Supporting Preschool and Kindergarten Students' Self-Regulation Through HighScope Curriculum Enhancements: Plan-Do-Review and Conflict Resolution U411C160089

Independent evaluator final report

U411C160089

This report was produced by an independent evaluation team lead by Lori Skibbe, for more details about this report please contact skibbelo@msu.edu.
For more information about details of the overall program -development, implementation and internal monitoring and evaluation) please contact Fernando H. Andrade Adaniya (fadaniya@highcope.org)

Introduction

The primary goal of the proposed project was to evaluate the impact of HighScope's curricular enhancements on children's self-regulation and academic development during preschool. **Key goals** of this project were to 1) Develop enhanced versions of PDR and CR, which are promising new strategies to support self-regulation, 2) understand whether teachers can implement PDR and CR with fidelity, 3) examine whether enhanced PDR and CR can increase children's self-regulation skills, and 4) consider whether these intervention efforts can also support children's academic achievement.

Methods

Participants in 2019-2020

Participating parents/caregivers and children (n = 266, Male = 51.1%) were recruited from their preschool programs in the fall of the school year. Children ranged in age from 45 to 68 months (M = 52.27 months, SD = 4.01 months). Children's race and ethnicity were reported by their parents/caregivers and included the following: 49.1% were identified as Black/African American, 47.9% as White/Caucasian, 1.9% as Asian/Pacific Islander, and 1.5% as American Indian/Alaskan. Parents/caregivers indicated that 9.8% of children were Hispanic/Latino of any race. Most children (93.5%) indicated that they understood and spoke English like a native speaker. In addition, the majority of mothers (82.5%) reported that they did not complete college (1.5% reported no school attended, 8.0% some high school, 32.3% graduated high school, 40.7% some college), although some mothers (12.2%) reported receiving a bachelor's degree, with a few (5.3%) attending graduate school as well. Teachers reported that all children were able to be assessed using conventional measures in field, although 34% of children were reported by their parents/caregivers to have some developmental concerns. All children enrolled in the study qualified for free and/or reduced lunch from their schools. Measures collected during the fall included subtests targeting working memory, inhibitory control, and attention from the TX-KEA (Montroy et al., 2020) as well as a measure of behavioral self-regulation (McClelland et al., 2014). Children were also administered the Letter-word identification and Applied Problems subtests from the Woodcock Johnson-IV (Schrank et al., 2014). Unfortunately, no data were collected in the spring of 2020, as schools shut down due to COVID, so it was not possible to judge the effects of the intervention.

Participants in 2020-2021

Ninety-seven preschool-aged children (Male = 46) were tested remotely in the fall. Families were called by our team to set up baseline child assessments that took place during the fall semester of 2020. Parents/caregivers were asked to fill out a questionnaire that included questions about their families' demographic and home characteristics, for which they were compensated $25. Approximately 31% ($n$ = 43) of families did not attend their remote assessments appointment, did not complete the parent survey, and subsequently dropped out of the study. Participating children were between 46 and 62 months old (M = 53.2 months, SD = 4.4 months). Approximately 55% of parents/caregivers reported their children's race as Black/African American, 33.7% White/Caucasian, 7.9% biracial, 2.2% other, and 1.1% Asian/Pacific Islander; 7.2% of parents/caregivers reported their children were Hispanic/Latino of any race. English was the primary language spoken at home for 94.6% of children, with 12% also speaking another language. All participating families reported being proficient in English,

which was the language of our assessment battery. Approximately 72% of mothers did not complete college (3.3% reported no school attended, 5.4% some high school, 23.9% graduated high school, 39.1% some college), with 16.3% of mothers reporting completing an undergraduate degree, and 12% attending graduate or professional school. About 35% of parents/caregivers reported that their children had mild medical or developmental concerns, ranging from speech and language impairments to socioemotional and learning difficulties. However, no children were excluded from our study based on an inability to participate due to medical or developmental reasons. Data from this year were used to validate our process for remote data collection (Ahmed et al., 2021; McRoy, under review).

Main Test of the Intervention

Participants in 2021-2022

Three hundred sixty children (Male = 162) were recruited for the present work. Of those, 152 children were in classrooms whose teachers received the intervention and 208 were attending control classrooms. Participating children were between 35 and 63 months old (M = 50.3 months, SD = 6.3 months) in the fall. Approximately 26.4% of parents/caregivers reported their children's race as Black/African American, 29.2% White/Caucasian, 11.5% were multi-racial, 3.3% listed their race as 'other', and 3.1% noted that they were Asian/Pacific Islander; 26.5% parents/caregivers did not report on their child's ethnicity. In addition, 4.2% of parents/caregivers reported their children were Hispanic/Latino of any race. English was the primary language spoken at home for 94.8% of children. All participating families reported that their child was proficient in English, which was the language we used to assess them. Of the mothers who filled out the demographic survey (n = 246), 46.4% of mothers did not complete college (3.9% some high school, 20.6% graduated high school, 21.9% some college); 11.4% of mothers reported completing an undergraduate degree and 10.6% attended graduate or professional school. About 26% of parents/caregivers reported that their children had mild medical or developmental concerns, but no children were excluded from the present work.

Measures of Self-regulation

Behavioral Self-Regulation was measured using the revised version of the Head Toes Knees Shoulders task (HTKS-R; McClelland et al., 2021). This measure has been shown to be valid and reliable in prior work (Gonzales et al., 2021). Possible scores on this measure range from 0 to 118, with higher scores indicating greater self-regulation.

Inhibitory Control was measured using The Day/Night Stroop task. This measure is appropriate for children who are 3 to 7 years age. The Day/Night Stroop task is considered to be both reliable (Rhoades et al., 2009; von Stauffenberg & Campbell, 2007) and valid (Carlson & Moses, 2001). Possible scores using this task range from 0 to 28 and higher scores represent greater inhibitory control.

Working Memory was measured using the The Digit Span task (Wechsler, 1991). This measure requires children to store, maintain, and manipulate information. Given the age of participants, only the forward span items were used. This task is associated with excellent reliability (Lipsey et al., 2017). Scores on this measure represent the maximum number of digits children could recall, with higher scores indicating greater working memory.

Cognitive Flexibility was measured using the The Dimensional Card Change Sort Task (DCCS; Zelazo, 2006), a measure which is appropriate for children who are 3 to 5 years of age.

This measure has excellent reliability (Beck et al., 2011). Possible scores using this task range from 0 to 12, with higher scores indicating greater cognitive flexibility.

Academic Achievement

Early literacy skills were assessed using the Letter-Word Identification subtest of the Woodcock-Johnson Tests of Early Cognitive and Academic Development (Schrank & Wendling, 2015). Initial items capture children's letter knowledge; later items require children to decode increasingly complex words. Scores are presented using *W*-scores, which incorporate Rasch-based measurement techniques to account for item difficulty as well as children's age. Reported reliabilities for this subtest are excellent (Range = .92–.94).

Early math skills were assessed using the number sense subtest of the Woodcock-Johnson Tests of Early Cognitive and Academic Development (Schrank & Wendling, 2015). This subtest assessed mathematical knowledge using word problems, pictures, and numbers. Participants were asked to listen to each item, determine the procedure to solve the problem, and successfully complete the computations. Similar to the approach used for early literacy skills, all analyses were conducted using *W* scores. The questions became increasingly difficult as participants progressed through the task. External validation efforts have demonstrated excellent reliability among young students and high test-retest reliability (Cronbach's $\alpha = .93$).

Preliminary Analyses

Before testing program impact, we conducted a confirmatory factor analysis (CFA) of pre-test and post-test self-regulation measures to ensure strong factor solutions before using latent variables in our analyses. Analyses were conducted in Mplus version 8.4 (Muthen & Muthen, 1998-2019) and full information maximum likelihood (FIML) estimation was used to account for missing data. Standard errors were adjusted to account for the nesting of children within classrooms using the Mplus CLUSTER command. We fit the CFA using fall and spring assessments of Digit Span Forward, Day/Night Stroop, DCCS, and HTKS-R, and correlated the residual variances among self-regulation latent variables across wave. The CFA indicated a good model fit: ($\chi^2(15) = 23.06$, $p = .08$, RMSEA = .04, CFI = .99, TLI = .98, SRMR = .03); standardized factor loadings ranged from 0.47 to 0.78 and all were significant.

We considered whether there were any significant group differences between the treatment and control groups on variables known to impact our outcomes (self-regulation, math skills, early literacy). There were significant group differences in child age at the outcome assessment ($t(289) = 2.67$, $p = .004$, such that children in the control group were, on average, older than children in the treatment group), child sex ($t(358) = -2.02$, $p = .022$, with more representation of females than males in the treatment group) and child disability status ($t(243) = 1.67$, $p = .048$, with more children with disabilities represented in the control group). There were no significant group differences in maternal education level ($t(244) = -1.63$, $p = .052$). Based on these analyses, we controlled for these variables in our analyses when possible.

Program Impact Analyses

We first attempted to use a Multi-level Structural Equation Modeling (MSEM) approach in software Mplus using the multi-level add-on TWOLEVEL. MSEM allows for latent variable outcomes, such as our Aim 1 outcome of self-regulation as a latent variable indicated by 4 components (Digit Span Forward assessment, Day/Night Stroop total score, DCCS total score, HTKS-R total score). It also is suitable for our Aim 2 outcomes of math and literacy

achievement. However, we encountered convergence issues in models for all outcomes (latent self-regulation, observed math skills, observed early literacy achievement). Therefore, model simplification was necessary.

As a result, all program impact analyses were conducted within a regression framework using Mplus version 8.4 (Muthen & Muthen, 1998–2019). Specifically, a dichotomous variable representing treatment status (0 = control; 1 = treatment) was entered as a predictor of all child outcomes while adjusting the standard errors to account for the nested structure of the data (i.e., children in classrooms). Full information maximum likelihood (FIML) estimation was used to account for missing data. The child- and family level variables that significantly differed by treatment status were covaried in our models to reduce potential bias.

Self-Regulation

Treatment status did not significantly predict changes in children's self-regulation from fall to spring of preschool (b = 0.103, *S.E.* = 0.060, p = 0.086). In this model, fall and spring self-regulation were treated as latent variables and their residual variances were correlated across wave. This model fit the data well ($\chi^2(51) = 71.28$, p = 0.032, RMSEA = 0.033, CFI = .971, TLI = .956, SRMR = .039).

Academic Achievement

Treatment status did not significantly predict changes in children's math achievement (b = -0.066, *S.E.* = 0.048, p = 0.166) or literacy achievement (b = -0.041, *S.E.* = 0.048, p = 0.390) from fall to spring of preschool. Model fit for math and literacy achievement models were just-identified due to having the same number of observations as free parameters.

Discussion of Results

Across all analyses, we consistently failed to detect any evidence for treatment effects for math and early literacy. Less clear, however, is the impact of HighScope's curricular enhancements on children self-regulation development. Based on our models, children enrolled in treatment classrooms exhibited greater growth in self-regulation across preschool. Although differences in growth were not statistically significant, the magnitude of effect is consistent with what has been found in previous preschool self-regulation intervention programs (e.g., Blair & Raver, 2014; Mattera et al., 2021). Additionally, the measures associated with the self-regulation latent variable demonstrated good model fit, indicating that this was an appropriate way to conceptualize children's self-regulation development.

References

Ahmed, S. F., Skibbe, L. E., McRoy, K., Tatar, B. H., & Scharphorn, L. (2022). Strategies, recommendations, and validation of remote executive function tasks for use with young children. *Early Childhood Research Quarterly, 60*, 336–347.

Blair, C., & Raver, C. C. (2014). Closing the achievement gap through modification of neurocognitive and neuroendocrine function: Results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. *PloS one*, *9*(11), e112393.

Gonzales, C. R., Bowles, R., Geldhof, G. J., Cameron, C. E., Tracy, A., & McClelland, M. M. (2021). The head-toes-knees-shoulders revised (HTKS-R): Development and psychometric properties of a revision to reduce floor effects. *Early Childhood Research Quarterly, 56*, 320–332.

Lipsey, M. W., Nesbitt, K. T., Farran, D. C., Dong, N., Fuhs, M. W., & Wilson, S. J. (2017). Learning-related cognitive self-regulation measures for prekindergarten children: A comparative evaluation of the educational relevance of selected measures. *Journal of Educational Psychology*, *109*, 1084–1102.

Mattera, S., Rojas, N. M., Morris, P. A., & Bierman, K. (2021). Promoting EF with preschool interventions: Lessons learned from 15 years of conducting large-scale studies. *Frontiers in Psychology*, *12*, 640702.

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., & Pratt, M. E. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Frontiers in Psychology*, *5*, 599.

McClelland, M. M., Gonzales, C. R., Cameron, C. E., Geldhof, G. J., Bowles, R. P., Nancarrow, A. F., . . . Tracy, A. (2021). The head-toes-knees-shoulders revised: Links to academic outcomes and measures of EF in young children. *Frontiers in Psychology, 12*, 14.

McRoy, K., Skibbe, L., Ahmed, S., & Ozkum, B. (under review). *Guidelines for assessing academic and executive function for young children using a remote testing platform.*

Montroy, J. J., Zucker, T. A., Assel, M. M., Landry, S. H., Anthony, J. L., Williams, J. M., . . . Taylor, H. B. (2020). The Texas kindergarten entry assessment: Development, psychometrics, and scale-up of a comprehensive screener. *Early Education and Development, 31*, 701–738.