

9-1-2024

Assessing Staff-Youth Interaction Quality: Using the CLASS to Evaluate Afterschool Programs

Joel M. Hektner

North Dakota State University, joel.hektner@ndsu.edu

Debarati Kole

North Dakota State University, debarati.kole@ndsu.edu

Vimbayi Chinopfukwa

College of Saint Benedict and Saint John's University, VCHINOPFU001@CSBSJU.EDU

Shawn L. Carlson

Purdue University, carlsosl@purdue.edu

Follow this and additional works at: <https://open.clemson.edu/jyd>



Part of the [Child Psychology Commons](#), [Civic and Community Engagement Commons](#), [Community-Based Learning Commons](#), [Developmental Psychology Commons](#), and the [Leadership Studies Commons](#)

Recommended Citation

Hektner, Joel M.; Kole, Debarati; Chinopfukwa, Vimbayi; and Carlson, Shawn L. (2024) "Assessing Staff-Youth Interaction Quality: Using the CLASS to Evaluate Afterschool Programs," *Journal of Youth Development*. Vol. 19: Iss. 3, Article 3.

Available at: <https://open.clemson.edu/jyd/vol19/iss3/3>

This Research and Evaluation Study is brought to you for free and open access by Clemson OPEN. It has been accepted for inclusion in Journal of Youth Development by an authorized editor of Clemson OPEN. For more information, please contact kokeefe@clemson.edu.



Volume 19, Issue 3, Fall 2024

ISSN 2325-4017 (online)

Assessing Staff-Youth Interaction Quality: Using the CLASS to Evaluate Afterschool Programs

**Joel M. Hektner
Debarati Kole
Vimbayi Chinopfukutwa
Shawn L. Carlson**

Introduction

More than 1.6 million children and youth in the United States attend 21st Century Community Learning Centers (21st CCLCs), which are federally funded after-school programs available in each state, particularly for students who attend high poverty and low-performing schools (US Department of Education, 2020). These 10,125 centers are intended to do much more than provide supervision and safety for children after-school while their parents are still at work. Their mission includes academic support and enrichment, social and emotional learning, drug and violence prevention, and physical activity and nutrition education (Afterschool Alliance, 2021). The centers do not all follow a prescribed curriculum, approach, or structure, even within their respective states, but they do all share common goals and report data on progress toward those goals (Weiss, 2013). Each state is mandated to conduct evaluations of its 21st CCLC programs, and these evaluations have grown in sophistication over the past decade (US Department of Education, 2020). However, these evaluations typically focus on mean-level outcomes, without considering the microlevel social interactions between staff and youth and among peers that are instrumental in producing those outcomes. The purpose of the present study is to explore the use of the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008a) observation protocol in a statewide evaluation of 21st CCLCs and its potential to be useful in designing and targeting training and professional development opportunities for 21st CCLC program staff.

A national evaluation of 21st CCLC programs in 2005 showed improved feelings of school safety among the elementary students who participated in the programs. However, there was no significant impact on participants' academic achievement (James-Burdumy et al., 2005). The 2007 national evaluation report also reported no effect on students' academic outcomes and an increased rate of negative behavior of program participants, especially boys with disciplinary problems. In contrast, state evaluation reports from the last decade have shown improvement in students' academic achievement. For example, Michigan's most recent 21st CCLC evaluation reported improvements in participants' overall grade point averages and homework completion (Wu & Van Egeren, 2023). In New York, improvements were reported in student performance relating to attendance, academic work, discipline, and social behaviors (Dodd & Bowen, 2011). Teachers in West Virginia reported that most 21st CCLC participants had improved classroom behavior, completion of homework, and class participation (Hammer & Whisman, 2017).

Youth self-reports have also been used in evaluations of 21st CCLC programs. Participants of 21st CCLCs in Oklahoma in 2015–2016 reported that they were able to make and stay friends with other youth most of the time. Regarding academic efficacy, participants reported they had good work habits and felt more efficacious in science and technology than in math and reading (Gersh et al. 2017). Eighty percent of youth in Michigan's 21st CCLC programs reported that the activities they do in the program helped them do better in school. Participants also noted the program's impact on a wide range of social-emotional skills, including understanding how other people feel and taking responsibility for their actions (Wu & Van Egeren, 2023). Other positive feelings were reported by 21st CCLC participants in Nebraska, including program belonging and engagement. These youth also reported they enjoyed going to the program, made friends, learned new things, and felt respected by the adults in the program (Johnson et al., 2022).

None of these 21st CCLC evaluations measured the quality of the interactions between staff and youth, which is not an ultimate outcome in itself but may be a key factor in most of the outcomes of interest. Smith and colleagues (2012) recommended the assessment of adult-youth interactions in evaluations of out-of-school-time (OST) programs because the great variation in program content and staff qualifications makes relying on those two factors difficult. Previous research has shown that supportive and close teacher-student relationships can predict students' academic achievement (e.g., Ansari et al., 2020; McCormick & O'Connor, 2015). The CLASS has been used extensively in research on regular classroom instruction, and the US Early Childhood Learning and Knowledge Center (2021) uses the pre-K version for national evaluations of Head Start. But we can find no published reports of its use in OST settings such as 21st CCLCs. With foundations in socioecological theories of development (e.g., Bronfenbrenner & Morris, 1998) and constructivist theories of learning (e.g., Rogoff, 1990; Vygotsky, 1978), the CLASS is a structured observation protocol that focuses on the interactions that take place between staff and youth. It provides scores on 10 dimensions, which are grouped into three domains. The domain of (1) *emotional support* includes the dimensions of *positive climate*, *negative climate* (reverse scored), *teacher sensitivity*, and *regard for student perspectives*; (2) *classroom organization* includes the dimensions of *behavior management*, *productivity*, and *instructional learning formats*; and (3) *instructional support* includes *concept development*, *quality of feedback*, and *language modeling* (Pianta et al., 2008a).

In regular school classrooms, a number of studies have shown positive relationships between CLASS scores and student outcomes, including achievement in math and reading (e.g., Johnson et al., 2016; LoCasale-Crouch et al., 2018) and gains in social cognitive and executive skills (Hamre et al., 2014). Research in preschool and early elementary classrooms has also shown that CLASS scores tend to be higher in most domains during sessions focused on science than those focused on other content (Cabell et al., 2013; Thorpe et al., 2020). While some of the reasons for this may be unique to science, other aspects may be transferrable to other content areas with appropriate mentoring.

Allen and colleagues (2015) reported improvements in student achievement on standardized tests when the CLASS was used as part of a teacher coaching program called My Teaching Partner-Secondary. Teachers video recorded portions of their class, and coaches selected clips illustrating positive interactions

and those that could have benefitted from a different approach. The coach and teacher then had a dialog on enhancements the teacher could work on, using the CLASS dimensions as a framework. In a similar approach, Kim and colleagues (2019) developed the Teacher Instructional Practices and Processes System (TIPPS) to allow for cultural variations in low- and middle-income countries. The use of TIPPS and CLASS in this way can help in cultivating equitable education systems by closing opportunity gaps, and thus achievement gaps related to race, ethnicity, dual language use, and social status (e.g., Garwood et al., 2017; Hamre & Pianta, 2005; Hindman & Wasik, 2013; Milner, 2012). These approaches also fit well with the continuous improvement model, in which teachers implement a practice, identify areas for improvement, and then revise the practice to fit the context in an iterative cycle (Knight & Skrtic, 2021). Of course, any comprehensive evaluation would incorporate other measures as well, such as feedback from stakeholders and direct assessment of participant outcomes.

The settings, structures, and activities found in 21st CCLCs are somewhat different from regular school classrooms and are much more diverse across programs than classrooms are across schools and districts. Thus, it is not a given that the CLASS would be feasible and appropriate for use in these programs. On the other hand, a primary goal of CCLCs is academic support and enrichment, and the domains and dimensions of the CLASS would appear to be relevant in any learning environment. Moreover, continuous improvement projects using ongoing observations and coaching of staff have also been documented in OST programs (Sheldon & Hopkins, 2008), and so it would seem logical that the CLASS could be useful in this regard also.

To determine whether the CLASS could be useful both as an evaluation tool of 21st CCLC programs and as a way to motivate and focus efforts to improve practice, we addressed the following research questions in the current study:

1. Can CLASS feasibly be used to gather reliable observation data on staff-youth interactions in 21st CCLCs? What adaptations, if any, need to be made to the CLASS observation protocol in order to use it in 21st CCLCs?
2. How do CLASS scores from 21st CCLCs serving elementary-age students compare to previously published scores from regular elementary school classrooms? This question is intended to provide some context for interpreting the meaning of CLASS scores to OST programs rather than to use regular classrooms as a standard against which to compare OST programs.
3. Do CLASS scores vary with content of the session (i.e., sessions focused on math or science versus other sessions)? How can these variations inform staff professional development?

Method

As part of a statewide evaluation of 21st CCLCs, observations were conducted at 30 sites across the state of North Dakota, including urban and rural areas. Across all sites in the state (not only the sites observed), 74% of the children served were White, 12% were Native American, 5% were Black, 4% were Latino/a, and the rest identified as one or more other groups or did not identify. Roughly equal numbers of boys and girls were served, and 33% of the children received free or reduced-price meals at school. Prior to beginning observations, eight observers completed a two-day training in using the K-3 CLASS and then passed a test, achieving at least 80% agreement with master coders across all ten dimensions on each of five video-recorded classroom sessions. Agreement is defined as being within one point of each other. Observers were located in four different regions of the state and included four parent educators, two graduate students, one professor, and one 21st CCLC regional administrator. None of the observers had any personal or professional relationships with the staff they observed. Staff were informed that their

CLASS scores would not be kept in their personnel file and would not be used in performance evaluations. Institutional Review Board approval was granted by the first author's institution, and informed consent letters were sent to all teachers and parents at participating sites.

Although only one observer is needed at each session, we chose to have two observers conduct each observation so that we could test for and ensure high reliability. At each site, observers observed for twenty minutes, taking notes on all interactions between staff and youth. Then they spent 10 minutes reviewing their notes and determining a score for each of the ten dimensions. They then repeated that cycle one more time. Doing only two observation cycles per session was an adaptation necessitated by the shortened hours of after-school programs. When the CLASS is used in regular school classrooms, four to six of these cycles are standard (Pianta et al., 2008a). This was the only deviation from the standard protocol we made. Per the protocol, the CLASS scores are based on the behavior of all the adults in the room during the observation cycle; the observers did not necessarily focus on just one staff member. As a result, in a few cases the second cycle included different staff members than the first. Both structured activities and unstructured free time were observed, as well as transitions. Only rooms including students from kindergarten to third grade were observed, to match the version of the CLASS observers were trained to use, and because that grade range is the most prevalent in 21st CCLCs in the state.

All 10 dimensions described above were scored. Scores on each dimension range from 1 to 7. Observers are trained to consider a number of different indicators for each dimension, and to categorize each indicator as being in the low, middle, or high range. For example, the indicators for the positive climate dimension are relationships, positive affect, positive communication, and respect. If nearly all indicators for a dimension are within the low range (e.g. the teacher and/or students display no or few positive affect), the dimension is scored as 1; medium range (e.g. sometimes, the teacher and/or students display positive affect), 4; and high range (e.g. frequent displays of positive affect by the teacher and/or students), 7. If there is a mix of indicators from different ranges, then the dimension is scored with a 2, 3, 5, or 6, depending on whether the mix is predominantly in the low range with some middle, the middle range with some low, the middle range with some high, or in the high range with some middle, respectively (Pianta et al., 2008a).

Interrater reliability between observers in the present study was high, as indicated by an overall agreement rate of 92.1%. The CLASS manual reports average reliability of 87% across several national studies (Pianta et al., 2008a). In the current study, if the observers did choose different scores for a particular dimension, their scores were averaged.

For each session, scores across the two cycles were averaged to obtain a score for each of the ten dimensions, and then domain scores were calculated as means of their respective dimensions. Observers also recorded the grade level(s), format, and content of the session. Content was recorded as one or more of the following: art, math, language arts, science, or social studies.

Example of a CLASS-Scored Observation

During two cycles of one observation, we observed staff leading students in circle time (deep breathing and individual sharing of perspectives) followed by an art project. Throughout, staff showed positive affect and respect for students, calling them by name, as well as sensitivity, by, for example, helping a student save their project when they were picked up early by parents. The project was introduced as creating an invention to capture a leprechaun (as it was close to Saint Patrick's Day). The youth were asked to draw out a plan on paper first before coming to the art cart to get supplies. Many were very excited about the project, but the staff reinforced appropriate behavior. While students were building their traps, the staff member visited each one and asked questions promoting deeper thought processes, such as, "What's the point of the straws?" and offered assistance such as tying strings. At the end of the session, the staff member announced that in about 5 minutes they would need to clean up and move into choice time. This session was scored high in the dimensions making up the emotional support and classroom organization domains and in the

midrange on the instructional support dimensions. Recommendations to improve instructional support might include asking the staff how they could build in more opportunities for analysis and reasoning (such as asking students to be ready to explain how their invention works and what its strengths and weaknesses are). They could also work in more advanced vocabulary (such as *lure* and *camouflage*) and try to incorporate stronger connections to real-world experiences and to other academic lessons (such as testing which shapes in a structure are the strongest).

Results

Table 1. *Statewide Means on Each CLASS Domain and Dimension*

Domain or Dimension	<i>M</i>	<i>SD</i>
Emotional Support	6.01	0.50
Negative climate	1.17	0.27
Positive climate	6.28	0.64
Regard for student perspectives	4.99	1.27
Teacher sensitivity	5.98	0.77
Classroom Organization	5.66	0.83
Behavior management	5.90	0.94
Instructional learning formats	5.26	1.09
Productivity	5.83	0.94
Instructional Support	3.31	1.18
Concept development	3.34	1.33
Language modeling	3.23	1.17
Quality of feedback	3.36	1.19

N = 30 sites.

Statewide means on each CLASS domain and dimension are reported in table 1. In the emotional support domain, the positive climate and teacher sensitivity dimensions were both in the high range, and the negative climate dimension, when reversed, had the best score of all ten dimensions. Regard for student perspectives was the only dimension in this domain to fall in the middle range. Regard for student perspectives includes showing flexibility, incorporating students' ideas, following students' lead, allowing students choice, giving students responsibilities, encouraging their talk, eliciting their ideas, and allowing their movement. Taking these four dimensions together, the overall domain score for emotional support was 6.01, just within the high range. Other published studies using CLASS in K-5 classrooms in different regions around the country reported scores on these dimensions from the mid-4's to the mid-5's (and low-1's to low-2's for negative climate; Pianta et al., 2008a). Thus, the 21st CCLCs in this study performed comparably or better than regular classrooms in this domain.

The overall *classroom organization* score was just slightly lower, but still high enough to round up to the high range. *Instructional learning formats* was the only dimension in this domain to score below 5.5. This dimension focuses on ways in which the staff member maximizes the students' interest, engagement, and ability to learn. It includes effective facilitation, questioning, and use of a variety of modalities and materials. It is also indicated by youth showing active participation, listening, and focused attention. These scores are also consistent with published means from regular classrooms (Pianta et al., 2008a).

On *instructional support*, regular school classrooms tend to score lower than in the other domains, generally in the low-middle to middle range (Pianta et al., 2008a). The 21st CCLCs in this study had similar but slightly lower scores, with all *instructional support* scores between 3.0 and 3.5.

A series of t-tests on cycle-level data revealed some significant differences ($p < .05$) in scores with respect to content of sessions. Sessions focused on math (12% of all sessions) had significantly higher scores in *teacher sensitivity* and *productivity* than sessions focused on anything else (see figure 1). However, math sessions also had significantly lower scores in the *instructional support* domain and in all three dimensions included in it. On the other hand, sessions focused on science (35% of all sessions) were unambiguously more positive than other sessions, with higher scores in the instructional support domain and its dimensions of *concept development* and *quality of feedback*, in addition to the dimensions of *instructional learning formats* and *positive climate* (see figure 2).

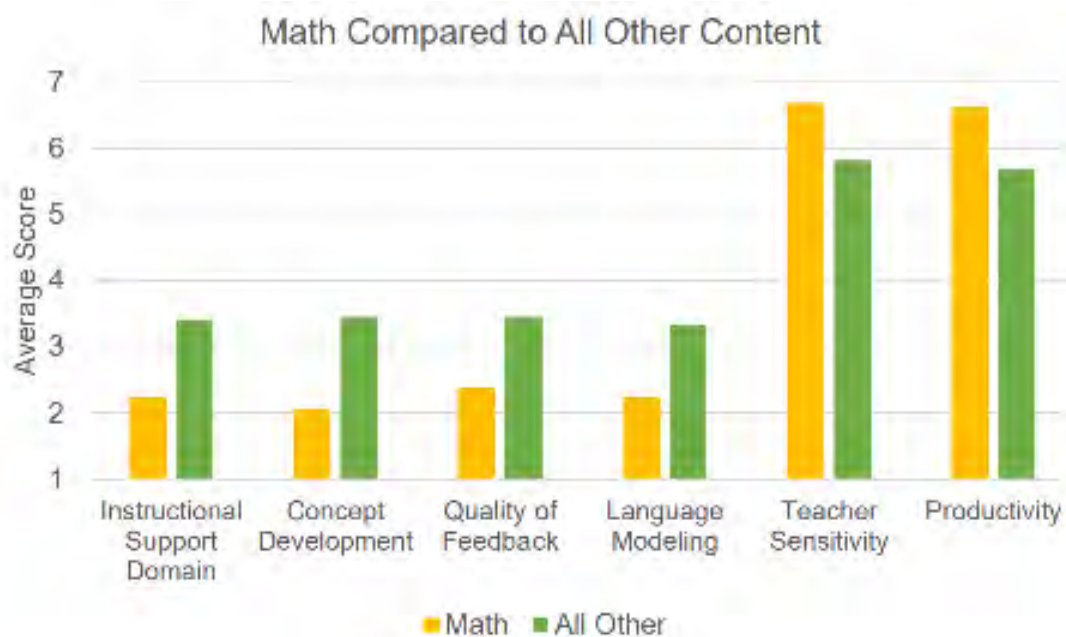


Figure 1. Math Compared to All Other Content. All differences $p < .05$.

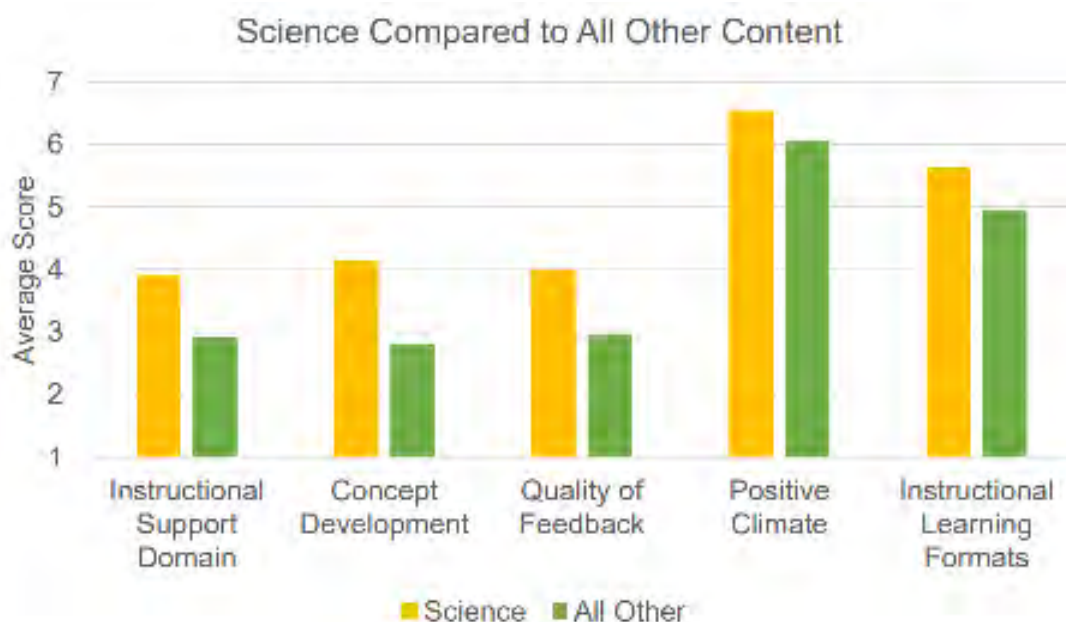


Figure 2. Science Compared to All Other Content. All differences $p < .05$.

Discussion

The current study showed that it is feasible to use the CLASS observation protocol to reliably evaluate the social and learning environment in after-school programs that have learning goals, and that the ranges of scores obtained in these settings are generally consistent with those in regular classrooms. Adaptations were needed to accommodate the different structure of after-school programs, including observing for only two cycles per session instead of four. Given the shortened duration of after-school programs compared to the regular school day, two cycles were sufficient to include a substantial portion of the after-school period. At some sites, the second observation cycle within a session may have included different staff members than the first cycle. Other OST programs, such as those in the morning before the school day or those that focus purely on sports and recreation, may require further adaptations or may not be appropriate for the use of the CLASS tool.

These observations showed that the 21st CCLCs in this study provided high levels of emotional support and were also strong in classroom organization. About half of the observed cycles included children from two or more grade levels, and these multiage sessions may have contributed to the strength seen in these areas. Multiage sessions may provide more opportunities for role modeling and positive relationships to develop, which is also evident in studies examining multiage regular school classrooms (Cozza, 2023; Johnson & Johnson, 1994; Kinsey, 2001; McClellan & Kinsey, 1999). Interacting with others can also help these students build foundational skills such as the ability to communicate effectively and think critically, take others' perspectives, and develop self-regulation skills (Afterschool Alliance, 2023).

The centers were much weaker in providing instructional support and lagged slightly behind regular school classrooms in this area. This outcome may not be surprising, given the different nature and purpose of 21st CCLCs compared to regular school classrooms. But it also points to opportunities for growth, which can be enhanced through professional development offerings. The findings that science sessions, in particular, were stronger in instructional support may provide starting points for trainings on how staff can strengthen their work in this area. In science sessions, staff tend to use diverse and engaging instructional learning formats as well as scaffolding, which promotes higher order thinking skills as students focus on problem-solving and creativity (Hamre et al., 2014; Pianta et al., 2008a). Staff could be trained to recognize the richer vocabulary and open-ended questions that they naturally use during science classrooms and then apply those same skills to other content. Staff trainings may also focus on the diversity of learners since they lead both single-age and multiage groups. Staff may expect to find greater diversity in cognitive, social, and physical development among multiage groups than single-age groups. To maximize each student's learning and achievement during sessions, staff may consider individual and group differences by using a variety of instructional strategies to help youth accomplish challenging tasks (Gayfer, 1991; Vincent, 1999; Vygotsky, 1978).

After each observation in the current study, we provided each observed staff member and their supervisor with a report of their CLASS scores along with a scoring guide provided by the publisher. In many cases, the staff member then worked with their supervisor or another mentor to identify areas they wanted to improve in and strategies they would implement. Although we collected anecdotal reports on this process, implementing and evaluating a formal professional development coaching program was beyond the scope of the current study. Programs planning to implement observations and coaching for professional development will need to consider power dynamics and equity, perhaps separating the process from high-stakes performance evaluations as in the current study. If the CLASS tool is to be used, programs will also need to find certified observers or fund their own personnel to get training and become certified.

A template for how the CLASS observational results could be used in professional development efforts of staff in youth development programs can be adapted from *My Teaching Partner* (Downer et al., 2024; Pianta et al., 2008b) and *My Teaching Partner-Secondary* (Allen et al., 2015). These professional development programs aimed at elementary and secondary classroom teachers, respectively, use the same

domains as the CLASS. Teachers work individually with coaches, submitting videos of their teaching. After scoring the video with the CLASS, the coach guides the teacher on identifying areas of strength and strategies to strengthen areas of lower performance. The focus on staff-youth interaction and content-neutral nature of these coaching sessions ease the translation of these professional development programs from classroom teaching to youth development work. A model for how this approach to staff professional development could be scaled up in after-school programs through leadership, structural changes, and capacity building is provided by the Youth Program Quality Intervention Study (Smith, et al., 2012).

Additionally, the incorporation of staff coaching with CLASS results to improve practice in OST could possibly lead to more equitable education systems. Continuous coaching that focuses on improvement in CLASS domains provides an opportunity for staff and OST programs to strengthen not only instructional opportunities but also relationships among staff and students. This in turn can have positive ripple effects on children's social, emotional, and academic development (Allen et al., 2015; Glover et al., 2023; Jennings & Greenberg, 2009; Zan & Donegan-Ritter, 2014).

In addition to focusing training efforts on continuous quality improvement, another benefit of the use of the CLASS across the state was that it provided a common language and set of goals for 21st CCLC staff. Implementation of common language and goals across a large system, such as a connected framework of after-school programs across a state, can further enhance data-driven decisions to improve academic and nonacademic outcomes for youth (Wohlstetter et al., 2008). Further research is needed to determine whether improvements in CLASS scores in 21st CCLCs will be seen after training occurs, and, if so, whether those improvements will in turn translate into positive outcomes for youth. Overall, results of this study suggest that the CLASS can be a useful tool to pinpoint how after-school programs can promote positive youth development through high-quality interactions between staff and students. The tool could be one part of a more comprehensive program evaluation that would also incorporate assessments of progress toward program goals through the use of stakeholder surveys, outcome tracking, and other methods.

References

- Afterschool Alliance. (2021). *21st CCLC is a critical source of funding for many local afterschool and summer learning programs*. Retrieved 2024 from <http://www.afterschoolalliance.org/policy21stcclc.cfm>
- Afterschool Alliance. (2023). *21st Century Community Learning Centers: Accelerating learning. Supporting families. Earning results*. <http://afterschoolalliance.org/documents/21stCCLC-Overview-2023.pdf>
- Allen, J. P., Hafen, C. A., Gregory, A. C., Mikami, A. Y., & Pianta, R. (2015). Enhancing secondary school instruction and student achievement: Replication and extension of the My Teaching Partner-Secondary intervention. *Journal of Research on Educational Effectiveness*, 8(4), 475–89. <http://doi.org/10.1080/19345747.2015.1017680>
- Ansari, A., Hofkens, T. L., & Pianta, R. C. (2020). Teacher-student relationships across the first seven years of education and adolescent outcomes. *Journal of Applied Developmental Psychology*, 71. <https://doi.org/10.1016/j.appdev.2020.101200>
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (5th ed., pp. 993–1028). Wiley.
- Cabell, S. Q., DeCoster, J., LoCasale-Crouch, J., Hamre, B. K., & Pianta, R. C. (2013). Variation in the effectiveness of instructional interactions across preschool classroom settings and learning activities. *Early Childhood Research Quarterly*, 28(4), 820–30. <http://dx.doi.org/10.1016/J.ECRESQ.2013.07.007>
- Cozza, B. (2023). *An alternative review of multi-age (multi-grade) teaching and learning in global communities. Global perspectives on education research, Vol. II: Facing challenges and enabling spaces to support learning*. Routledge.

- Dodd, A. T., & Bowen, L. M. (2011). 21st century community learning centers—improving the academic performance of at-risk students: A Bronx tale. *Journal of Health and Human Services Administration*, 34(1), 10–41.
- Downer, J.T., Doyle, N.B., Pianta, R.C., Burchinal, M., Field, S., Hamre, B.K., LoCasale-Crouch, J., Howes, C., LaParo, K., & Scott-Little, C. (2024). Coaching and coursework focused on teacher–child interactions during language/literacy instruction: Effects on teacher outcomes and children’s classroom engagement. *Early Education and Development*, 35(5), 1032–62. <http://doi.org/10.1080/10409289.2024.2303604>
- Garwood, J. D., Vernon-Feagans, L., & Family Life Project Key Investigators. (2017). Classroom management affects literacy development of students with emotional and behavioral disorders. *Exceptional Children*, 83(2), 123–42. <http://doi.org/10.1177/0014402916651846>
- Gayfer, M. (Ed.). (1991). *The multigrade classroom: Myth and reality. A Canadian study*. Canadian Education Association.
- Gersh, A., Smith, C., Garner, A., & Macleod, C. (2017). *Oklahoma 21st Century Community Learning Centers statewide evaluation report: 2015-2016 annual report*. The David P. Weikart Center for Youth Program Quality.
- Glover, T. A., Reddy, L. A., & Crouse, K. (2023). Instructional coaching actions that predict teacher classroom practices and student achievement. *Journal of School Psychology*, 96, 1–11. <https://doi.org/10.1016/j.jsp.2022.10.006>
- Hammer, P. C., & Whisman, A. (2017). *21st Century Community Learning Centers 2015-2016: A descriptive evaluation*. West Virginia Department of Education Office of Research, Accountability, and Data Governance.
- Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children’s development. *Child Development*, 85(3), 1257–74. <http://doi.org/10.1111/cdev.12184>
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, 76(5), 949–67.
- Hindman, A. H., & Wasik, B. A. (2013). Vocabulary learning in Head Start: Nature and extent of classroom instruction and its contributions to children’s learning. *Journal of School Psychology*, 51(3), 387–405. <http://dx.doi.org/10.1016/j.jsp.2013.01.001>
- James-Burdumy, S., Dynarski, M., Moore, M., Deke, J., Mansfield, W., Pistorino, C., & Warner, E. (2005). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers program. Final report*. U.S. Department of Education.
- Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence relation to student and classroom outcomes. *Review of Educational Research*, 79(1), 491–525. <https://doi.org/10.3102/0034654308325693>
- Johnson, D. W., & Johnson, R. T. (1994). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Allyn and Bacon.
- Johnson, A. D., Markowitz, A. J., Hill, C. J., & Phillips, D. A. (2016). Variation in impacts of Tulsa pre-K on cognitive development in kindergarten: The role of instructional support. *Developmental Psychology*, 52(12), 2145–58. <http://dx.doi.org/10.1037/dev0000226>
- Johnson, J., Skoglund, B., & Smith, A. (2022). *Annual report 2019-2020: Nebraska 21st Century Community Learning Centers*. University of Nebraska Medical Center Munroe-Meyer Institute.

- Kim, S., Raza, M., & Seidman, E. (2019). Improving 21st century teaching skills: The key to effective 21st century learners. *Research in Comparative and International Education*, 14(1), 99–117. <http://doi.org/10.1177/1745499919829214>
- Kinsey, S. (2001). Multiage grouping and academic achievement. *ERIC Digest*. <https://files.eric.ed.gov/fulltext/ED448935.pdf>
- Knight, D. S., & Skrtic, T. M. (2021). Cost-effectiveness of instructional coaching: Implementing a design-based continuous improvement model to advance teacher professional development. *Journal of School Leadership*, 31(4), 318–42.
- LoCasale-Crouch, J., Jamil, F., Pianta, R. C., Moritz Rudasill, K., & DeCoster, J. (2018). Observed quality and consistency of fifth graders' teacher-student interactions: Associations with feelings, engagement, and performance in school. *SAGE Open*, 8(3), 1–11. <http://doi.org/10.1177/2158244018794774>
- McClellan, D. E., & Kinsey, S. J. (1999). Children's social behavior in relation to participation in mixed-age or same-age classrooms. *Early Childhood Research and Practice*, 1(1). <https://ecrp.illinois.edu/v1n1/mcclellan.html>
- McCormick, M. P., & O'Connor, E. E. (2015). Teacher-child relationship quality and academic achievement in elementary school: Does gender matter? *Journal of Educational Psychology*, 107(2), 502–516. <https://doi.org/10.1037/a0037457>
- Milner, H. R. IV. (2012). Beyond a test score: Explaining opportunity gaps in educational practice. *Journal of Black Studies*, 43(6), 693–718. <http://doi.org/10.1177/0021934712442539>
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008a). *Classroom Assessment Scoring System Manual K-3*. Teachstone.
- Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008b). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, 23(4), 431–51.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.
- Sheldon, J., & Hopkins, L. (2008). Supporting success: Why and how to improve quality in afterschool programs. *Insight: Lessons Learned from the CORAL Initiative*. James Irvine Foundation and Public/Private Ventures.
- Smith, C., Akiva, T., Sugar, S. A., Lo, Y. J., Frank, K. A., Peck, S. C., Cortina, K. S., & Devaney, T. (2012). *Continuous quality improvement in afterschool settings: Impact findings from the Youth Program Quality Intervention study*. Forum for Youth Investment.
- Thorpe, K., Rankin, P., Beaton, T., Houen, S., Sandi, M., Siraj, I., & Staton, S. (2020). The when and what of measuring ECE quality: Analysis of variation in the Classroom Assessment Scoring System (CLASS) across the ECE day. *Early Childhood Research Quarterly*, 53, 274–86. <https://doi.org/10.1016/j.ecresq.2020.05.003>
- U.S. Department of Education. (2020). *21st Century Community Learning Centers (21st CCLC) analytic support for evaluation and program monitoring: An overview of the 21st CCLC performance data: 2018-2019* (15th report). Washington, DC.
- U.S. Early Childhood Learning and Knowledge Center. (2021). *National overview of grantee CLASS scores by year*. U. S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start. <https://eclkc.ohs.acf.hhs.gov/data-ongoing-monitoring/article/national-overview-grantee-class-scores-year>
- Vincent, S. (Ed.). (1999). *The multigrade classroom: A resource handbook for small, rural schools. Book 1: Review of the research on multigrade instruction*. Northwest Regional Educational Lab. Portland, OR.
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Harvard University Press.

- Weiss, H. (2013). Fifteen years of evaluation of 21st Century Community Learning Centers: A driver for program quality and capacity in the field. In T. K. Peterson (Ed.), *Expanding minds and opportunities: Leveraging the power of afterschool and summer learning for student success*. <https://www.expandinglearning.org/expandingminds/article/fifteen-years-evaluation-21st-century-community-learning-centers-driver>
- Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness and School Improvement*, 19(3), 239–59. <https://doi.org/10.1080/09243450802246376>
- Wu, H. C. J., & Van Egeren, L. A. (2023). *Michigan Nita M. Lowey 21st Century Community Learning Centers evaluation: 2021-2022 annual report*. Michigan State University.
- Zan, B., & Donegan-Ritter, M. (2014). Reflecting, coaching and mentoring to enhance teacher–child interactions in Head Start classrooms. *Early Childhood Education Journal*, 42(2), 93–104. <https://doi.org/10.1007/s10643-013-0592-7>