

Measuring What Matters: A Generalized Structural Equation Modeling Approach to Examine Teacher Self-Efficacy and School Efficacy for Students With Complex Support Needs

Research Overview:

Building on national survey data from 622 U.S. educators, this study employed Generalized Structural Equation Modeling (GSEM) to examine how School-wide PBIS (SWPBIS) implementation levels and students' educational placement shape two key constructs: teacher self-efficacy and school-level implementation of behavioral supports for Students with Complex Support Needs (SCSN).

Purpose:

Multi-tiered behavioral support frameworks are increasingly positioned as universal solutions across educational systems. In the United States, Positive Behavioral Interventions and Supports (PBIS) is among the most widely adopted of these frameworks designed to serve all students, including those with the most significant disabilities or complex support needs (CSN). Yet Students with CSN (SCSN) remain largely excluded from meaningful PBIS participation, despite the framework's explicit commitment to serve everyone. This phenomenon is not unique to the US; internationally, students with the most intensive needs are routinely last to benefit from universal intervention frameworks, raising concerns for inclusive education policy and practice worldwide.

The current study builds on two prior studies (Kurth et al., 2024; Zagona et al., 2025) that analyzed survey responses from 622 U.S. PreK–12 educators. The studies found that knowledge, beliefs, and practices related to CSN involvement in PBIS varied significantly based on two factors: the level of PBIS implementation at a school and the educational placement of SCSN. One major weakness of these pre-existing results, however, was in their reliance on item-by-item analysis, which prevented them from drawing strong conclusions around interdependency of underlying constructs and what factors may impact them.

Methods:

The current study addresses this limitation by applying a Generalized Structural Equation Modeling (GSEM) framework to the same national dataset, and simultaneously estimates two latent constructs using Item Response Theory (IRT)-based measurement models. Teacher Self-Efficacy (Construct 1) consists of 23 five-point Likert-scale items, and is modeled using a Graded Response Model (GRM). SWPBIS Efficacy for SCSN (Construct 2) consists of 22 binary items and is modeled using a 2-Parameter Logistic (2-PL) IRT model. Stata 18 was used for all aspects of the analysis, including the verification of unidimensionality, and evaluation for potential presence of Differential Item Functioning (DIF). The final model can be characterized as a MIMIC model, with covariates such as educators' race/ethnicity, roles, and school grade

level taken into account along with our primary predictors of interest: PBIS implementation level (0–3 tiers), and educational placement of SCSN.

Results and Discussion:

The final model reveals a theoretically coherent but empirically differentiated pattern across the two constructs. For Teacher Self-Efficacy (Construct 1), only educators in schools implementing all three PBIS tiers demonstrated significantly higher self-efficacy for supporting SCSN ($\beta = .204$, $SE = .087$, $p = .019$). Schools implementing one or two tiers showed no significant difference from non-implementing schools. This threshold effect is important: it suggests that partial implementation — which describes the majority of PBIS schools — does not generate the sustained educator confidence needed to support SCSN. For international systems considering phased rollouts of tiered behavioral frameworks, this raises a pointed question about what minimum conditions are actually needed to shift educator capacity.

For SWPBIS Efficacy for SCSN (Construct 2), the pattern differed. Both one-tier ($\beta = .265$, $p = .007$) and three-tier ($\beta = .806$, $p < .001$) implementations were associated with significantly higher fidelity, with three-tier implementation showing a substantially larger effect. Schools with full implementation demonstrated meaningfully stronger fidelity in the practices most relevant to SCSN, including behavioral data collection, accessible posting of expectations, and inclusive acknowledgment systems.

DIF findings added an important equity dimension. For Teacher Self-Efficacy (Construct 1), African American educators showed significantly lower self-efficacy at equivalent latent trait levels ($\beta = -.157$, $p = .018$), even after controlling for school context — likely reflecting structural inequities in access to training and professional networks rather than individual differences. For SWPBIS Efficacy for SCSN (Construct 2), Hispanic/Latinx educators ($\beta = -.230$, $p = .044$) and high school educators ($\beta = -.224$, $p = .006$) reported lower school-level fidelity, pointing to well-documented secondary-level implementation gaps and raising equity concerns about differential access to structured behavioral supports in racially and linguistically diverse school communities.

Together, these findings advance the field methodologically and substantively. The shift from descriptive statistics and item-by-item survey analysis to validated latent variable modeling reveals patterns — particularly the self-efficacy threshold effect — that prior work could not detect. More broadly, these results offer a caution for education systems globally: implementing a universal framework does not guarantee equitable access for students with the most significant needs. Who actually benefits may be contingent, in part, on implementation depth, educator background, and school-level structural conditions. These are challenges shared across national contexts, and this study offers both a methodological template and preliminary empirical findings relevant to that global conversation.