

Year 3 Evaluation Summary

Prepared by Meagan Henry, PhD, and Ariana Riccio Arista, PhD

The following report summarizes EDC’s evaluation of NSF-funded Game Design and Employment Workshops. Reflections across all three years of workshop implementation are incorporated throughout the report with an emphasis placed on summative findings and recommendations following Year 3. While each year of evaluation data for the workshops cannot be directly compared due to curriculum iteration and staffing changes, an overview of how the workshops evolved and lessons learned will be presented as part of this report.

Table of Contents

Successes and High-level Recommendations	2
Evaluation Questions and Adaptations Across Project Years	3
Table 1. Evolution of Evaluation Goals	3
EQ1. Was training effective to prepare staff to achieve student learning objectives?	4
Evaluation Takeaway	4
Background	4
Methods	6
Results	6
Table 2. Mean (Standard Deviations) Ratings from Post-training Staff Surveys Measuring Staff Feelings of Preparedness to Achieve Workshop Learning Objectives	7
Table 3. Quantitative Staff Survey Responses Assessing Pre-workshop Training Effectiveness in Staff Training Goals	9
EQ2. Did available resources and ongoing support enhance staff success?	10
Evaluation Takeaway	10
Background	10
Methods	11
Results	11
Table 4. Ratings of Daily Debriefs—Learning Objectives	12
EQ3. Were there changes in Computational Thinking as a result of student workshop participation?	12
Evaluation Takeaway	12
Background	13
Methods	14
Results	16
Figure 1. Year 3 Pre- and Post- Student CT Assessment Scores (n = 34) and Maintenance Scores (n = 33)	16
Closing Recommendations	18
References	20

Successes and High-level Recommendations

Plan for Flexibility and Ongoing Support Throughout Workshops

- Leveraging technology in the moment (Slack, Zoom breakout rooms) and allowing time for daily debriefs among teaching staff was beneficial in achieving student outcomes.

Invest in Neurodiversity-Affirmative Practices

- Positive outcomes in staff teaching practices and student self-advocacy were observed following this workshop's commitment to staff training that fostered a community of inclusion and positive neurodivergent (ND) identities.

Provide Ample Time to Achieve Learning Objectives

- The workshops took on a number of learning objectives to affect participant technical and pre-professional skills over a short period of time. While staff were successful in fostering positive ND identities and supporting students in the moment, time to achieve all workshop goals was limited across all three years of implementation.

Prioritize Staff Hiring and Training Practices

- Over the three years, successes came from hiring and training staff with strong abilities upholding neurodiversity-affirming principles. Future workshops also would benefit from this approach of prioritizing staff identities, prior knowledge, and training for working with neurodivergent learners.
- By Year 3, significant investments were made in developing a specific staff training protocol to address student needs. Staff were empowered to create supportive learning environments, which had a positive effect on many of the workshop learning objectives. However, time and staff expertise was more limited in other more technical areas such as computational thinking.

Evaluation Questions and Adaptations Across Project Years

Taking a formative evaluation approach allowed evaluators at EDC to continuously provide feedback to the project to improve the design of the game design workshops and staff training. As a result, the evaluation questions established in Year 1 (Y1) of the project evolved over Year 2 (Y2) and Year 3 (Y3), and required adaptations in research design and measurement to ensure findings would be useful to the project (Table 1).

Evaluation in Y1 examined student assessments of computational thinking (CT) and staff feedback around diversity profiles and perceptions of student learning across Y1 learning objectives (LOs). As the workshops evolved, the focus of the evaluation broadened to adapt to changes in workshop design, including virtual delivery of the workshops, expansion of student learning objectives, changes in staff recruitment and partnerships, and shifts in staff training goals (STGs). By the end of the Y1 workshop, it became clear that intensive investments in staff recruitment, training, and ongoing support were critical to hosting successful workshops and thus our evaluation goals shifted to capture these important adaptations.

Evaluation for Years 2 and 3 reflected those changes by capturing additional staff feedback across updated learning objectives and adjustments to student assessments of computational thinking. The evaluation plan was adapted to assess (1) how comprehensive staff support (in addition to diversity profiles proposed in Y1) supported student learning, (2) how effective additional staff training modules were in preparing staff members, and (3) changes in student CT skills as a result of workshop participation.

Table 1. *Evolution of Evaluation Goals*

Focus Areas	Proposed Evaluation Questions (Y1)	Final Report Evaluation Questions	Data Sources
Supporting Student Diversity & Neurodiversity-Affirmative Instruction	Do instructors believe strategies [based on attentional profiles] are effective for engaging diverse students?	EQ1. Was training (including staff training goals) effective to prepare staff in achieving student learning objectives?	Staff survey ratings and open responses
Staff Training to Support Workshop Implementation	Do instructors implement the “diversity blueprints” with fidelity?	EQ2. Did resources available (diversity profiles) and ongoing support (daily debriefs,	Staff survey ratings and open responses

		Slack communications) enhance staff success?	
Computational Thinking Skills	Does participation in an informal STEM curriculum with “diversity blueprints” lead to improvements in computational thinking?	EQ3. Were there changes in computational thinking as a result of student workshop participation?	CT assessment scores; staff survey open responses; workshop video observations

EQ1. Was training effective to prepare staff to achieve student learning objectives?

Evaluation Takeaway

Findings across staff surveys highlight the limitations of quantitative surveys for this type of measurement. While staff provided generally high ratings throughout measurement points, qualitative survey responses were critical to evaluation and training revisions across all years. When given space to respond qualitatively, staff consistently provided areas for improvement and acknowledged student progress not reflected in survey ratings.

The emphasis placed on expanding and improving staff training across Years 1–3 proved successful as reflected in high ratings and staff comments, particularly around self-advocacy and principles of neurodiversity-affirmative education. Staff feedback highlighted the multifaceted challenge of training staff to work with neurodiverse learners in a complex way. Teaching these workshops called for a variety of skill sets in addition to working with neurodiverse learners, including teaching in online environments and focusing activities on specialized STEM content such as computational thinking with technical tools for game design. The feasibility of a training program to tackle wide-ranging LOs as well as staff training goals (STGs) proved challenging for some when staff were hired without backgrounds in the content area or prior experience working with neurodiverse learners.

Background

Staff training over Years 1, 2, and 3 was adjusted in response to the needs of learners and the staff delivering the workshop, with increased training to best teach and support neurodiverse learners. Learning objectives were developed to specifically address the needs of neurodiverse learners through the approach of game design workshops. An increasing emphasis was placed on developing LOs in partnership with a participatory group of autistic and non-autistic team members based on results from earlier workshop pilots. See the linked external document for a full view from the research team of [Staff Recruitment and Training Changes](#) over Years 2 and 3.

Year 1 staff training addressed delivering the 2021 Workshop LOs: game design concepts, social justice through games, social and emotional skill practice, and career exploration in game design. Training resulted in staff reporting in post-workshop surveys that they felt more prepared to teach game design concepts than social-emotional skills or social justice themes. Based on feedback from staff surveys, changes were made between Workshops 1 and 2 to address staff training needs, with some improvements in staff preparedness being observed by the evaluation team after Workshop 2.

In Year 2, staff hiring yielded a mix of experience levels among teachers and support staff. Pre-workshop training consisted of a 2-hour session and optional 2–3-hour training focused on Autism and universal design (UD) . The main training was adjusted to better ensure research goals were prioritized throughout teaching, and to address a change in the 2022 Workshop LOs. Year 2 LOs that were refined to more directly target the needs of neurodiverse learners included (1) Computational Thinking, (2) Self-advocacy, (3) Collaboration, (4) Time Management, and (5) Career Exploration. While weekly staff surveys indicated that the training supported them in achieving LOs, open-ended feedback indicated a need for additional training in working with neurodiverse learners and additional time for reviewing curriculum prior to instruction.

In Year 3, a more comprehensive process was devised for staff hiring, with a stronger focus on previous experience working with Autistic youth and practices for creating inclusive environments. Following the evidence that the Y2 training in 2022 was insufficient to prepare staff for working with neurodiverse youth, staff training was expanded in the Y3 training in 2023 to include new content and materials. Staff training was expanded to 22 hours, with a stronger emphasis on neurodiversity-affirming principles in practice. These practices, collaboratively agreed upon by the research team’s participatory group of autistic and non-autistic members, center around the following principles (O’Brien et al., 2024):

1. “Nothing about us without us”—all processes should be guided by autistic persons’ feedback, whether through participatory feedback or student feedback
2. Strengths-based approach—everyone is unique and should not be reduced to any sweeping deficit or strength
3. Value community diversity—creating opportunities to foster and bolster community diversity, collaboration, and self-advocacy strengthens our teaching
4. Value all forms of communication and engagement—we will not prioritize or encourage one form of communication over another
5. Provide multiple paths for learning (UD approach)—if a student is unable to access a workshop activity, the onus is on the staff to make it accessible via another pathway

The main learning objectives from Y2 continued to guide the curriculum: (1) Computational Thinking, (2) Self-advocacy, (3) Collaboration, (4) Time Management, and (5) Career Exploration. In addition to training for the delivery of LOs, staff pre-workshop trainings were evaluated on the following staff training goals: (1) workshop policies and procedures, (2) principles of Universal Design, (3) principles of neurodiversity-affirmative education, and (4) how to foster a respectful community of students and staff on Zoom.

Methods

The 2023, Y3 staff responses were received from 7 members (out of 9 total): 2 teachers, 2 occupational therapists, 1 social worker, and 2 counselors. They facilitated two workshops in the summer of 2023, and were surveyed three times over a period of 8 weeks: pre-workshop (after completing staff pre-training), post-workshop 1, and post-workshop 2.

Perceptions of student learning and staff preparedness to deliver the five student LOs was measured with survey items on a 5-point scale (strongly disagree–strongly agree). For each LO, staff rated the effectiveness of

1. workshop activities and instructional materials to support teaching,
2. the pre-workshop training, and
3. daily staff debriefs throughout the workshop.

The impact of staff training also was measured against the four staff training goals of the Y3 pre-workshop training. Survey items measured perceptions of staff preparedness and ability to apply knowledge from their training to practice, asking them to reflect on their feelings at different points of the workshop:

1. Following the pre-workshop training
2. Following daily debriefs
3. During the workshop while applying STGs as part of their role

Results

In Y3, staff reported above-scale-midpoint preparedness for achieving LOs in all areas after the pre-workshop training, with the greatest confidence in **Time Management** and **Self-advocacy**, and lowest confidence in **Collaboration** and **Career Exploration** (Table 2). The top two LOs also received the highest ratings of activities and teaching methods during the workshop; perceptions of the usefulness of the pre-training workshop strengthened over time, with Self-advocacy (mean score 4.83 out of 5) being the most highly rated LO in post-workshop scores while Collaboration (mean score 3.83 out of 5) continued to be the lowest. Open responses provided additional context, sometimes telling a different story from survey ratings.

Table 2. Mean (Standard Deviations) Ratings from Post-training Staff Surveys Measuring Staff Feelings of Preparedness to Achieve Workshop Learning Objectives

	Pre-Workshop 2023: Do you believe the pre-workshop training you received prepared you to achieve this learning objective with students?	Post-Workshop 2023: Do you believe the pre-workshop training you received prepared you to achieve this learning objective with students?
Learning Objective 1: Computational Thinking	4.1 (1.1)	4.33 (0.8)
Learning Objective 2: Self-advocacy	4.4 (1.3)	4.83 (0.4)
Learning Objective 3: Collaboration	3.7 (1.2)	3.83 (0.8)
Learning Objective 4: Time Management	4.6 (0.7)	4.33 (0.8)
Learning Objective 5: Career Exploration	3.8 (1.4)	4.5 (0.8)

N = 7 pre-workshops; N=13 combined responses across post-workshops

Response scale: strongly disagree (1) to strongly agree (5)

Staff open-ended responses about **Self-advocacy** reflected the high ratings, with staff reflections on the training sharing "[Self-advocacy] was covered thoroughly throughout [all] trainings and it was really helpful when self-advocacy was modeled for us in the way that we should model it to the students." This emphasis was perceived as beneficial to learning outcomes as well, shown in staff descriptions of the resulting impact on students: "I thought all the students did a very good job at advocating for themselves! They all improved a lot by the end of the workshop."

Despite low ratings, **Collaboration** was a LO that shifted in priority throughout the progression of the workshop to align with feedback from neurodivergent students and staff. Incorporating neurodiverse perspectives provided useful context in designing for neurodiverse learners, with workshops attempting to strike a balance between encouraging collaboration for learners in group projects (in Y1 and Y2) and activities designed without "forced collaboration" (in Y3). The challenges of collaboration were also reflected in staff comments as "hard getting students to participate or collaborate," with some difficulties in emphasizing student

collaboration due to “format challenges (over Zoom) and some time constraint challenges (having to meet quite a lot of learning objectives in quite a short time).”

In terms of the design of the workshop, staff also noted that “there were opportunities for students to work together but more often they did not choose to collaborate so not much opportunity for divvying up group work,” suggesting potential learner differences and preferences for this group. Creating chances for students to work together was “something [staff] realized early in the workshop and made adjustments to fit the need,” but many felt that this aspect was lacking, and “more opportunities could be added.” However, an occupational therapist noted positive student progress amid challenges, stating, “it was so nice seeing students get to know each other more and bond with each other! I loved seeing how comfortable they got as the workshop progressed.”

Staff feedback on **Computational Thinking** showed areas for improvement, despite scores over 4 points both before and after the workshop. See EQ3 for a full discussion of CT measures, including staff perceptions.

Positive experiences during the workshop also may have increased overall survey ratings, as many described positive outcomes and noted that they received support to effectively deliver LOs. For example, despite some staff noting a lack of emphasis on **Career Exploration** in the pre-workshop training, open responses about this LO were among the strongest by the end of the workshop, with much praise for the impact of visits from industry speakers who talked with students. One teacher said, “I didn't feel fully prepared to lead those initially but I think they went well and provided the opportunity for students to start thinking about careers they're interested in.”

As shown in Table 3, feedback consistently improved across all staff training goals, from the beginning of the workshop (when surveyed following the pre-training) to the end of the workshops (when surveyed post-training).

Table 3. Quantitative Staff Survey Responses Assessing Pre-workshop Training Effectiveness in Staff Training Goals

Question	Pre-Workshop After Pre-Training Mean (SD) n=9	Post-Workshop Mean (SD) n=13, combined data of Workshops 1 & 2
Staff Training Goal 1: Workshop Policies and Procedures		
I feel/felt informed and knowledgeable about workshop policies and procedures following my pre-workshop training.	4.33 (1.32)	4.77 (0.44)
I believe I will be able to/I was able to apply workshop policies and procedures during the workshop in my staff role.	4.44 (1.01)	4.92 (0.28)
Staff Training Goal 2: Principles of Universal Design		
I feel/felt informed and knowledgeable about the principles of Universal Design following my pre-workshop training.	4.33 (1.32)	4.77 (0.44)
I believe I will be able to/I was able to apply the principles of Universal Design during the workshop in my staff role.	4.33 (1.32)	4.67 (0.50)
Staff Training Goal 3: Principles of Neurodiversity-Affirmative Education		
I feel/felt informed and knowledgeable about the principles of neurodiversity-affirmative education following my pre-workshop training.	4.44 (1.33)	4.92 (0.28)
I believe I will be able to/I was able to apply the principles of neurodiversity-affirmative education during the workshop in my staff role.	4.44 (1.33)	4.92 (0.28)
Staff Training Goal 4: How to Foster a Respectful Community of Students and Staff on Zoom		
I feel/felt informed and knowledgeable about how to foster a respectful community of students and staff on Zoom following my pre-workshop training.	4.44 (1.33)	4.92 (0.28)
I believe I will be able to/I was able to foster a respectful community of students and staff on Zoom during the workshop in my staff role.	4.44 (1.33)	5.00 (0.00)

Response scale: strongly disagree (1) to strongly agree (5)

These results suggest that as staff gained experience working with students, their sense of their own knowledge and confidence grew. Given the efforts made to provide support for staff throughout the workshop in many of these target areas, these high ratings also may be related to staff experiences, which are reported in detail in the following section.

EQ2. Did available resources and ongoing support enhance staff success?

Evaluation Takeaway

Flexibility, ongoing staff support throughout workshops, and adopting a neurodiversity-affirmative framework throughout workshop development were critical to workshop success for both students and staff across all three years of implementation. The effort placed on resources, such as diversity profiles, and modes of support such as daily debriefs and just-in-time communication through Slack proved effective and highly beneficial to staff.

Background

Across all three years of the workshops, staff coordinating the workshops provided consistent support to practitioners such as teachers, tech counselors, and social workers. While some of this support came in the form of resources such as diversity profiles, ongoing communication over Zoom and Slack also was a significant part of the workshop experience. Daily debriefs among staff at the end of the workshop sessions to troubleshoot and discuss improvements to instruction were present across all three years. Staff also used messaging via Slack to communicate in the moment and coordinate between one another (i.e. asking a social worker or tech counselor to work individually with a student in a breakout room).

Diversity profiles or “diversity blueprints,” as they were called in the initial proposal, were a major focus of the project team’s plan for supporting students’ individual strengths, goals, and support needs. Over the three years of this project, the composition of diversity profiles evolved based on staff feedback and workshop goals. In Year 1, staff were provided with a description of each students’ strengths and challenges with a focus on areas for additional support. While comprehensive, staff feedback indicated that these profiles were too lengthy to effectively reference in the moment and they were used primarily by the social work staff in the event of a major issue. Profiles were adapted in Year 2 to be considerably more brief and easy to reference. However, staff feedback again indicated that diversity profiles were interesting and informative to read at the beginning of each workshop, but they were not often referenced.

The most significant iteration to diversity profiles occurred in Year 3. Y3 diversity profiles took the form of introductory slides where students were asked to share communication styles and

learning preferences, interests, and other personal information prior to the workshop. In 2023, students made their own diversity profiles to introduce themselves, compared to previous years when students did not select the information shared. This led to more frequent use of diversity profiles that were aligned to students' own identities.

Daily debriefs were delivered similarly across all three years, with time set aside at the end of workshop sessions for staff to reflect on the day's events and plan how to best support individual student needs. Slack appeared to be used consistently and successfully across all three years of the workshops.

Methods

The same participants and data set described above in EQ1 informed EQ2. Quantitative feedback was collected measuring the impact of daily debriefs as they related to each of the five LOs and four STGs (phrased as, "Do you believe the daily staff debrief sessions prepared you to achieve this learning objective with students?"). Staff also were given open-response opportunities to provide feedback on the support they received throughout the workshop, how effective they were in collaborating with their peers, and thoughts about the resource of student diversity profiles. Qualitative data were examined for themes relating to support across questions focused on debriefs, as well as in responses to other questions where support and resources were mentioned.

Results

Support throughout the workshops was among the most often mentioned piece of feedback, with an overwhelmingly positive response. In terms of modes of support offered, the use of student diversity profiles was among the most highly rated across all three years. The Y3 student-centric iteration of diversity profiles was well-received by staff and highlights a further commitment to neurodiversity-affirmative practices and an emphasis on working with individual student needs. Staff responded positively to this approach, saying, "I like how students are given the option to be creative and share as much or as little as they want."

Staff reported varying levels of application during the workshop. Nearly all staff intended to use the profiles in the pre-workshop survey, but many reported in open responses post-workshop that it was difficult to use them in the moment because of the demands of teaching. Feedback also was provided that the format and amount of information was not easy to access and understand while also facilitating workshop curriculum activities.

Daily debriefs on the topics of Self-advocacy (4.69) and Time Management (4.54) were those rated most effective. While Collaboration was the lowest rated LO among debrief topics (4.23), daily debriefs still proved to be the most effective way for teachers to approach this challenge

in practice when compared to ratings of activities and materials for instruction (4.15) and pre-training (3.92).

Table 4. Ratings of Daily Debriefs—Learning Objectives

	LO1 debriefs score	LO2 debrief score	LO3 debriefs score	LO4 debriefs score	LO5 debriefs score
N	13	13	13	13	13
Mean	4.38	4.69	4.23	4.54	4.46
Median	5	5	4	5	5
Standard deviation	0.768	0.751	0.832	0.660	0.877
Minimum	3	3	3	3	3
Maximum	5	5	5	5	5

(LO1: CT; LO2: Self-advocacy; LO3: Collaboration; LO4: Time Management; LO5: Career Exploration)

This suggests that the daily debriefs offered continued support for a topic for which staff were already well-prepared (Self-advocacy) and just-in-time support for topics that some may have struggled to apply in the workshop (Time Management and Collaboration with Zoom practices to support learners). One staff reflection sums this up nicely, stating, “The prep and debrief session helped organize and keep track of ways to support students each day, like check-ins and things to just be aware of.”

Staff reported high ratings across STGs after pre-training; however, ratings increased across all STGs post-workshop after having worked with students and received ongoing support (see Table 2, above). The STG with the greatest gain from pre- to post-workshop was staff ability to foster a respectful community of students and staff on Zoom.

Among staff open responses, many also described using the Slack channel in the moment for supporting each other and students, and appreciating staff ability to “pop in” across Zoom breakout rooms to assist with a variety of needs. Slack was described as “a great platform!” and an assistant teacher said, “the use of Slack was critical in allowing for dynamic shuffling of responsibilities and allowed for reactive teaching, where students could get individualized support quickly and efficiently.”

EQ3. Were there changes in Computational Thinking as a result of student workshop participation?

Evaluation Takeaway

Although no significant changes in CT assessment scores were found across the years of the workshop, findings from staff survey responses and video observations of workshop units related to CT can provide additional insight into how CT was taught as a subject and

understood by students in Y2 and Y3. Our team experienced challenges when measuring CT using standardized quantitative assessments across all years despite iterating our assessment methods. This included experiencing ceiling effects for students who appeared to enter the workshops with prior CT knowledge and high-level reasoning skills. Given the limitations of our CT assessment, the evaluation team conducted video observations of CT learning and how teachers incorporated CT as a LO and checked for understanding, which proved to be among the most effective approaches for understanding CT in action across the workshop.

Background

Computational Thinking was a focus of the workshop curriculum across all three years, with game design activities intended to teach problem solving and critical thinking skills through iterative design, playtesting, and constructive feedback. Students were tasked with creating a game using open-access Web-based design software (differing across years) and games varied widely across students in terms of completeness and complexity. The workshop did not intend to teach coding or programming skills directly.

EDC's evaluation of the workshops investigated CT as a LO, measuring learning outcomes related to the workshops and staff perceptions and practices around CT. CT is most frequently defined as the concepts, practices, and perspectives for problem-solving that are derived from computer science (Wing, 2006; Yadav et al., 2014). For this project, CT was focused on generally agreed-upon main topics of abstraction, decomposition or breaking bigger problems into smaller problems, developing algorithms, applying conditional logic, and iterative testing and debugging (Aho, 2011; Grover & Pea, 2013; Tedre & Denning, 2016). CT was measured in multiple ways, including: (1) assessments administered at pre-, post- (Years 1–3) and maintenance timepoints (Y3 only), with interviews to describe problem-solving strategies during the tests (in Y3 only); (2) surveys of staff perceptions of learning; and (3) through observations of approximately 20% of workshop video footage (in Y2 and Y3 only).

Assessments

CT was measured for all participants in Years 1–3 primarily using adapted versions of the Computational Thinking Abilities—Middle Grades Assessment (CTA-M) (Wiebe et al., 2019). The CTA-M assessment remained the CT measure of choice throughout the study, with a strong research base for this age group. The CTA-M provided items focused on sequencing and debugging tasks across concept areas including loops, conditionals, and functions. In all three years using this measure, ceiling effects were exhibited for some learners, with little change in CT abilities from pre- to post-workshop.

In Y1, CT was assessed using an 8-item subset of CTA-M, at pre- and post-test. Y1 used 16 items total, with students randomized to 8-item A and B versions; however, results had low

internal consistencies across the two sets, some of which appeared to be due to ceiling effects. Low correlations between items purportedly measuring similar constructs raised questions about the measure's reliability.

A new iteration of the CTA-M was administered in Y2. All participants received the same 10-item subset of the CTA-M at pre- and post-test, with the same focus areas; however, the Y2 version removed items that exhibited ceiling effects in 2021 and added items focused on functions, which were previously excluded due to their complexity. Additionally, as a part of the CT measure, students were asked to share how they arrived at the answer they selected. With these modifications, the CT measure exhibited acceptable internal consistency. However, improvements in CT were not observed in either 2021 or 2022 and ceiling effects were again observed for some students.

To address ceiling effects, additional items with increased complexity were added in Y3, including more difficult items from a CT measure for high school students and verbal explanations of assessment responses. These items, from Part 4 of the Exploring Computer Science curriculum (ECS) (Weintrop & Wilensky, 2015), focused on programming and included tasks in comparing algorithms, the concept of a variable, creating a program, and comprehending code. For these items, learners responded to multiple choice questions and open response prompts to show their thinking. As described above in staff feedback, teacher training efforts in CT also were adjusted in response to Y1 and 2 outcomes, including the development of more applied strategies to practice computational thinking in Y3.

Surveys—Staff Perceptions of CT Learning

Staff surveys were administered every year at pre- and post-workshop, as described in EQ1. Questions related to CT were asked in a section about the learning objective, asking staff to reflect on the effectiveness of their own instruction, the pre-training, and daily debriefs in their ability to achieve the learning objective with students. These ratings questions were followed by an open-response question asking, “In a few sentences, please provide support for your above ratings.”

Workshop Video Observations

In Y2 and Y3, workshops were recorded via Zoom. In Y2, sessions were recorded with only the main room captured, not breakout rooms. In Y3, all sessions of both workshops and all breakout rooms were recorded.

Methods

In Y3, learners (n = 33) ranged in age from 14–22 years old at the start of the workshops (M = 17.1, SD = 2.49) with 85% identifying as male. The racial makeup of the group was 30% Mixed

Race, 30% White/Caucasian, 18% Hispanic/Latiné, 12% Asian, and 9% Black or African American. Learner background in terms of family income was also diverse, with 18% at the poverty level (less than \$25,000), 21% low income (between \$25,000-\$74,000), 27% middle income (between \$75,000-\$149,000) and 15% upper middle income (between \$149,000-\$299,000). No learners reported family income over \$300,000 and the remaining 18% preferred not to answer the family income question.

Assessments

Year 3 learners in Workshops #1 and #2 were administered the CT assessment at four collection points: pre-workshop, post-workshop, one month post-workshop, and 6 months post-workshop as a maintenance measure. Learners were given the adaptation of the CTA-M, Part 4 of the ECS, and asked to provide verbal explanations of their answer choices in interviews with researchers. Analyses of CT assessment scores included comparisons of scores pre- to post- and differences between learners. Approximately 20% of interview transcripts with learners from Y3 workshops 1 and 2 also were reviewed in an effort to better understand learners' thought processes, but as students did not provide enough information for evaluators to assess their understanding of the CT prompt, the interview data were not included in the reporting of results.

Surveys—Staff Perceptions of CT Learning

The same staff participants and data set described above in EQ1 informed EQ3. The open-response question about computational thinking as a learning objective (“In a few sentences, please provide support for your above ratings”) was examined to better understand how staff in Y3 perceived student learning around CT.

Workshop Video Observations

Video footage of the game design workshops was reviewed and annotated for instances of computational thinking instruction from staff or demonstrations of CT from students. Instances of CT were defined for the purposes of video review as use of explicit vocabulary (e.g., algorithms, loops, debugging), definitions (delivered in lectures and repeated in other contexts), problem solving strategies (e.g., “break it down,” “test it out”) and related mindsets (e.g., rule systems, abstraction, design process). Thematic analysis of the workshops was developed from descriptive annotations and notable quotes from staff and students to illustrate the nature of their discussions related to CT. It is important to note a critical limitation of the methodology in that it captured only information from students who expressed themselves verbally or who had a chat read aloud by a staff member. However, we felt that observation of teaching methods in particular was a useful addition to the evaluation plan for understanding how CT was taught across both years.

Approximately 20% of recorded time (12 hours out of 60 total) in Y2 was reviewed by one researcher. However, given that only one hour of the workshop was designated as focused on CT according to the agenda, other topic areas adjacent to CT were chosen to supplement the review, including game design basics, playtesting, troubleshooting, game mechanics collaborative exercise, and the final showcase. In Y2, only one instructor’s breakout room recordings were available for review. Thus, additional footage from other instructors working individually with students was not included in considerations of student demonstrations of understanding.

For Y3 recordings, 23% of recorded time (25 hours out of 108 total) was reviewed by one researcher, with 15 hours focused on instructional time with staff across two workshops delivered consecutively and approximately 10 hours in breakout rooms with students participating. The detailed course agenda in Y3 provided guidance for selecting videos with CT topics as part of the learning objective (e.g., computational thinking, debugging as part of troubleshooting), which accounted for 10 hours of instruction and breakout room activities reviewed compared to one hour in Y2. The additional topics chosen to supplement the review were aligned to those chosen from available Y2 recordings, including game design, troubleshooting, and gameplay (playtesting). Breakout sessions were reviewed across four different staff members leading the discussions (one of whom was a tech counselor helping with troubleshooting) to get a representative sense of teaching styles and the breadth of conversations taking place with students.

Results

Findings related to CT are reported below, with assessments and survey results reflecting Y3 only, while video observations provide results for Y2 and Y3.

Assessments

As in previous years, in Y3 ceiling effects were observed in participant scores where learners who began the program with high test scores in CT continued to receive high test scores across the additional time points (Figure 1).

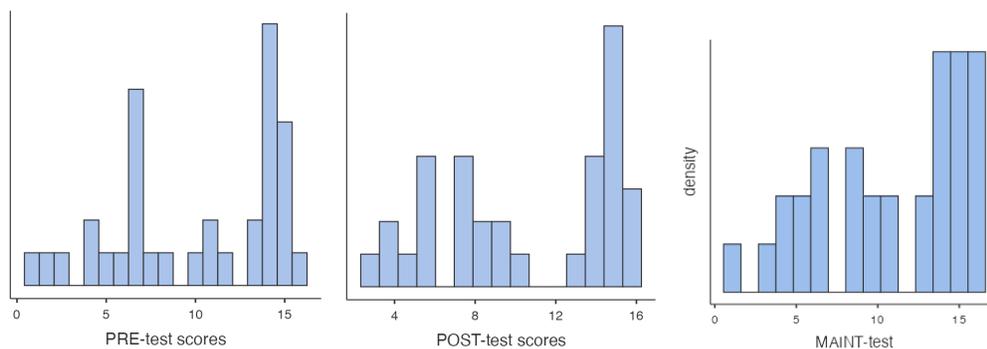


Figure 1. Year 3 Pre- and Post- Student CT Assessment Scores ($n = 34$) and Maintenance Scores ($n = 33$)

While upward trends were observed, there were no statistically significant changes between pre- and post- assessment scores. Analyses of CT scores compared to demographic differences and prior experiences were not found to be statistically significant. Further analyses were conducted to examine survey data collected by the research team for any correlational findings. Items about students' prior experience, knowledge in game design, and feelings of self-efficacy were analyzed, but did not yield significant results against CT scores.

Cronbach's alpha and omega statistics to assess reliability of the CTt only and the CTt plus the ECS Part 4 measure showed acceptable reliability at all testing time points (Baseline, Pre-, Post-, and both Maintenance points) with α s between .8 and .89 and ω s between .81 and .89.

Surveys—Staff Perceptions of CT Learning

Staff feedback on computational thinking showed areas for improvement despite scores over 4 points on a 1–5 scale, both before and after the workshop. Some staff described success among students acquiring “computer skills” and noted “peer to peer tech help” as evident of this learning outcome, while others noted, “We did go through some tech-related things and the whole training took place on computers, but I don't think we used computer logic really,” and “Computational thinking could have been more highlighted. Programming had a spot in the workshop, but not as much as I think was necessary.” These responses reflect CT as a developing area within teacher training, with varied definitions and understanding of topics within the discipline.

Given the complex nature of teaching CT and the varied backgrounds of staff, there was a range in responses regarding preparedness. Some staff shared a lack of confidence in their ability to deliver the subject matter on game design, with an occupational therapist (OT) describing it as “hard to focus on things I don't really understand (game design),” while another OT felt that they would “be able to provide ample support to the tech counselors and teachers when helping students improve their computational thinking.” This was also reflected in broader reflections on what staff would change for the future, with one social worker saying, “I would have spent more time learning how to use Bitsy so I could have helped with game design,” and an OT stating that they would “try to understand more of the actual content in itself so I could be more helpful in regards to workshop activities.” While social workers and OTs were not expressly hired to teach CT skills, these reflections indicate that some foundational knowledge of CT concepts would be helpful for these staff members to more fully assist their students in learning new material.

Workshop Video Observations

Overall, despite having less time specifically allotted for CT instruction, staff in Y2 incorporated CT content into their curriculum more intentionally and frequently than did staff in Y3. This included explicit instruction such as incorporating CT vocabulary, or tying examples and lectures back to previous points. For example, in a demonstration of MakeCode Arcade (the software used by some learners in Y2), an instructor narrated their process as it related to CT, saying, “I don’t need all of this information. This is where the idea of abstraction comes in because I’m only taking the most important parts. So I’m going to focus on a sprite ... and set my background image.” Instructors in Y2 also incorporated CT metacognitive strategies into instruction, such as describing how to break problems down into smaller tasks when working with students in a breakout room. During times when it was more challenging to manage the class discussion (e.g., students going off topic, arguing, interventions from social workers), instruction was shorter with missed opportunities to incorporate CT.

By contrast, Y3 staff tended to use more pragmatic language and approaches when working with content, rather than terms and strategies related to CT (e.g., focusing on programming and point-and-click solutions rather than getting into specifics of loops, conditions, and debugging when opportunities were available). While more missed opportunities to tie back instruction to CT concepts were observed in Y3, staff created opportunities to highlight other learning objectives related to neurodiversity-affirming principles. For example, in units focused on troubleshooting and debugging, a stronger emphasis was placed on emotional regulation and career exploration than CT. This was seen in moments of frustration when staff would remind students to take a break if needed and to try focusing on the emotional experience of playing a game and the “feeling you get” when finally mastering something challenging.

Students in Y2 and Y3 came in with varied backgrounds in CT and programming or game design, making observations of their understanding difficult to directly compare. Perhaps as a result of the explicit instruction in Y2, staff created a neurodiversity-affirming opportunity for a student to share clear examples of their own thinking and processes for navigating the world as it related to CT. While these types of discussions were not observed in Y3, students in those workshops exhibited CT skills during problem-solving activities, such as making suggestions related to decomposition (e.g., “go back to basics and look at what something is made of”) during debugging. Additionally, whereas in Y2 staff offered CT metacognitive strategies for problem solving, in Y3 students suggested using them. In both Y2 and Y3 students did not use explicit CT vocabulary when demonstrating understanding, opting instead for pragmatic language directly related to software being used. For a full reporting, see the external document of [video observations findings](#).

Closing Recommendations

The series of workshops hosted by this project aimed to address important topics recognized as critical to the success of autistic students in post-secondary education and the workplace. Key lessons were learned across three years of workshop implementation. Findings and recommendations aligned with our three evaluation questions are outlined below.

Evaluation Questions	Findings & Recommendations
<p>EQ1. Was training (including STGs) effective to prepare staff in achieving student learning objectives?</p>	<p>Positive outcomes in staff teaching practices and student self-advocacy were observed following this workshop’s commitment to staff training that fostered a community of inclusion and positive neurodivergent identities in Year 3. Staff commitment to inclusive practices was meaningful for students and a prime example of the value of neurodiversity-affirming practices in education.</p> <p>Findings across staff surveys highlight the limitations of quantitative surveys for evaluators when assessing constructs such as staff preparedness. While staff provided generally high ratings throughout measurement points, qualitative survey responses were critical to evaluation and training revisions across all years. When given space to respond qualitatively, staff consistently provided areas for improvement and gave important examples of workshop successes not reflected in survey ratings.</p>
<p>EQ2. Did resources available (diversity profiles) and ongoing support (daily debriefs, Slack communications) enhance staff success?</p>	<p>Hiring practices and curriculum design valued the lived experiences as neurodivergent educators, and the effort placed on modes of support such as daily debriefs and just-in-time communication through Slack proved significant. This was especially true for topics such as collaboration, that proved challenging in practice and were rated lower in the pre-training.</p> <p>Teaching these workshops called for a variety of skill sets in addition to working with neurodiverse learners, including teaching in online environments and focusing activities on specialized STEM content such as computational thinking with technical tools for game design. Providing adequate training and time to tackle wide-ranging student LOs and staff training goals proved challenging for some when staff were hired without backgrounds in the content area or prior experience working with neurodiverse learners. For example, one staff member said, "This is the first time I have worked with neurodiverse teenagers and I felt a little</p>

	reserved sometimes especially because I have no knowledge on game design either."
<p>EQ3. Were there changes in computational thinking as a result of student workshop participation?</p>	<p>Challenges in CT measurement as well as available time for CT instruction may have affected overall evaluation results. Ceiling effects, differing student knowledge at baseline, and the need to alter CT quantitative assessments methods across Years 1–3 provided challenges in our ability to detect change in CT knowledge and skills across the intervention period. Given the pre-/post- differences among learners with high prior knowledge in CT, baseline knowledge may be a stronger predictor of students' performance on game design tasks rather than participation in this workshop. Also, considering the multifaceted approach to curriculum development and staff training, which focused on self-advocacy and pre-employment skills, the topics assessed by the CTA-M and ECS Part 4 may not have been aligned with the CT content taught.</p> <p>As observed in workshop videos, staff varied in their approach to teaching CT concepts (for example, using more explicit, intentional language and metacognitive strategies in Y2 compared to pragmatic language and hands-on approaches in Y3). Given that the field of instruction and teacher training in CT is still growing and is characterized by a lack of common definitions among educators, CT as a learning objective was a considerable challenge. Further, while the CTA-M assessment used as part of this evaluation has been validated externally and reliably measured CT knowledge across learners in this study, the field of measurement in CT is also nascent, with many questions remaining around how to recognize and measure CT for diverse learners.</p>

References

Aho, A. (2012). Computation and computational thinking. *The Computer Journal*, 55(7), 832–835. <https://doi.org/10.1093/comjnl/bxs074>

Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational Researcher*, 42(1), 38–43. <https://doi.org/10.3102/0013189X12463051>

O’Brien, S., Grossman, E., Kilgallon, E., Batkin, D., Grevalle, C. D., Delos Santos, J., Dwyer, P., Arista, A. R., Kofner, B., Yan, A., Hwang-Geddes, L., Shevchuk-Hill, S., Dave, S., Gravitch, K., Tricarico, N., Hayes, A., Rico, J., Leon, B., Malik, A., ... Gillespie-Lynch, K. (2024, May 15–18).

“You took [our] feedback and made it better”: Autistic advocacy in game-design and employment workshops [Poster presentation]. International Society for Autism Research Annual Meeting, Melbourne, Australia. <https://doi.org/10.13140/RG.2.2.16404.92802>

Tedre, M., & Denning, P. J. (2016). The long quest for computational thinking. *Koli Calling '16: Proceedings of the 16th Koli Calling International Conference on Computing Education Research*, 120–129. <https://doi.org/10.1145/2999541.2999542>

Weintrop, D., & Wilensky, U. (2015). Using commutative assessments to compare conceptual understanding in blocks-based and text-based programs. *ICER '15: Proceedings of the Eleventh Annual International Conference on International Computing Education Research*, 101–110. <https://doi.org/10.1145/2787622.2787721>

Wiebe, E., London, J., Aksit, O., Mott, B. W., Boyer, K. E., & Lester, J. C. (2019, February). Development of a lean computational thinking abilities assessment for middle grades students. In *Proceedings of the 50th ACM technical symposium on computer science education* (pp. 456-461).

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. <https://doi.org/10.1145/1118178.1118215>

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational thinking in elementary and secondary teacher education. *ACM Transactions on Computing Education*, 14(1), 1–16. <https://doi.org/10.1145/2576872>