Evaluation of *We Have Skills!*
A Multimedia Classroom Level Social Skills Program for Elementary Students

Grant Award Number: R305A150046

Methods and Data Documentation

IRIS Educational Media (dba Trifoia)

& Oregon Research Institute

Investigators

Adam Wendt, Project Director, IRIS Educational Media, Inc.

Keith Smolkowski, Principal Investigator, Oregon Research Institute

February 17, 2022

Table of Contents

Purpose of the Documentation 3
Project Description and Research Design Overview 3
Participants 4
Intervention 4
  Curriculum Description 4
Procedures 4
Measures 4
  Elementary Social Behavior Assessment 5
  Walker–McConnell Scale of Social Competence 5
  Teacher Sense of Efficacy Scale 5
  Teacher Concerns Inventory 6
  Implementation Fidelity 6
  Direct Observations 6
Data Structure 7
References 9

Purpose of the Documentation

This document describes methods used for *Evaluation of We Have Skills! A Multimedia Classroom Level Social Skills Program for Elementary Students* funded by Institute of Education Sciences (IES) to IRIS Media, Inc. (Grant Award Number: R305A150046). This report supplements the published papers with additional details, such as variable names for measures included in archived data. This document may repeat text from manuscript without quotation marks or attribution from Smolkowski et al. (in press).

Project Description and Research Design Overview

The proposal to the Institute of Education Sciences described this project as a four-year evaluation of the efficacy of *We Have Skills!* (WHS) for developing student social and academic behaviors compared to standard instruction typically provided in elementary school programs (comparison). The study was originally designed to address the following research questions:

1. Does WHS lead to greater gains in 1st grade students' classroom behavior skills and in academic performance, as well as their teachers' sense of efficacy and effectiveness compared to social skills instruction routinely delivered in 1st grade classrooms? Do students' gains in classroom behavior skills and academic performance maintain into 2nd grade?
2. Do (a) student characteristics, such as risk of behavioral and academic problems, English language learner, IEP status, and socioeconomic status and (b) teacher characteristics, such as length of teaching experience and education level, moderate the relationship between WHS and student and teacher outcomes?

During the course of the project, and for the impact manuscript, we refined the research questions based on three considerations. First, few participating schools provided academic information or data for students after they transitioned into second grade. Second, the original research questions included too many tests of moderation, so we focused our analysis of moderation on only students' initial classroom adjustment. We also found little rationale for other moderation tests upon a review of the literature, and some potential moderators were unavailable from most schools (e.g., economic disadvantage). Finally, the original research questions did not include questions about observations of teacher behavior.

The impact manuscript focused on these three research questions:

1. Relative to students in control classrooms, students in classrooms that use WHS will demonstrate improved academically related behavior and classroom adjustment.
2. As some students arrive at school versed in behavioral skills, baseline levels of student behavior will moderate treatment-group differences in student outcomes.
3. Teachers implementing WHS will increase their descriptive praise and corrections of skills taught by WHS, report fewer concerns about student behavior, and report greater self-efficacy for classroom management than control teachers.

The study also included a detailed cost and cost-effectiveness analysis. The cost-effectiveness analysis was not included in the original project description.

This study took place from 2015 to 2019 with early elementary school teachers and their students in 60 schools from multiple states (AZ, CA, CO, NV, OH, OR, NM, UT, WA).

Investigators randomly assigned teachers to either an immediate-intervention condition or a wait-list, business-as-usual control condition in which no additional intervention or services were offered. Random assignment occurred within each school, when possible, district, state, and wave (cohort) and to experimentally control for extraneous influences, such as funding levels, proportion of students who receive free or reduced-price lunch, and similar variables. The study was designed with power (.80) to detect medium-small to medium effect sizes (0.23 to 0.33).

Participants

As reported in Smolkowski et al. (in press), 127 teachers and 2,817 students within 60 schools participated in the study. Project staff assigned 66 teachers and 1,515 students to the WHS condition (23.0 students/teacher) and 61 teachers and 1,302 students to the control condition (21.3 students/teacher). Although first-grade teachers were invited to participate, about 39% taught kindergarten or a combination of kindergarten and Grade 1. Teachers and students were similar across conditions in a range of demographic and other characteristics. For sample details, please see Smolkowski et al. (in press).

Intervention

Curriculum Description

WHS delivers instruction an academically related social skills, which have been called academic enablers or academic classroom behaviors. WHS is designed to improve key student classroom skills empirically linked to K-3 students' academic success. Pilot testing of this program demonstrated improved student classroom behavior and increased teacher self-efficacy. For a complete description of WHS, its development, and a review of foundational literature, see Smolkowski et al. (in press) and Marquez et al. (2014).

Briefly, WHS delivers universal instruction to improve students' academic and social skills, thereby increasing teacher self-efficacy in managing the classroom. The intervention targets kindergarten to second-grade teachers and their students. It consists of a curriculum with eight lessons and classroom materials, videos and songs for classroom instruction, and an assessment to screen students on the skills taught by WHS. WHS emphasizes a set of 7 classroom behavior skills, covering teacher-preferred, school adjustment skills—*listening to the teacher*, *following directions*, *doing your best work*, *asking for help*, *following the rules*—and interpersonal peer skills that strengthen student school engagement, such as *getting along with others* and *managing strong feelings*. Its highly efficient multimedia resources provide explicit instruction and model of each skill, minimizing teachers' time for preparation and delivery.

Procedures

The study maintained the same procedures for recruitment, implementation, and assessment across the four waves of teachers and students who participated in the study. All procedures were consistent with school and district policies. See Smolkowski et al. (in press) for details about recruitment, the two intervention conditions, implementation, and the use of related programs in study classrooms.

Measures

Teachers completed surveys in the fall and spring of their participation year, reporting their demographics, self-efficacy for classroom management, and concerns about student behavior management. Teachers reported students' background characteristics, student

proficiency with the skills taught by WHS, and their classroom adjustment. Teachers completed implementation fidelity checklists during their use of WHS. Staff members also observed teachers' use of praise and corrections and reported observer general impressions of students' behavior.

Elementary Social Behavior Assessment

The ESBA, derived from the research of Walker and colleagues (Hersh & Walker, 1983; Walker, 1986), assesses the student skills taught by WHS (Walker et al., 2015). The ESBA has been validated in English (Pennefather & Smolkowski, 2015) and adapted to and evaluated in Norwegian (Arnesen et al., 2018) and Swedish (Wikman et al. 2021). Teachers rate items such as "Listens to and respects the teacher" and "Follows the teacher's directions," with short definitions to operationalize each behavior. The definition of the "Listens" item, for example, includes the steps taught to students: "faces you while you are talking," "keeps eyes and ears on you," and "is attentive and waits turn to talk." With the ESBA, teachers rate students on a 3-point scale: skill mastered (3), needs improvement (2), and cause for concern (1).

The average of the 12 items produced a reliable score ($\alpha = .94$) at baseline and remained stable over time among comparison students ($r = .74$). Previous studies found that the ESBA fit a single factor (Pennefather & Smolkowski, 2015), was reliable ($\alpha = .94$–$.95$), correlated well with criterion measures ($r > .77$; Arnesen et al., 2018), and could demonstrate change (Marquez et al., 2014).

Walker–McConnell Scale of Social Competence

With the Walker–McConnell Scale of Social Competence and School Adjustment–Elementary Version (WM; Walker & McConnell, 1995), teachers respond to 43 positively worded behaviors about student social competence. The project collected only the 19-item Classroom Adjustment Behaviors subscale, which included questions about academic competence (i.e., listening skills, participation, responsiveness, and quality of work). Example items include "Other children seek child out to involve her/him in activities" and "Has good work habits." The WM uses a scale from never (1) to sometimes (3) to frequently (5). In this study, the WM had strong score reliability ($\alpha = .96$ at pretest) and stability ($r = .73$, control sample). Previous research on the WM has demonstrated suitable score reliability for subscales ($\alpha > .89$), test-retest reliability ($r > .76$ over 5 weeks), and concurrent validity with the Social Skills Rating System (Gresham & Elliott, 1990).

Teacher Sense of Efficacy Scale

Teachers completed the short form of the Teacher Sense of Efficacy Scale (TSES; Tschannen-Moran & Hoy, 2001), a measure of their self-efficacy for classroom management and instruction. The scale contains 12 items, such as "How much can you do to control disruptive behavior in the classroom?" and "How much can you do to motivate students who show low interest in school work?," with a scale from nothing (1) to a great deal (9). The TSES is reliable and valid (Tschannen-Moran & Hoy, 2001). The TSES has been associated with student achievement (Moore & Esselman, 1992; Ross, 1992), teacher planning and organization (Allinder, 1995), student referrals to special education (Soodak & Podell, 1993), and teaching commitment (Trentham et al., 1985). The summary mean of the 12 items was found to be reliable in the current study ($\alpha = .90$ at pretest).

Teacher Concerns Inventory

The Discipline and Motivation subscale of the modified Teacher Concerns Inventory (TCI; Fimian, 1984) captured teachers concerns about student behavior. The subscale, rated on a scale ranging from no strength or not noticeable (1) to major strength or very noticeable (5), included six items, such as "I feel frustrated because of discipline problems in my classroom" and "I feel frustrated having to monitor pupil behavior." The TCI has demonstrated strong reliability ($\alpha > .75$) and good construct and content validity (Fimian, 1984). The scale average was reliable in the present study ($\alpha = .80$ at pretest).

Implementation Fidelity

The WHS Implementation Fidelity Checklist asked teachers to document their completion of WHS lessons. The checklist for the introductory Lesson 1 asked just four questions. The checklists for Lessons 2 to 8 asked 16 questions about whether teachers completed instruction, practice, feedback, and problem solving activities. Three instruction items asked about use of classroom discussion, the video, and skill cards. A practice item asked if teachers practiced the target skill 3 to 5 times per day. Six feedback items assessed the use of descriptive praise, descriptive correction, skills tickets, tally sheets, happy notices, and certificates of mastery. Finally, teachers reported their use of problem-solving discussion, role-play, songs, games, coloring pages, and skills booklets.

Direct Observations

Independent observers assessed each teacher's behavior after they began using WHS. The lessons suggest teachers praise and correct students with descriptions of their behavior to let them know what they did well and which skills they needed to improve. The observations framed this feedback with directions and nondescriptive praise and corrections, but descriptive praise and corrections were the primary target of the observations. Descriptive praise included comments such as "Thank you for raising your hand" or "Great job listening when others are talking." Nondescriptive praise may have consisted of "Great job" or "Way to go." Corrections were similar but used to remind students about behavior they needed to improve, such as "Please remember to keep your hands and feet to yourselves" (descriptive correction).

Observers, received an overview of the system, an explanation of the codes, and a description of procedures. They used an observation codebook for reference. A trainer and observers practiced coding video and debriefed observations as a group. They then practiced in nonstudy classrooms with a trainer. The trainer and observers established a minimum 80% rate of agreement prior to observing within project classrooms. A trainer periodically retested observers to maintain agreement.

Observers collected data from 114 of the 127 classrooms: 58 in the intervention condition and 56 from the control condition. Scheduling, travel, or other limitations prohibited observations in some classrooms. Observers attempted to code classrooms for 60 minutes in five 10-minute intervals with 2-minute breaks. Some observation periods were cut short by interruptions, so we summarized praise or corrections as rates per minute of observation time.

After collecting all data, we characterized the interrater reliability of the data with intraclass correlations. A set of 42 paired observations produced intraclass correlations of .88 for descriptive praise and .55 for descriptive corrections, representing nearly perfect and moderate reliability, respectively (Landis & Koch, 1977). Nondescriptive praise and feedback produced intraclass correlations of .84 and .43, respectively.

Observers also rated their impressions of student behavior after observation sessions in 107 classrooms. They rated students, as a whole classroom, on seven behaviors: (a) listens to and respects the teacher, (b) follows teacher's directions, (c) works with effort, (d) does seat assignments as directed, (e) makes assistance needs known appropriately, (f) follows rules, and (g) gets along with peers. Ratings ranged from rarely (1) to always (6). An intraclass correlations of .77 suggested substantial interobserver reliability (Landis & Koch, 1977). In seven questions, observer impressions data were unavailable due to technical challenges.

Data Structure

The multilevel WHS data set includes assessments nested within students and students within teachers. The data set includes unique identifiers for each level of nesting. Table 1 describes the variables in the WHS data set including unique identifiers and design variables, student demographic characteristics, student outcome measures, teacher demographic characteristics, teacher outcome measures, teacher-level fidelity of implementation, and teacher-level direct observation measures.

Table 1. *Variable Names and Descriptions for the WHS Data Set*

Variable Name	Description
Unique identifiers and design variables	
StudentID	Student ID
TeacherID	Teacher ID
Tx	Study Condition (1 = WHS, 0 = BAU)
Cohort	Cohort
T	Time coded 0 at Pretest
Time	Time coded 1 at Pretest
Student demographic characteristics	
StDemo_1	Student Demo: Sex (1 = Male, 2 = Female)
StDemo_2	Student Demo: Ethnicity (1 = Hispanic, 2 = Not Hispanic)
StDemo_3	Student Demo: Race†
StDemo_4	Student Demo: Receives ESL services (1 = Yes, 0 = No)
StDemo_5	Student Demo: Has IEP (1 = Yes, 0 = No)
Student outcome measures	
ESBA	Teacher-report Elementary Social Behavior Assessment (mean of 12 items)
WM	Teacher-report Walker-McConnell Classroom Adjustment Behaviors (19 items)
Teacher demographic characteristics	
Demo_01	Teacher Demo: Sex (1 = Male, 2 = Female)
Demo_02	Teacher Demo: Age
Demo_03	Teacher Demo: Educational background††
Demo_04	Teacher Demo: Race†
Demo_05	Teacher Demo: Taught first grade only (1 = Yes, 0 = No)

Variable Name	Description
Demo_06	Teacher Demo: Years of teaching experience
Teacher outcome measures	
TSES	Self-report Teacher Sense of Efficacy Scale (mean of 12 items)
TCI	Self-report Teacher Concerns Inventory (mean of 6 items)
Teacher-level fidelity of implementation	
InstDisc	Fidelity Check (rate of completion): Instruction - Discussion
InstVid	Fidelity Check (rate of completion): Instruction - Video
InstSkillCard	Fidelity Check (rate of completion): Instruction - Skill card
Prac35	Fidelity Check (rate of completion): Practice 3-5 times per day
FBpraise	Fidelity Check (rate of completion): Feedback - Descriptive praise
FBcorrect	Fidelity Check (rate of completion): Feedback - Descriptive correction
FBskills	Fidelity Check (rate of completion): Feedback - Skills tickets
FBtally	Fidelity Check (rate of completion): Feedback - Tally sheets
FBhappy	Fidelity Check (rate of completion): Feedback - Happy notices
FBmastery	Fidelity Check (rate of completion): Feedback - Certificate of mastery
ACTprobs	Fidelity Check (rate of completion): Activities - Problem-solving discussion
ACTrole	Fidelity Check (rate of completion): Activities - Role-play
ACTsong	Fidelity Check (rate of completion): Activities - Song
ACTgames	Fidelity Check (rate of completion): Activities - Games
ACTcolor	Fidelity Check (rate of completion): Activities - Coloring page
ACTskills	Fidelity Check (rate of completion): Activities - Skill booklet
Teacher-level direct observation measures	
DescPrRate	Observed rate of descriptive praise per minute
DescFDoRate	Observed rate of descriptive corrections per minute
BehScale	Observer impressions of student behavior per minute
NonDescPrRate	Observed rate of nondescriptive praise per minute
NonDescFDoRate	Observed rate of nondescriptive corrections per minute

Note. †Race is coded 1 = White/Caucasian, 2 = Black/African American, 3 = Asian, 4 = American Indian/Alaskan Native, 5 = Pacific Islander/Native Hawaiian, 6 = More than one race, 7 = Other, and 8 = Prefer not to answer.

††Education background is coded 1 = Bachelor's, 2 = Master's, 3 = Doctorate or professional, and 4 = Other.

References

Allinder, R. (1995). An examination of the relationship between teacher efficacy and curriculum-based measurement and student achievement. *Remedial and Special Education*, *16*, 247–254.

Arnesen, A., Smolkowski, K., Ogden, T., & Melby-Lervåg, M. (2018). Validation of the Elementary Social Behavior Assessment: Teacher ratings of students' social skills adapted to Norwegian, Grades 1 to 6. *Emotional and Behavioural Difficulties*, *23*(1), 39–54. https://doi.org/10.1080/13632752.2017.1316473

Fimian, M. J. (1984). The development of an instrument to measure occupational stress in teachers: The Teacher Stress Inventory. *Journal of Occupational Psychology, 57*, 277–293.

Gresham, F. M., & Elliott, S. (1990). *Social Skills Rating System*. American Guidance Service.

Hersh, R., & Walker, H. M. (1983). Great expectations: Making schools effective for all students. *Policy Studies Review*, *2*, 147-188.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Marquez, B., Marquez, J., Vincent, C. G., Pennefather, J., Sprague, J. R., Smolkowski, K., & Yeaton, P. (2014). The iterative development and initial evaluation of *We Have Skills!*, an innovative approach to teaching social skills to elementary students. *Education and Treatment of Children*, *37*(1), 137-161.

Moore, W., & Esselman, M. (1992). *Teacher efficacy, power, school climate and achievement: A desegregating district's experience*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Pennefather, J. T., & Smolkowski, K. (2015). Validation of the Elementary Social Behavior Assessment: A measure of student prosocial school behaviors. *Journal of Assessment for Effective Intervention*, *40*(3), 143–154.

Smolkowski, K., Walker, H., Marquez, B., Kosty, D., Vincent, C., Black, C., Cil, G., & Strycker, L. A. (in press). Evaluation of a social skills program for early elementary students: We Have Skills. *Journal of Research on Educational Effectiveness*. https://doi.org/10.1080/19345747.2022.2037798

Soodak, L. C., & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *Journal of Special Education*, *27*(1), 66–81.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805.

Walker, H. M. (1986). The AIMS (Assessments for Integration Into Mainstream Settings) assessment system: Rationale, instruments, procedures and outcomes. *Journal of Clinical Child Psychology*, *15*(1), 55–63.

Walker, H. M., & McConnell, S. (1995). Walker–McConnell Scale of Social Competence and School Adjustment: Elementary version. Singular Publishing Group.

Wikman, C., Allodi, M. W., & Ferrer-Wreder, L. (2021). Psychometric properties of the Elementary Social Behavior Assessment in Swedish primary school: A teacher rated index of students' prosocial school behaviors. *Frontiers in Education*, *6*(681873). https://doi.org/10.3389/feduc.2021.681873

Wikman, C., Allodi, M. W., & Ferrer-Wreder, L. (2021). Psychometric properties of the Elementary Social Behavior Assessment in Swedish primary school: A teacher rated index of students' prosocial school behaviors. *Frontiers in Education*, *6*(681873). https://doi.org/10.3389/feduc.2021.681873