using ecological-transactional theory with an interaction-centered model of language and behavioral development (Bronfenbrenner, 1994; Chow, 2018; Sameroff, 2009). In the interaction-centered model (Chow et al., 2020), effective behavior management provides an environment that allows SLPs to increase the quality and quantity of language learning opportunities. We situate this integrative model within an ecological-transactional framework to contextualize our understanding of how proactive behavior supports can maximize the utility of language intervention within dynamic learning environments. This model framework emphasizes the dynamic relations between behavior and language that highlight both the unique dyadic interactions children have with SLPs, and the integrated, nested systems of instructional learning environments (Chow et al. 2018). Because intervention occurs via SLP-child interaction, effective behavior management can set the stage for high-quality intervention. SLPs who are not equipped to effectively support positive behavior and reduce problem behavior may not be able to deliver as effective (quality and/or dose) as those who are prepared to address behavior. In the next section, we provide a rationale for using content area podcasts and how self-efficacy may predict response to the intervention materials.

MULTIMEDIA CONTENT AREA INSTRUCTION

One way to support the behavior management needs of SLPs is via multimedia content area instruction. Multimedia instruction has been used to improve preservice teacher knowledge and skills (Hirsch et al., 2020; Peeples et al., 2018). One specific format of multimedia instruction that has been deemed an evidence-based practice for improving preservice teacher knowledge and learning are content acquisition podcasts (CAPs; Dieker et al., 2014), which will serve as a treatment option in the present proposal. Following adult learning theory, CAPs are grounded in instructional design principles and the science of multimedia learning (Mayer, 2008). These principles are designed to reduce extraneous cognitive processing and maximize the active learning process (Mayer, 2021). However, researchers that use CAPs have not studied the effects of combining asynchronous multimedia learning (CAPs) with synchronous components of instruction for participants who may need additional support. On average, participants may benefit from CAPs more than reading articles on the same topic (Peeples et al., 2019), but to what extent do individuals for whom CAPs are not effective gain access to meaningful instruction? It may also be the case that some people may learn just as readily from reading articles or engaging in other forms of learning as with a CAP but understanding the potential role of sequencing components like articles and CAPs can be important to design and pedagogical decisions. To address this issue, we propose to use a sequential multiple-assignment randomized trial (SMART), an experimental design that has promise in education research (Chow & Hampton, 2022), to develop an adaptive intervention using interactive video modules in order to tailor instruction to better meet the needs of individual students.

SELF-EFFICACY

Teaching efficacy is a teacher's belief in his or her abilities to enact teaching practices and skills successfully and is rooted in Bandura's social cognitive theory (Bandura, 1977). For the purposes of this study, we define teaching efficacy as a teacher's judgment of their capabilities to bring about desired outcomes of student engagement and learning (Armor

et al., 1976; Bandura, 1977; Tschannen-Moran & Woolfolk Hoy, 2001). Although this construct is related to the larger concept of self efficacy, teaching efficacy is specific to the learning environment of students and to adult's beliefs about their abilities within the learning environment (Tschannen-Moran et al., 1998). Theoretical and empirical work suggest that teaching efficacy is best captured by three domains: student engagement efficacy, instructional strategies efficacy, and classroom management efficacy (Tschannen-Moran & Hoy, 2001).

Across these domains, teaching efficacy has strong implications for learning environments and significant implications for instructional behaviors, motivations, and resilience (Tschannen-Moran & Woolfolk Hoy, 2001). For example, teachers with a strong sense of efficacy are less critical of students when they make errors, are more willing to experiment with new methods to better meet the needs of their students, and are more persistent when things do not go smoothly (Ashton & Webb, 1986; Allinder, 1994; Berman et al., 1977; Guskey, 1988; Stein & Wang, 1988); this logic applies to speech-language pathologists who serve students with speech and language impairments as well as other disabilities. Additionally, teaching efficacy is positively related to higher quality and more effective classroom and behavior management skills. To illustrate, teachers who report high teaching efficacy are also more effective at planning and organizational tasks, increased use of student-centered learning, and higher-quality literacy instruction compared to teachers who report lower levels of efficacy (Justice et al., 2008). In contrast, teachers who feel insecure in their ability to manage student behaviors experience more negative or callous feelings toward their students than teachers with higher behavior management efficacy and are more likely to use reactive classroom management practices than proactive practices (Brouwers et al., 2001; Garcia-Ros et al., 2015; Wang et al., 2015). Teachers with low levels of self-efficacy are also likely to experience stress, depression, anxiety, helplessness, and pessimistic thoughts about their accomplishments and professional development (Bandura, 1997; Schwarzer & Hallum, 2008). As such, efficacy in behavior management within instructional contexts (such as speech-language therapy) may be an important facilitator of effective practice.

ADAPTIVE INTERVENTION DESIGNS

Adaptive interventions designs are multicomponent multistage interventions that adjust intervention components through progress monitoring (Almirall et al., 2018). Progress is evaluated at each *decision point*, the earliest point in intervention that a stable decision about response to initial treatment can be determined. Progress on a *tailoring variable* is measured using intervention tools designed to be collected during ongoing treatment. A *decision rule,* a set of criteria or a cut-score from the tailoring variable, is used to prescribe the next treatment option for individual participants. Taken together, an adaptive intervention is a sequence of intervention *decision points* where a *decision rule* is applied to a *tailoring variable* to determine the next intervention tactic to best optimize outcomes for each individual participant. A SMART is an efficient method used to optimize each of the key components of an adaptive intervention (Chow & Hampton, 2023; Nahum-Shani & Almirall, 2019). Within a SMART, multiple adaptive interventions are embedded within the design which are subsequently evaluated to determine the optimal embedded adaptive intervention. These embedded adaptive interventions are not the individual pathways in the SMART but encompass a combination of decision rules and associated intervention tactics at each stage. A SMART allows for experimental manipulation around the key decision points in an AI such that even one decision point that results in two categories (responders and non-responders) adds as many as eight, experimentally manipulated, embedded adaptive interventions. For example, one embedded adaptation we aim to test is the added value of intensifying the intervention for non-responders, which in our case, is adding live feedback/coaching as a second-stage treatment. These adaptive interventions can then be further optimized through evaluation of moderators at each of the key decision points in the study which results in even greater tailoring to individual characteristics (Hampton & Chow, 2022).

PURPOSE

The purpose of this registered report[1] was to conduct a pilot SMART that aimed to identify the best pathways to knowledge in behavior management for preservice SLPs. We selected behavior specific praise and visual activity schedules given their low-effort nature, and the ability to easily prepare these practices outside of instructional time and without substantially altering typical SLP intervention practice. Given that school SLPs receive minimal training on behavior management and positive behavior supports in their preservice programs while also encountering a variety of different topographies of challenging behavior (Chow et al., 2023), identifying effective and efficient methods for providing this important content to SLPs during their training is important for SLP practice and their student outcomes.

We proposed three components that represented three levels of intensity of intervention (article, video module [VM], live feedback/coaching). They also represented modalities of intervention delivery that are common practice, that can be readily adopted by practitioners and SLP preparation and continuing education programs, and the content of this intervention study is based on a recently published practitioner tutorial in a speech-language pathology journal. To empirically test the relative effects and sequences for each adaptive intervention, we used a SMART design in this context to better understand the levels of support during the learning process that SLPs need to successfully acquire adequate knowledge and self-efficacy.

RESEARCH QUESTIONS

1. Do participants gain more behavior management knowledge when assigned to first review the video module than those first assigned to read an article covering the same content?
2. Do participants who *do not* respond to the first assignment gain more behavior management knowledge when assigned to an intensified assignment as compared to those who switch to the alternative assignment?
3. Do participants who *do* respond to the first assignment gain more behavior management knowledge when assigned to add the alternative assignment than those assigned to continue with the initial assignment?
4. Which embedded intervention sequence is the most effective for optimizing participants' knowledge?
5. Does participant initial behavior management self-efficacy predict response to the first assignment?
6. Does participant time engaged with the first assignment moderate the second assignment tactic (intensify, continue, or switch)?

METHOD

Participants

We recruited participants simultaneously, from SLP programs at four universities (three large state universities, one private) in the Mid-Atlantic and Southern regions of the United States. Inclusion criteria were individuals who are currently enrolled in a graduate or advanced undergraduate training program in good standing and those willing to complete 2–3 hours of reading and/or virtual activities (within 2–3 sessions) and all survey instruments. We also included an option for focus group participation following the intervention. To recruit a sample of trainees who are new to the topic of behavior management, we excluded participants if they have completed certification in behavior management or positive behavior supports (i.e. registered behavior technician [RBT] certification, board certified behavior analyst

1 The Stage I Registered Report was accepted at *Exceptional Children* and can be accessed here: https://osf.io/ydp6k/.

[BCBA], special education teaching license). We also excluded participants who reported prior or current enrollment in a behavior management course including special education, classroom management, or autism intervention. The project's Institutional Review Board approved this research (IRB #0001028). Participants were offered an incentive to participate: $25 for completed surveys at each time point and an additional $25 for focus group participation following the intervention sequence.

Following initial recruitment, we received 88 interested potential participants. Of the 88 potential participants, 66 participants met inclusion criteria and completed the consent form. Of the 66 participants consented, 13 participants did not continue with the study following consent. The 53 remaining participants were asked to self-identify race, and results indicated that the participant group identified as the following: 64.2% White, 15.1% Hispanic, 13.2% mixed-race, 3.7% Black, and 3.7% Asian. Advanced undergraduate students (56.6% of participants) and graduate students (43.4% of participants) were eligible to participate. The mean age of the participants was 23.57 years (SD = 3.87), and women comprised 94% of the participant group. This resembles data from the American Speech-Language Hearing Association that reports 96% of SLPs are female (ASHA, 2019). All participants were speakers of English. Following pre-test assessments, two additional participants withdrew from the study due to a lack of availability.

Random Assignment

We conducted an initial group assignment using a 1:1 ratio stratified by site using a random number sequence. We recruited all participants as one cohort and randomized individuals at the same time. Only the data manager was aware of the random sequence. We repeated this process for the second randomization following the first-stage intervention; however, we stratified randomization by response-status subgroups and again by site. See Figure 1 for the present study's design. Specific randomization procedures are as follows.

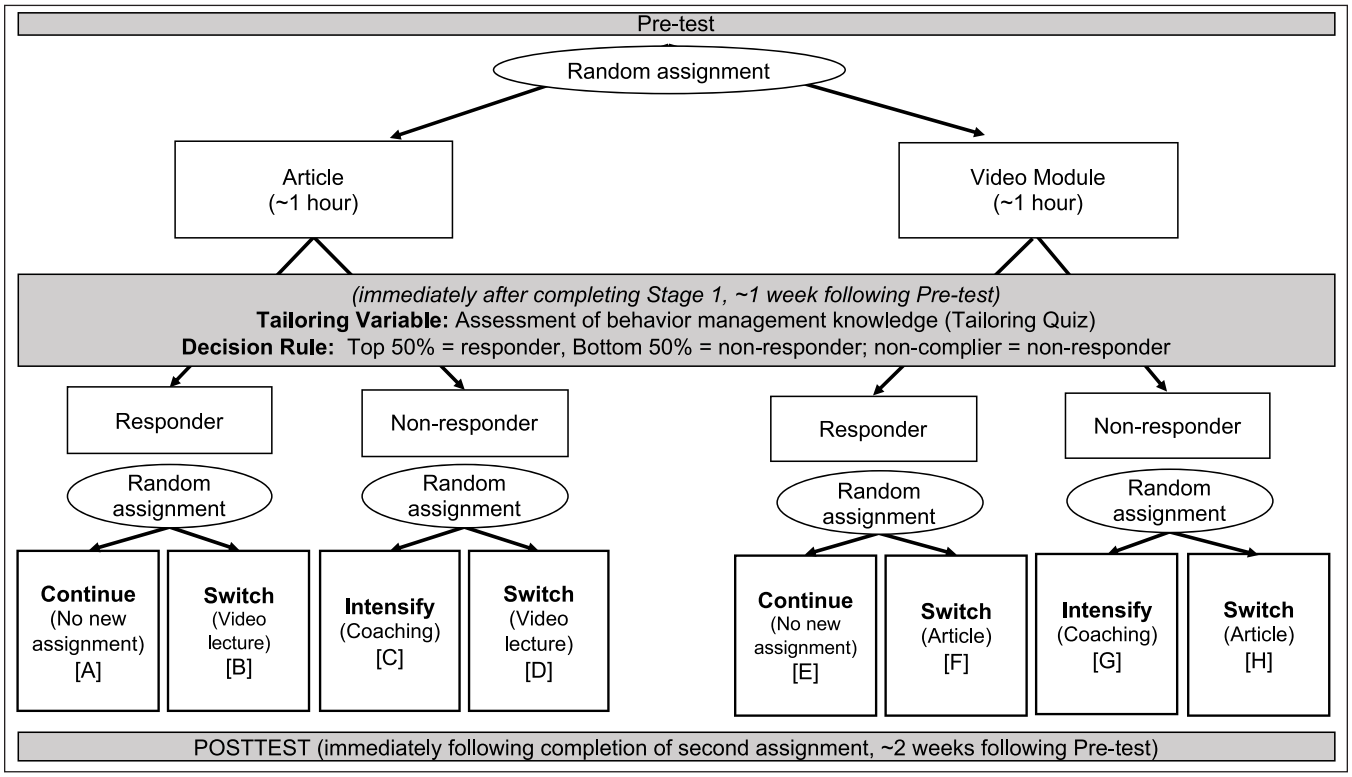

Figure 1: Study Design.

First-stage intervention options. Following completion of initial surveys, we randomly assigned participants first to (a) read a published practitioner article on behavior-specific praise and visual activity schedules (Chow et al., 2020), or (b) complete two video modules – one on behavior specific praise and one on visual activity schedules. We sent participants instructions for completing the initial surveys, confirmed they could access the materials, then gave them one week to complete the activity they were assigned to. Participants then completed a brief knowledge scale (Tailoring Quiz, see S4). We used this measure, designed to be part of the training package, to classify the participants as responders or non-responders to the initial assignment. This classification allowed us to more deeply tailor to individual needs to optimize learning. We then determined response-status categories (responder; non-responder) by identifying the midpoint of the tailoring variable distribution and used the median score as the cut point so that approximately 50% of participants will be considered responders and 50% non-responders. By forcing the cut-point to evenly distribute our participants, we hoped to evaluate the clinical utility of the measure while retaining an equal distribution of participants in each of the final cells. We randomized responders to continue or switch instructional strategies. We re-randomized non-responders to either intensify the initial assignment or switch to the other. The intensify strategy, regardless of initial instructional strategy, was 1:1 virtual coaching (described below; Fig 1, cell C). The switch strategy assigned participants to the same procedures, but alternate protocol, as stage 1 intervention. The continue strategy assigned participants to continue to receive access to the initial intervention, and no new lessons were provided.

The hour-long coaching/feedback session covers the same content as the video module in more detail with content checks embedded. More specifically, this included 10-minutes of discussion about the student's current placement context and behavior challenges, 30-minutes of viewing example videos (five, 1–3 minute videos) and discussing the strategies/opportunities for strategy use, 10-minutes to answer questions, and 10-minutes to discuss a specific example from their placement (Fig. 1, cell C & G). See Table 1 for module components that we also used to monitor intervention fidelity. Participants who are assigned to switch assignments received the alternative assignment from the first instruction (i.e. article or video module) regardless of response status (Fig. 1, cells B, D, F, and H). We re-randomized participants classified as responders to either continue (no supplemental instruction, Fig. 1, cells A and E) or switch (cells B and F). For responders, we were most interested in the added benefit of the supplemental instruction. For non-responders, we were most interested in which repair tactic optimizes outcomes for the most participants. See S1 for the video module content and the article and video module alignment. All the individual intervention components were designed to last for one hour in duration.

COMPONENT	TIME	PURPOSE	COMPONENT DESCRIPTION
Introduction	10 min	Understand context	Discuss the student's current placement, use of PBS strategies, and current challenging behaviors
What is the practice?	2 min.	Explicit instruction	Define and explain what the practice is by including examples and non-examples.
Why should I use the practice?	3 min.	Explicit instruction	Provide the evidence behind the practice and explain why this VAS/BSP is important to use during instructional delivery.
How do I use the practice?	5 min.	Explicit instruction	Explain how to use the practice (e.g., frequency and under what conditions).
What does the practice look like in action? (I do)	5 min.	Modeling	Show an exemplar video of the practice in action with real SLP or research team members. We stop the video at critical points and point out important aspects of the practice and critique the practice in use via self-reflection.

(Contd.)

COMPONENT	TIME	PURPOSE	COMPONENT DESCRIPTION
Evaluating the practice in action (We do)	5 min.	Guided practice	Show another video of the practice in action that shows the practice in use in an exemplar way at times, but at other times the video shows examples of missed opportunities or misuse of the practice. Pause and direct participants to critique the video at certain time points. After the participant has a chance to reflect and critique, stopping points throughout the video, show a video of a research team member self-reflecting on the same content.
Case Study (You do)	10 min.	Independent practice	Guide the participant through a case study where we provide descriptions of a mock intervention session, including descriptions of fictitious students. We include videos and realistic images. The participant has to complete several activities related to planning, implementation, and assessment of the use of the practice.
Placement example	10 min	Application practice	Guide the participant through a relevant example from their current placement, including descriptions of the student, the challenging behaviors, and current strategies. The participant has to complete several activities related to planning for implementing the strategies into this context.
Wrap Up	10 min.	Review	Provide a review of the session and answer questions.

Table 1: Coaching Session Overview (Intensify Condition).

Note. This table will be used as a yes/no checklist to monitor intervention fidelity.

MEASURES

This study includes (1) the primary outcome measure, (2) two tailoring variables, one that determines response status and one moderator, and (3) a measure of social validity. We also collected participant demographics (e.g., gender, race/ethnicity, age, undergraduate major, current degree program) to assess pretest equivalence. See supplementary materials for the pre-posttest assessment, tailoring variable, and self-efficacy measure.

Pre and Posttest Knowledge Test

The primary outcome measure is aligned both with the content of the article and video modules (where participants are first randomized to either condition) and previously used behavior knowledge assessments that have demonstrated reliability (see Kennedy et al., 2016; Hirsch et al., 2020). This Knowledge Test measure includes knowledge items specific to the content as well as application items that included video-based application questions. Participants observed one- to three-minute video segments of SLP sessions and answered questions about the use or opportunities to use the target strategies. Final measure consisted of 25 items scored on a weighted scale of 0–2 points each. See supplementary materials for all items. We developed a scoring manual for each of the items in the assessment. Raters trained using practice assessments and the first 10 participants' data, reaching 90% point-by-point agreement. Two independent raters, masked to timepoint, independently scored each of the Knowledge Tests. Raters then met to resolve any discrepancies on scoring to ensure agreement. See S2 for this measure.

Tailoring Variables

Response status. We used a 20-item Tailoring Quiz (tailoring variable, see S4) to determine response status to the first-stage intervention. We designed this scale to quickly assess knowledge acquired from the article or the video modules. First, for the purposes of this study, we categorized participants who did not complete the Tailoring Quiz as non-

responders. Then, to determine response status, we used the distribution of scores at the decision point and categorized 37% (n = 19; score of 17 or higher) as responders and 63% (n = 32; score of 16 or lower) as non-responders.

Behavior management self-efficacy. We created a behavior management efficacy scale as our planned baseline tailoring variable, drawing from previous measures of classroom management efficacy (Dellinger et al., 2008; Emmer & Hickman, 1991). We adapted a classroom self-efficacy scale for SLPs because no self-efficacy scale for SLPs includes a behavior management. This planned baseline moderator variable would determine if behavior management efficacy predicted response to each embedded adaptive intervention. See S3.

Time engaged. We assessed whether participant time engaged during the first stage treatment (article or video modules) predicted response to intervention posttest outcomes. This variable served as a proxy for a participant's time spent on/engagement in the first stage treatment. Conceptually, we aimed to determine if time engaged in the first stage treatment matters for how participants respond to the second stage of the intervention. We measured time engaged using a self-report survey question of the number of minutes engaged at the end of the assignment in addition to engagement statistics generated by the intervention platform based on the number of minutes viewed. We asked students to report time engaged directly with the material and total time engaged in the content including self-reflection, discussion with others, or additional research.

Social Validity

We originally intended to assess social validity using an adapted version of the System Usability Scale and Questionnaire for User Satisfaction (Delamarre et al., 2021). This assessment was developed for virtual training (Chin et al., 1988), and more recently adapted for virtual teacher training with strong internal reliability (Bangor et al., 2008; Cronbach's alpha > .90). The aim was to pair these results with the focus group at posttest, as recommended by Delamare et al. (2021). However, given the small final sample of participants who completed project activities, we elected to conduct a more intentional feedback survey and a focus group that targeted feasibility, attrition, and recommendations for future studies.

Data Analysis

We collected all survey data in REDCap (Harris et al., 2019), a secure, web-based application for data entry, and conducted all analyses in RStudio running R version 4.3.1 (RStudio Team, 2020; R Core Team, 2023). We examined descriptive data of our measures prior to all analyses. We identified outliers using modified z-score analyses (Iglewicz & Hoaglin, 1993) and we used a winsorized mean to adjust for these outliers. This occurred only for the minutes engaged data where some extremely long times were extracted from the learning module. We assumed no clustering issues given we stratified at the university level and our assessment of distributions for each dependent variable confirmed no meaningful differences between subgroups at baseline. We controlled for Type 1 errors at posttest using the Benjamani-Hochberg procedure (Benjamini & Hochberg, 1995). We tested for any group differences on demographics and determined that it was unnecessary to include any covariates in the final models. While we aimed to complete an intent-to-treat analysis, attrition and missingness limited our ability to do so. In response to the observed high levels of attrition, we determined that all missing data would be excluded from the final analyses (rather than imputing missing data) because 1) we did not have sufficient baseline data for a robust imputation, 2) we observed patterns in the missingness that indicated the missingness was not at random, and 3) it was relevant to analyze and interpret the results from those who fully participated and consider this compared to those who choose not to participate throughout the study. One participant was excluded from the analysis because they consented to the study but failed to complete more than 20% of the initial surveys. All other participants were retained for each timepoint for which they contributed data. We expected low levels of attrition and missing data, due to incentives and the short intervention period. Due to non-equivalence and higher-than-tolerable levels of differential attrition at both randomization points, we adjusted the proposed analyses as described above.

Research Question 1

We used a generalized linear regression model to answer the first research question about behavior management knowledge based on first stage treatment (Lei et al., 2012). Group membership was entered into the model using the first randomization (0 = Article, 1 = Video Module) and used to predict outcome scores on the behavior management knowledge assessment. Raw pre-test scores on the behavior management knowledge assessments were entered as a covariate.

Research Questions 2–4

We did not use a generalized linear regression to answer the second and third research questions about behavior management knowledge based on second stage randomization as planned (Lei et al., 2012). This was due to missingness and small participant numbers in the final comparison groups. More complete data are required to test differences between treatment response status (RQs 2 and 3) or to determine the best sequence of intervention components (i.e., the best embedded adaptive intervention; RQ4).

Research Questions 5–6

Moderator analyses were planned to mimic the main effect analysis with an added predictor using the proposed continuous moderator variable. Research question 5 analysis would use the preservice SLP's self-efficacy score as a predictor of outcome knowledge between initial randomization groups, averaging across all second stage assignments. Research question 6 would use the participant's assignment 1 time engaged as a continuous predictor of best second assignment tactic (continue, switch, intensify), averaging across all first stage assignments. Though we intended to answer research question 5 via two separate analyses for each response status (Nahum-Shani et al., 2012; 2019), we did not have enough data to run these analyses separately.

Power. Using a SMART power calculator (Kim et al., 2016) and assuming approximately 50% participants will be responders at the decision point, we estimated 68 participants were necessary to evaluate the initial effects and to ensure at least six participants fall in each final cell. We used effect size estimates from previous studies that tested the efficacy of our initial treatment stage (article vs. video module; see Kennedy & Thomas, 2012; d = .86 to 1.62). We did not compare individual cells; all analyses considered the primary or secondary randomization subgroups. Assuming power of 85% and a pretest covariate, and assuming .05 attrition during stage 1, an effect of .86 or greater (i.e., a minimal detectable effect or MDE) is detectable for $p < .05$ for the effect of the initial assignment to first-stage treatment, which is our main effect. The MDE for the best tactical decision for responders, is 1.72 assuming approximately 50% are responders following Stage 1 and assuming another .05 attrition. Though the MDE for the primary randomization is .86, previous studies that have compared effects of a reading/notetaking condition to a video module condition report an overall ES of d = .98 (see Kennendy & Thomas, 2012).

RESULTS

We present the results of the analyses that we were able to conduct given high levels of attrition and contextualize the broader findings to include issues of implementation of SMARTs in higher education contexts. Note that these analyses were conceptualized and conducted prior to any data collection. In the results, we present the analyses as planned, but readers should interpret the results as underpowered. For example, we were powered (a priori) to detect primary effects with a sample of 68 participants; out initial sample size in the final analysis was 53. We did not conduct post-hoc power analyses.

Sample

Of the total included sample of 53 participants, 27 were randomized to receive the article as the initial treatment, and 26 were randomized to receive the module as the initial treatment. After the initial treatment stage, 29 participants completed the Tailoring Quiz and were re-randomized. Of the initial 53 participants, we deemed 22 "non-complaint" given that they did not complete the Tailoring Quiz, which is required to determine response status. We categorize them as non-responders for the purposes of re-randomization but also refer to this group as non-compliant participants henceforth to distinguish this group from the participants who demonstrated relatively less knowledge following the first-stage intervention. Because these participants represent a large proportion of the non-responder samples, we report their results separately at times to distinguish the differences between those who struggled to learn the content and those who struggle to stay engaged in the study.

Initial Randomization

After controlling for pretest knowledge scores, we found no differences on the outcome knowledge assessment between the participants that were initially randomized to the module condition compared to the article condition (p = .901). We had initially planned to continue with comparisons of responders and non-responders to each intervention tactic (e.g., switch; intensify); however, too few participants remained in the study for this to be a viable statistical analysis.

Engagement

We explored whether minutes of engagement in the content predicted outcome knowledge scores regardless of group. After controlling for pretest knowledge scores and using winsorized means (to account for extremely long times of engagement possibly due to a participant leaving the module running), minutes of engagement in the content, as indexed by the time spent actively participating in the online learning platform, did not significantly predict outcome scores (p = .342).

Self-Efficacy

We also explored whether self-efficacy predicted knowledge outcome. In this sample, self-efficacy did not significantly predict the outcome. As a predictor or a moderator, there was no significant relation between self-efficacy and knowledge. All data and analyses can be found at this link: https://osf.io/ydp6k/.

Feedback Survey

In response to our high attrition and to understand participant perceptions of the intervention, we created an open-ended feedback survey that asked the following questions:

1) What were the primary barriers, if any, to your participation in the study?
2) Would you recommend this study to a friend/peer? Why or why not?
3) What is one aspect of the content presented in the material that you anticipate will be useful to your future practice?

We sent the survey out to all participants. Of the 30 that initiated the survey, 27 completed all items. In terms of primary barriers, 14 (52%) identified time and their busy schedule as a barrier, though five responses mentioned time but that it was a minimal barrier/minor issue. Three participants indicated an issue with preference (e.g., they would rather watch a video than read an article; having to learn the materials on their own); three reported logistical issues (e.g., the multiple emails, logging into the system); two reported a lack of experience (e.g., no experience with kids; being undergrad with less knowledge). Five participants explicitly reported no barriers to participation.

All 27 respondents indicated that they would recommend the study to a friend/peer. Reasons included usefulness in practice, that they learned a lot, the information was highly valuable, and the content was not present in any of their classes in their programs. Of the 27, 20 responses were general recommendations for the content, statements of usefulness and positive impacts on their practice, and perceptions of usefulness for other practitioners. Seven respondents specifically mentioned how the study deepened their knowledge and/or mentioned the usefulness of the explicit instruction. Two participants also mentioned the compensation being beneficial. When asked about one aspect of content they anticipate would be useful to their future practice, 11 (41%) identified VAS specifically, 8 (30%) identified BSP specifically, 7 (26%) identified both specifically, and one identified having a script of the video as useful.

Post-Intervention Focus Group

To better understand participation issues and to gain an element of confirmation for this reasoning, we conducted an optional focus group with participants who remained in the study through its completion; those that opted in received an additional $25 compensation. Five participants consented to a virtual focus group that lasted approximately one hour. The first author conducted the focus group, and a graduate research assistant took focus group notes. We conducted the focus group in a systematic process, reviewing the purpose of the focus group and the process we would engage in to anonymize the data prior to qualitative coding of the focus group notes. Given this was not a primary research aim of the study, we did not transcribe the focus group verbatim. Instead, a graduate assistant took detailed notes. Then, two team members independently read the notes and extracted initial themes by each focus group question. Then, the team members met to review themes.

To summarize, the primary themes that emerged were: (1) learning about behavior specific praise was useful and seen to be broadly applicable outside of the study context, (2) overall the study content (e.g., article, modules) were easy to understand, (3) timing of the study was difficult, and (4) program-based incentives would have likely increased participation and reduced attrition.

More descriptively, the focus group participants found both the modules and the articles helpful, and the participants who read the article either did prefer (1 student) or think they would have preferred the module (2 participants). When asked about what they liked least about the study, or how we could improve the procedures, participants felt like some of the deadlines were not clear, and that an advanced organizer would have been helpful, so they had a better idea of the full scope of the study. In this version of the study, other than in the initial email, we did not clearly map out the full study for participants and multiple points at each stage of the study but alluded to future content that was based on randomization and performance on quizzes throughout. The most relevant issue for our study, our future planning, and next steps to avoid similar levels of attrition was the timing of the study. Two participants agreed that the pacing of the study was good (the time given for each of the components), and all the five participants agreed that starting the study earlier in the semester would have likely led to fewer people dropping out. The two reasons that stood out for the five participants for why they remained in the study were the incentives ($75 each participant) and their interest in the content. Notably, the participants who were in the focus group all had career goals to work in schools or were interested in how working with children and/or children with disabilities might inform their interest in working with adult populations. Participants suggested making the study required as a part of the SLP program(s).

DISCUSSION

Using a SMART, we aimed to identify the best embedded adaptive intervention for producing favorable learning outcomes for preservice SLPs in behavior management knowledge and skills. Though we carried out the study protocol as designed in Stage 1 of this registered report, several issues did not allow us to conduct all the proposed analyses as intended. Due to the nature of registered reports and open science practices, we conducted the analyses where we

had sufficient data to run the appropriate analyses, and adapted some elements the study to (1) better understand the scope of the attrition that ultimately limited the data analyses, and (2) to plan for future iterations of SLP professional development studies that can better account for factors we were unaware of or were out of our control when conducting the study as designed.

One major issue we encountered was attrition. Given the individualized design of our SMART and the timing of each of the assessment and intervention windows that provided windows for participants to complete the content (module, article), participants varied in the time they took for each of the components. For most participants, they waited until the end of each window to complete their tasks for that stage of the study, and this pushed the timeline further into the academic semester. At the end of the semester, close to finals and high-stakes assignments in their coursework, many participants did not complete the study.

Another implementation factor to consider is that the entire intervention, other than recruitment, was conducted virtually. Contact was made and information disseminated via email. There was no room to build personal relationships, develop rapport, or install components of accountability that may have increased buy in and participation. If conducted in the context of a course with an instructor who was invested in the content and student participation, it may have been a more conducive environment for the intervention for buy-in, but also for communicating clearly the steps of the intervention and what components were to be completed.

This study, including implementation failures, can inform future studies using SMART designs in preservice programs. Though we implemented the study as designed, the response to our planned study did not occur as assumed. We learned valuable lessons from the focus groups and our assumptions about timing and recruitment. Broadly, earlier in the semester or a more sensitive approach to the timing of the intervention would have eased implementation. Working directly with program faculty to make a part of the course/clinical experience, to require elements, or to provide extra credit, or even time within the program to complete the content would have likely increased engagement in the study. We had initial buy-in from the SLP programs, which is a nod to the likely adoption of the content given a more feasible, aligned delivery approach. Our survey confirmed that timing was the primary barrier. Though, the survey also revealed high levels of social validity in the study content and the direct implications for SLP practice. Participants specifically identified VAS and BSP as informative to their practice and recommended it to peers.

We learned about issues of resources and what it takes to carry out the basic SMART procedures and second randomization. The re-randomization procedures require careful consideration and communication with the research participants. Participants who may receive more or less intensive interventions must be prepared for all possible outcomes, which requires planful timing of assessments and intervention phases. In our current study the second stage intervention occurred proximal to final exams for the students which may have contributed to high rates of attrition. Pilot SMARTs are good for optimizing intervention components, yet the nuance of each varying component adds complexity to the study which may complicate enrollment and engagement in the study procedures. However, one benefit of the current study is that we identified a meaningful and potentially useful cut-score on the tailoring variable for use in a future SMART. With a preconceived cut-point, enrollment could be rolling and therefore individualized to specific participant needs. Now we have performance on the tailoring variable to make data-based cut score decisions. Future research will benefit from further specification of the tailoring variable in different contexts, populations, and timelines. In this study, we designed assessments like we would have in other randomized experiments. What this study design allowed us to realize is that variability of assessments, particularly those within the causal pathway of the intervention and corresponding to within-intervention decision making (e.g., decision point tailoring variable) is especially important to differentiate the sample into response categories. Further piloting of the measures, with more attunement to how each specific item would perform because of actual learning from the content is an important next step for assessments within SMART designs.

LESSONS LEARNED AND RECOMMENDATIONS FOR FUTURE SMARTS

We conducted this study with the intention of providing initial evidence that would support SLP training and curricular decision making. We encountered several implementation issues that researchers should consider if planning a SMART design or other adaptive intervention design. We view the primary contribution of this study as an assessment of the feasibility of SMARTs within preservice preparation programs, as well as the implications of the registered report process on challenging experimental studies, including adaptive intervention designs. We conclude this study with some observations about both registered reports and SMARTs and remain optimistic about the potential of SMARTs in the preservice educator arena.

First and relative to the registered report process, the purpose is important and conceptually sound in its link to improved science and transparency. However, the time it takes and the unpredictability to successfully execute and disseminate the product in the context of applied, experimental research and the peer-review process may outweigh the potential scientific and procedural benefits in some cases. For instance, we began this study in 2020 and are now able to distribute the final report five years later. This process included an accepted Stage I Registered Report at one journal and completion of the process in another journal that included an Editorial Team change at the initial journal. Procedurally, implementing the actual mechanics of a SMART is challenging and at times cumbersome (e.g., assessment and scoring of the tailoring variable to produce data for response status generation and re-randomization). We have hopes that these processes can be streamlined, including their fidelity and accuracy, with machine learning models.

Substantively, future SMART study can benefit from the challenges and lessons learned in the current study. The changing nature of the intervention may require additional support for participants including some face-to-face check-ins, regular delivery of incentive payments, and regular interviews with participants about their ongoing experience rather than one final conversation at the end of the study. Further, intervention-as-received by the participant varies greatly in a self-led intervention such as a video module or article reading. Future students can greatly benefit from more nuanced engagement metrics to understand the extent to which participants are fully engaging in the intervention materials. While the software that we used in this study allowed us to estimate the time that the participant had the materials open in a browser, actual attending/engagement is difficult to measure. Some participants reached the maximum time allotment (three hours) which far exceeded our engagement expectations. It is likely that these participants left the window open and running with little engagement for at least a portion of the three-hour window. Improved engagement monitoring could include requiring participants to engage in the modules in a lab space where observers can rate their engagement, regular quizzes embedded in the module to check for attention, or more nuanced estimates of engagement based on active window and active computer activity checks.

A strength of the current study is that we enrolled a full cohort at one time and did not use rolling enrollment. While rolling enrollment has some benefits to individualizing intervention timing for participant needs, a cohort model allowed us to 1) examine the effects with the same external pressures for all students (exams, etc.), 2) randomize the participants with stratification based on actual enrollment rather than planned enrollment tables which may over or underestimate true enrollment in any given block, and 3) create a cut-score for the tailoring variable based on the actual responses received to ensure equal distribution of participation across groups. However, rolling enrollment would allow researchers to react to high attrition by over-enrolling, allow researchers to respond to individual participant scheduling needs, and provide more individualized attention to participants during study shifts. Future researchers would benefit from careful consideration of the benefits and risks of each type of enrollment based on their specific study context and needs.

SUPPLEMENTARY FILES

S1. Article and CAP Alignment. https://doi.org/10.25894/rise.2834.s1

S2. Pre/Posttest. https://doi.org/10.25894/rise.2834.s2

S3. Self-Efficacy Tailoring Variable. https://doi.org/10.25894/rise.2834.s3

S4. Tailoring Variable (20 questions). https://doi.org/10.25894/rise.2834.s4

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Jason Chow: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft, Writing – Review & Editing; **Lauren Hampton**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Writing – Original Draft; Writing – Review & Editing.

AUTHOR AFFILIATIONS

Jason C. Chow orcid.org/0000-0003-1967-9714
Vanderbilt University, US

Lauren H. Hampton orcid.org/0000-0003-1967-9714
University of Texas at Austin, US

REFERENCES

Almirall, D., Kasari, C., McCaffrey, D. F., & Nahum-Shani, I. (2018). Developing optimized adaptive interventions in education. *Journal of Research on Educational Effectiveness*, *11*(1), 27–34.

American Speech-language Hearing Association. (2019). https://leader.pubs.asha.org/doi/10.1044/leader.AAG.24072019.32

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction*, *24*(6), 574–594.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Bronfenbrenner, U. (1994). Ecological models of human development. *International Encyclopedia of Education*, *3*(2), 1643–1647.

Chow, J. C. (2018). Comorbid language and behavior problems: Development, frameworks, and intervention. *School Psychology Quarterly*, *33*(3), 356–360.

Chow, J. C., Cunningham, J., & Wallace, E. (2020). Interaction-centered model for language and behavioral development. *Handbook of research on emotional & behavioral disabilities: interdisciplinary developmental perspectives on children and youth.* Routledge. https://doi.org/10.4324/9780429453106

Chow, J. C., Ekholm, E., & Coleman, H. (2018). Does oral language underpin the development of later behavior problems? A longitudinal meta-analysis. *School Psychology Quarterly*, *33*(3), 337.

Chow, J. C., & Hampton, L. H. (2019). Sequential multiple-assignment randomized trials: Developing and evaluating adaptive interventions in special education. *Remedial and Special Education*, *40*(5), 267–276.

Chow, J. C., & Hampton, L. H. (2022). A systematic review of sequential multiple-assignment randomized trials in educational research. *Educational Psychology Review*, *34*(3), 1343–1369.

Chow, J. C., Senter, R., & Morse, A. (2023). Speech-language pathologists' experiences with challenging behavior in practice. *Communication Disorders Quarterly*, *44*(2), 117–126.

Chow, J. C., & Wallace, E. S. (2021). Speech-language pathologists' behavior management training and reported experiences with challenging behavior. *Communication Disorders Quarterly*, *42*(2), 67–72.

Chow, J. C., & Wehby, J. H. (2018). Associations between language and problem behavior: A systematic review and correlational meta-analysis. *Educational Psychology Review*, *30*(1), 61–82.

Chow, J. C., & Wehby, J. H. (2019). Profiles of problem behavior in children with varying language ability. *Journal of Emotional and Behavioral Disorders*, *27*(2), 110–118.

Chow, J. C., Zimmerman, K. N., & Senter, R. (2020). Tailoring effective behavior management strategies for speech-language pathologists. *Language, Speech, and Hearing Services in Schools*, 1–13.

Cunningham, J. E., Chow, J. C., Meeker, K. A., Taylor, A., Hemmeter, M. L., & Kaiser, A. P. (2021). A conceptual model for a blended intervention approach to support early language and prosocial development. *Infants and Young Children, 36*(1), 53–73. https://journals.lww.com/iycjournal/fulltext/2023/01000/A_Conceptual_Model_for_a_Blended_Intervention.5.aspx?context=LatestArticles&casa_token=B_rEpTUHrAoAAAAA:rQnFDJK2sBsO9F6Q6gy7htCpkihAzgs5dzUk13bIvaYqTaLeShjpjLStfbeZ1actD7_u6kkM9iAoApBhbRH64Agp

Curtis, P. R., Frey, J. R., Watson, C. D., Hampton, L. H., & Roberts, M. Y. (2018). Language disorders and problem behaviors: A meta-analysis. *Pediatrics*, *142*(2), e2017351. https://publications.aap.org/pediatrics/article-abstract/142/2/e20173551/37527/Language-Disorders-and-Problem-Behaviors-A-Meta

Delamarre, A., Shernoff, E., Buche, C., Frazier, S., Gabbard, J., & Lisetti, C. (2021). The Interactive Virtual Training for Teachers (IVT-T) to practice classroom behavior management. *International Journal of Human-Computer Studies, 152*, 102646.

Dieker, L. A., Kennedy, M. J., Smith, S., Vasquez III, E., Rock, M., & Thomas, C. (2014). Use of technology in the preparation of pre-service teachers. *Document No. IC-11*.

Flower, A., McKenna, J. W., & Haring, C. D. (2017). Behavior and classroom management: Are teacher preparation programs really preparing our teachers? *Preventing School Failure: Alternative Education for Children and Youth*, *61*(2), 163–169.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, 8(3), 206–213. https://doi.org/10.1007/s11121-007-0070-9

Greenberg, J., Putman, H., & Walsh, K. (2014). Training our future teachers: Classroom management. Revised. National Council on Teacher Quality.

Hampton, L. H., & Chow, J. C. (2022). Deeply tailoring adaptive interventions: Enhancing knowledge generation of SMARTs in special education. *Remedial and Special Education*, *43*(3), 195–205.

Hampton, L. H., Chow, J. C., Bhat, B. H., & Roberts, G. (2023). Reporting and design considerations for SMART behavioral science research. *Educational Psychology Review*, *35*(4), 117.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., ... & REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208.

Hemmeter, M. L., Fox, L., & Ostrosky, M. (2021). (Eds). *Unpacking the Pyramid Model: A Practical Guide for Preschool Teachers.* Brookes Publishing. https://products.brookespublishing.com/Unpacking-the-Pyramid-Model-P1231.aspx

Hirsch, S. E., Chow, J. C., Randall, K. N., Nemer, S. L., & McKown, G. (2020). Evaluating the effect of embedded responses in multimedia-based instruction with preservice teachers. *Behavioral Disorders*, *46*(1), 18–28.

Hollo, A., Chow, J. C., & Wehby, J. H. (2019). Profiles of language and behavior in students with emotional disturbance. *Behavioral Disorders*, *44*(4), 195–204.

Hollo, A., Wehby, J. H., & Oliver, R. M. (2014). Unidentified language deficits in children with emotional and behavioral disorders: A meta-analysis. *Exceptional children*, *80*(2), 169–186.

Iglewicz, B., & Hoaglin, D. C. (1993). *How to detect and handle outliers* (Vol. 16). Asq Press.

Kennedy, M. J., & Thomas, C. N. (2012). Effects of content acquisition podcasts to develop preservice teachers' knowledge of positive behavioral interventions and supports. *Exceptionality*, *20*(1), 1–19.

Kim, H., Ionides, E., & Almirall, D. (2016). A sample size calculator for SMART pilot studies, *SIAM Undergraduate Research Online, 9*, 229–250.

Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., & Murphy, S. A. (2012). A" SMART" design for building individualized treatment sequences. *Annual Review of Clinical Psychology, 8*, 21–48.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, *63*(8), 760–769.

Mayer, R. E. (2021). Evidence-Based Principles for How to Design Effective Instructional Videos. *Journal of Applied Research in Memory and Cognition, 10*(2), 229–240.

McLeod, B. D., Sutherland, K. S., Martinez, R. G., Conroy, M. A., Snyder, P. A., & Southam-Gerow, M. A. (2017). Identifying common practice elements to improve social, emotional, and behavioral outcomes of young children in early childhood classrooms. *Prevention Science*, *18*(2), 204–213.

Nahum-Shani, I., & Almirall, D. (2019). An Introduction to Adaptive Interventions and SMART Designs in Education. NCSER 2020–001. *National Center for Special Education Research*.

Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W. E., Gnagy, B., Fabiano, G. A., ... & Murphy, S. A. (2012). Q-learning: A data analysis method for constructing adaptive interventions. *Psychological Methods, 17*(4), 478.

Oliver, R. M., & Reschly, D. J. (2010). Special education teacher preparation in classroom management: Implications for students with emotional and behavioral disorders. *Behavioral Disorders*, *35*(3), 188–199.

Peeples, K. N., Hirsch, S. E., Gardner, S. J., Keeley, R. G., Sherrow, B. L., McKenzie, J. M., ... & Kennedy, M. J. (2019). Using multimedia instruction and performance feedback to improve preservice teachers' vocabulary instruction. *Teacher Education and Special Education*, *42*(3), 227–245.

Poulou, M. (2007). Personal teaching efficacy and its sources: Student teachers' perceptions. *Educational Psychology*, *27*(2), 191–218.

Rubin, D. B. (2004). Multiple imputation for nonresponse in surveys (Vol. 81). John Wiley.

Sameroff, A. (2009). *The transactional model*. American Psychological Association.

Shortreed, S. M., Laber, E., Scott Stroup, T., Pineau, J., & Murphy, S. A. (2014). A multiple imputation strategy for sequential multiple assignment randomized trials. *Statistics in Medicine*, *33*(24), 4202–4214.

Sutherland, K. S., Conroy, M., Abrams, L., & Vo, A. (2010). Improving interactions between teachers and young children with problem behavior: A strengths-based approach. *Exceptionality*, *18*(2), 70–81.

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry*, *54*(5), 516–524.