STATUS, EQUITY, AND COLLABORATION: A PILOT STUDY OF COMPLEX INSTRUCTION–INFORMED PAIR PROGRAMMING

Ian Abrash[1], Judith Canner[1], Erica Geam[2], Noah Kan[1], Charlotte Kubitz[2], Erica Lopez[1], Jett Palmer[2], Isabel Pfeiffer[2], Allison Theobold[2], Victoria Woodard[3]
[1]California State University, Monterey Bay
[2]California Polytechnic State University, San Luis Obispo
[3]University of Notre Dame
jcanner@csumb.edu

As programming becomes integral to statistics and data science education, institutions increasingly use pair programming to support students' computational learning. Yet, research has shown that collaborative programming can reproduce participation inequities, particularly for students from historically marginalized groups. This pilot study examines collaboration patterns in introductory data science courses at two universities and evaluates a revised pair programming protocol grounded in critical theories. We analyze discourse from student pairs to identify how power, authority, and status shaped their interactions. Findings from the first year reveal specific challenges in achieving equitable collaboration, which inform planned revisions to the collaborative protocol. These results highlight opportunities for structuring pair programming to promote more inclusive participation.

INTRODUCTION

Collaborative learning is widely recognized as a way to foster students' engagement and sense of belonging in higher education (Kuh, 2009). Such experiences can be especially meaningful for students from marginalized groups, who often seek connection and validation within academic communities (Ellington & Frederick, 2010; Fullilove & Treisman, 1990; Oppland-Cordell & Martin, 2015). Within computer science education, pair programming has emerged as a key collaborative practice aimed at improving learning outcomes, belonging, and retention (Salleh et al., 2011). Yet, despite its benefits, collaborative learning environments often reproduce issues of status—the unequal distribution of power among peers (Cohen & Lotan, 1995).

Conflicts over who speaks and whose voices are recognized are indicative of power and status (Johnson, 2002), where students with higher perceived status disproportionately dominate the conversational floor and have their ideas judged as higher quality (Chiu & Khoo, 2003; Esmonde, 2009; Mulryan, 1994). Studies of group discourse further reveal how racialized and gendered microaggressions constrain participation and suppress the development of students' discipline-based identities (Byun, Shah, & Reinholz, 2023; Langer-Osuna, 2011; 2016; Reinholz et al., 2022). These patterns align with critical race theory, which frames classrooms as sites where broader social hierarchies of race and power are reproduced through everyday interactions and participation.

Complex Instruction (CI), a pedagogy specifically designed to address inequitable interactions within the classroom (Cohen & Lotan, 1997), offers a promising framework for addressing these inequities. Although originally developed for K–12 classrooms, CI has recently been adopted in higher education contexts (Goyer et al., 2021; Reinholz, 2024; Reinholz et al., 2024). Unlike traditional group work, CI explicitly seeks to disrupt status hierarchies that lead to unequal participation. It does so through structured roles, group norms, and deliberate instructor attention to the relative status of individuals. CI is grounded in three guiding principles: (1) using **"group-worthy" tasks** that require multiple abilities to complete, (2) establishing **norms and roles** that promote autonomy and interdependence, and (3) **attending to individual status** through ongoing instructor support and mechanisms of group accountability (Cohen & Lotan, 1997).

This paper reports on a pilot study within a three-year NSF-funded project (DUE-2334988, DUE-2334989) aimed at integrating CI into pair programming to foster more equitable collaboration. We define "equitable" collaboration as interactions in which students have equal opportunities to contribute meaningfully to the shared work. This contrasts with equality, which assumes identical outcomes for all students. An equitable approach acknowledges that, in some circumstances, students may need different forms of support to achieve comparable opportunities to participate.

The following section outlines the data collection context, the implemented pedagogy, and our analytical approach. We then present findings on how Complex Instruction shaped equity dynamics in pair programming and describe the refinements made to the collaborative protocol in response to these results. We conclude by offering directions for future research and considering the broader implications for creating more equitable collaborative learning environments.

METHODS

Our study took place at two institutions, one Hispanic-serving institution, California State University, Monterey Bay (CSUMB), and one predominantly white institution, California Polytechnic State University, San Luis Obispo (Cal Poly). These differing institutional contexts allow for the study of the relationship between the demographics of a classroom's student body and the pedagogy which fosters equitable group collaborations. At CSUMB, 47% of the student body identifies as Latine, with 59% woman-identifying, 49% first-generation, and 52% from underrepresented minority groups. At Cal Poly, 49% of the students identify as white, 23% as Latine, 16% first-generation, and equal proportions of women and men.

Pedagogical Context

We chose to focus on introductory data science courses as these courses act as an introduction to the discipline and practices of data science and are paramount in students' decisions to pursue a major or minor in Statistics and Data Science. The courses at both institutions focus on teaching students how to program in R with an emphasis on packages within the *tidyverse* ecosystem (Wickham et al., 2019) and Quarto documents, closely following the contents of *R for Data Science* (Wickham, Çetinkaya-Rundel, & Grolemund, 2023). Specifically, students learn data importing, wrangling, visualization, joins, and additional topics such as function writing and version control.

At CSUMB, this course is taught over a 15-week semester, meeting once a week for 110-minutes and is a requirement for the Statistics and Data Science (SDS) major, Data Science minor, and an option within the Statistics minor. The course is only taken by undergraduate students and though SDS majors typically take the course within their second year, students in minor programs typically take the course in the second semester of their third year or later. This course has no computing prerequisite requirement. At Cal Poly, the course is taught over a 10-week quarter, meeting twice a week for 110-minutes. The course is required for Statistics majors, typically taken in their second year, and is the most popular course for students completing the Statistics minor. While mostly taken by undergraduates, graduate students may take this course. At Cal Poly, the course has an introductory computer science prerequisite.

Pair Programming

To promote equitable participation in collaborative programming, we adapted the established practice of pair programming, commonly used in software engineering education. Pair programming traditionally involves two students working at a single computer while alternating between the "driver," who types and manages shared resources, and the "navigator," who provides oversight and suggestions (Beck, 2000). Although this structure often supports learning, research has shown that it can also lead to inequitable interactions, including domination by the more experienced partner or uneven opportunities for contribution (Lewis & Shah, 2015). These concerns motivated the development of a modified approach designed to foster more equitable participation.

Our intervention introduced a set of role modifications aligned with Complex Instruction (CI) principles. Rather than adopting the traditional driver–navigator model, we created two new roles: **Coder** and **Developer**. The Coder led the problem-solving process by determining the solution strategy, providing instructions for what the Developer should type, and managing supporting resources (e.g., textbook, R cheatsheets). The Developer executed the Coder's instructions, ran the code, and evaluated whether the output met the requirements of the prompt. By separating decision-making from typing, these roles intentionally inverted the traditional structure to prevent the typist from dominating the activity. Pairs periodically switched roles to ensure that both students engaged in both the strategic and procedural aspects of the task.

To further support equitable participation, we designed each activity to be "group-worthy," meaning that successful completion required multiple forms of ability beyond coding proficiency. Additionally, instructors actively attended to classroom status dynamics by observing interactions, identifying status differences, and assigning competence to lower-status students. By publicly highlighting these students' intellectual contributions, instructors helped interrupt negative status expectations and promote more balanced participation.

Study Participant Demographics

For student identities, we focus on teacher-identification, rather than student self-identification, as this is closer to how implicit bias functions (Staats, Capatosto, Wright, & Jackson, 2016). At CSUMB, there were 15 of 29 total undergraduate students within the course that volunteered to participate in the study. Among the participants, 55% were perceived to be women (by the instructor) and 34% were perceived as non-white. Though 15 participants agreed to participate in the study, of the seven men, only three regularly participated in class sessions and of the eight women, six regularly attended class sessions, with the other participants missing 50% or more of the recorded class sessions. Of the three men who regularly attended, two were highly experienced programmers and of the six women, four were highly experienced programmers based on information provided by the students. Therefore, groups tended to either be predominantly woman-man, with only two planned interactions with man-man participants occurring throughout the semester and each of the woman-woman pairings interrupted by a third member, with the exception of two. The majority of the woman-man interactions were between highly experienced programmers.

At Cal Poly, 11 of the 35 students enrolled agreed to participate in the study. Among the participants, 54% were perceived to be women and 46% were perceived to be non-white. Over the duration of the course, there were very few instances of students who agreed to participate in the study missing class. Similar to CSUMB, the majority of the recorded collaborations were between two women or one woman and one man, with only two unique collaborations between two men.

Analysis Methods

We recorded and transcribed consenting student pair-programming interactions (audio and video) for discourse‑analytic evaluation (Gee, 2014). The unit of analysis was a student contribution, defined as an uninterrupted thought spoken by a member of the pair. Contributions were examined within proposal negotiation units (PNUs) (Engle et al., 2014), which begin with a proposal and end with an utterance that accepts, modifies, or rejects the proposal. Each PNU was coded for the initial proposal, any intermediary negotiations, and the final uptake (promotion) or rejection (demotion). At Cal Poly, a total of 28 interactions were coded. At CSUMB, a total of 11 representative interactions have been coded at the time of submission.

This framework allowed us to examine both the substance of students' ideas and how those ideas were positioned by their partners. Drawing on the influence framework (Engle et al., 2014), we conceptualized power and authority through patterns of interaction aligned with different forms of merit and authority. In this paper, we focus specifically on *intellectual merit*, which reflects the degree to which a student's ideas are treated as high- or low-quality. We compare students' intellectual merit scores—calculated as proposal promotions minus demotions—across different gender pairings and examine whether prior programming experience (*technological authority*) shaped the group's discourse.

Additional Data Collected

To provide additional insight into student perceptions, we used the the Student Survey of Motivational Attitudes toward Data Science (Kerby-Helm et al. 2022) to evaluate if perspectives on attitudes towards coding and programming in data science changed over the duration of the course. In addition, a survey at the end of the semester collected student feedback on pair programming structures and perception of learning throughout the course. For the purposes of this paper, we will focus on comparison of gender differences in student feedback and perceptions of pair programming from the end of semester survey

RESULTS

In the section that follows, we identify three key points that illustrate how students engaged with the study in ways that suggest our protocol may need revision. These include students' perception and buy-in for the protocol, whether they recognized and adhered to the roles outlined in the protocol, and how intellectual merit was influenced by the collaborative protocol we implemented.

Students' perceptions of and buy-in for the collaborative protocol

Of the 26 total study participants, 13 out of 15 participants, including eight women, at CSUMB and 10 out of 11 participants, including six women, at Cal Poly completed the final experience survey. In general, most participants reported agreement that pair programming activities were helpful, personally beneficial, and they felt they were able to contribute (**Table 1**). At CSUMB, the exceptions included three women with high levels of programming experience, and one man who had chronic attendance issues. At Cal Poly, the woman who reported a low ability to contribute consistently had partners who positioned her with less authority. At CSUMB, no student with chronic absenteeism that completed the survey reported that the pair programming activities were a reason for their absences. Instead they identified other reasons unrelated to the class and its structure, either in the survey or directly to the instructor, that explained their absences (e.g., conference travel, graduate school visits, family obligations).

Table 1: Students rated their experiences on a four-point scale for the following statements: (1) I benefited from working with my peers in pairs during class activities (Benefited); (2) I was able to contribute to my group during pair programming activities in class (Contributed); and (3) The pair programming activities during class helped me to learn the material (Helped).

Perception	Perceived Gender	School	Strongly Disagree	Disagree	Agree	Strongly Agree
Benefited	Man	Cal Poly			50%	50%
		CSUMB		**20%**	60%	20%
	Woman	Cal Poly			33%	67%
		CSUMB	**12%**		50%	38%
Contributed	Man	Cal Poly			50%	50%
		CSUMB			60%	40%
	Woman	Cal Poly		**17%**	50%	33%
		CSUMB			75%	25%
Helpful	Man	Cal Poly			75%	25%
		CSUMB			80%	20%
	Woman	Cal Poly			67%	33%
		CSUMB		**25%**	38%	38%

Students' recognition of and adherence to the roles outlined by the collaborative protocol

The efficacy of a cooperative learning framework depends first on whether students adopt the structure and second on how it manifests in practice. We examined *role talk*, or verbal references to the pair-programming protocol, as an indicator of adoption. Three forms of role talk emerged in our data: *swaps, re-establishments, and indirect references*. When prompted by the task, students typically verbalized a swap (e.g., "Okay, now you type.", "Okay, I'm the coder now."). Re-establishments occurred when students clarified their current role (e.g., "Uh, who am I? Am I typing or am I saying what to do?"). Indirect references appeared when students invoked the roles without prompting (e.g., "I want to type it out, but I have to, like, explain it").

Role talk appeared in 23 of 28 collaborations (82.1%) at Cal Poly and 10 of 11 (90.9%) coded collaborations at CSUMB. At Cal Poly, of the groups mentioning the roles, 14 mentioned roles at least as many times as the number of assigned swaps (60.1%) and at CSUMB, five mentioned roles at least as many times as the number of assigned swaps (50%). If adherence to the protocol is defined as mentioning roles at all, then the vast majority groups adhered. But if adherence requires mentioning the roles at least as many times as the number of swaps, then just over half of the groups adhered.

Protocol influence on role talk and intellectual merit

Figure 1 shows the relationship between the absolute difference in partners' intellectual merit (disregarding which partner had greater merit) and the number of times roles were mentioned. We hypothesized that more role talk would be associated with more equitable collaboration (i.e., smaller differences in intellectual merit). However, the data show little to no overall relationship between the two variables.

The gender composition of pairs, indicated by color and shape, instead reveals a more distinct pattern. At CSUMB, the crosses – showing the average of the two quantitative variables by groups – indicate little variability in the number of times roles were mentioned, however, at Cal Poly there is much larger variability. Woman–woman pairs mentioned roles most frequently, averaging four mentions per collaboration with nearly twice that of woman–man pairs (2.28) and four times that of man–man pairs (0.8). This suggests that women were more attentive to the collaborative protocol. Still, across all pair types, differences in intellectual merit varied widely, indicating that attentiveness to roles did not necessarily translate to more balanced intellectual contributions. This finding motivated our closer examination of intellectual merit differences by the gender of the pair.

At CSUMB, the differences in intellectual merit among woman-man pairs were less variable and generally smaller in comparison to Cal Poly. In general, when mixed-gendered pairs were both highly experienced programmers, or the woman was more experienced than the man, the differences in intellectual merit tended to be small, though men still scored higher, with one observed exception of a highly experienced woman and a less experienced man with chronic attendance issues. Mixed gender pairs with similar high levels of experience averaged a difference of 4.33 in intellectual merit and mixed gender where the women had higher levels of experience averaged a difference of 5. For mixed gender groups where both individuals had low levels of experience, we saw larger gaps in intellectual merit across interactions. Therefore, while prior programming experience seemed to mitigate gender differences in intellectual merit when experience levels for the women were high, gaps manifested when both individuals had less experience in programming.

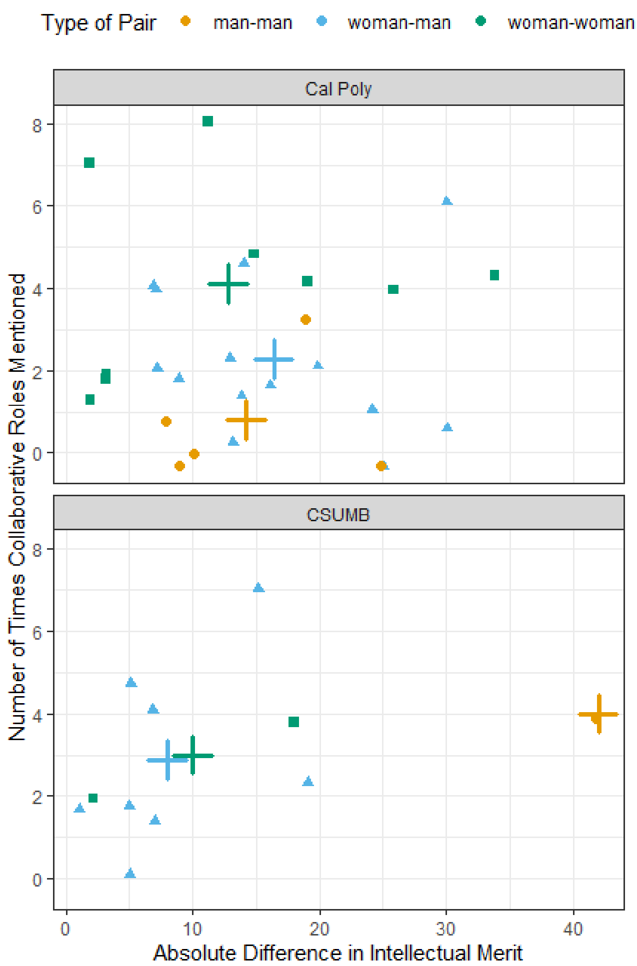

Figure 1: Plot of the absolute value of difference in intellectual merit versus the number of times collaborative roles were mentioned, color coded by group gender dynamic and faceted by school. Crosses represent the mean for each of the quantitative variables within each group.

DISCUSSION

The point discrepancies we observed reflect the interaction of two status systems: disciplinary / technological expertise and gender. These dynamics influenced behaviors such as taking over the keyboard, speaking outside one's assigned role, or dismissing a partner's ideas. We observed these patterns at both sites, though the specific mechanisms through which status operated differed.

At Cal Poly, disciplinary major played a central role in shaping intellectual merit, particularly within women–women pairs. Statistics majors were consistently positioned with higher intellectual merit than their non-major partners, producing the largest merit gaps when only one woman in the pair was a Statistics major. However, this expertise-based advantage disappeared in heterogendered pairs. When paired with men, women who were Statistics majors were not afforded the same intellectual merit. Instead, men often positioned their own ideas as more legitimate and asserted control over decision-making regardless of the woman's disciplinary expertise.

At CSUMB, technological authority was the more salient status marker. Students with greater programming experience often narrowed gaps in intellectual merit within their pairs, suggesting that technical expertise could elevate a student's status. Yet mixed-gender pairs still showed smaller but persistent gendered discrepancies. Women with technological authority also reported more negative experiences with pair programming in end-of-semester surveys compared to their peers. Additionally, chronic absenteeism contributed to uneven preparation and inconsistent adherence to assigned roles.

The persistence of status-based patterns at both institutions indicates that our current intervention did not fully counteract these dynamics, underscoring the need to revise the protocol to more effectively support equitable collaboration.

Modification to the Pair Programming Protocol

Findings from the pilot study pointed to confusion about the pair programming roles, which contributed to inconsistent role adherence and potentially enabled status dynamics to shape collaboration. In response, we revised the collaborative protocol in several key ways. First, the previous Developer role is now named the ***Computer*** to more accurately describe the student's function. Second, we clarified and rebalanced the responsibilities of each role: the Coder is now explicitly tasked with articulating their thinking, including points of confusion, while the Computer is responsible for encouraging this verbalization and ensuring that the Coder maintains control of the solution strategy.

To reinforce these distinctions, we introduced more tangible role structures. Each partner now receives a role-specific protocol sheet which they physically exchange during role switches, and only the partner currently occupying the Computer role may operate the keyboard. During the fall quarter at Cal Poly, we have observed stronger adherence to roles and more productive collaborative conversations with the revised protocols. CSUMB will adopt the updated protocol beginning in January 2026, supported by a redesigned attendance policy and a shift away from Friday-afternoon class sessions to promote more consistent student engagement. Collectively, these changes aim to bolster role clarity and reduce the influence of status characteristics—such as gender, major, and technological authority—on students' opportunities to contribute intellectually.

Future Directions and Further Investigations

As we complete this pilot study, several directions for future work have emerged. The analyses presented here aggregate data across the full quarter or semester; moving forward, we aim to examine how intellectual merit and other dimensions of power and authority shift over the duration of the course. We anticipate that role adherence will strengthen with practice and, if the revised roles function as intended, disparities should diminish over time. In addition, we plan to investigate whether the pair programming protocol mitigates status differences related to racial / ethnic identity, socioeconomic status, or other social markers. These extensions will help us assess the extent to which our intervention supports equitable collaboration across diverse student identities.

CONCLUSION

Across two institutions, this pilot study demonstrates how status characteristics shape who gains influence during collaborative programming. Even with a structured protocol informed by Complex Instruction, inequities persisted, suggesting that pair programming can reproduce broader social hierarchies without careful design. Our findings extend sociological theories of power and status into computational learning contexts and highlight the need for clearer roles, intentional scaffolding, and repeated practice to support equitable collaboration. The refinements introduced to our protocol represent an iterative step toward reducing status-based differences in intellectual merit, and future analyses will examine how these dynamics evolve over time. More broadly, this work underscores the importance of equity-focused instructional design and contributes to growing efforts to build collaborative learning environments that disrupt, rather than reinforce, inequities in STEM education.

REFERENCES

Chiu, M. M. & Khoo, L. (2003). Rudeness and status effects during group problem solving: Do they bias evaluations and reduce the likelihood of correct solutions? *Journal of Educational Psychology, 95*(3), 506–523. https://doi.org/10.1037/0022-0663.95.3.506

Conference Board of the Mathematical Sciences. (2016). *Active learning in post-secondary mathematics education*.

Cohen, E. G. & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. American Educational Research Journal, 32(1), 99-120. https://doi.org/10.3102/00028312032001099

Cohen, E. G., Lotan, R. A., Scarloss, B. A., & Arellano, A. R. (1997). Complex Instruction: Equity in Cooperative Learning Classrooms. *Theory into Practice*, 38(2), 80-86. https://doi.org/10.1080/00405849909543836

Ellington, R. M., & Frederick, R. (2010). Black high achieving undergraduate mathematics majors discuss success and persistence in mathematics. *The Negro Educational Review*, 61, 61–84.

Esmonde, I. (2009). Mathematics learning in groups: Analyzing equity in two cooperative activity structures. Journal of the Learning Sciences, 18(2), 247-284. https://doi.org/10.1080/10508400902797958

Fullilove, R. E., & Treisman, P. U. (1990). Mathematics achievement among African American undergraduates at the University of California, Berkeley: An evaluation of the Mathematics Workshop Program. *The Journal of Negro Education*, 59(3), 463–478. https://doi.org/10.2307/2295577

Gee, J.P. (2014) An Introduction to Discourse Analysis: Theory and Method. Routledge, Oxon.

Goyer, A., Lynch, A. & Wand, J. (2021) A Redesign of Precalculus at California State University, Monterey Bay, PRIMUS, 31:3-5, 492-503. https://doi.org/10.1080/10511970.2020.1746451

Johnson, B. (2002). *Introducing linguistics: Vol. 3. Discourse analysis*. Malden, MA: Blackwell.

Kuh, G. D. (2009), The national survey of student engagement: Conceptual and empirical foundations. *New Directions for Institutional Research*, 2009: 5-20. https://doi.org/10.1002/ir.283

Langer-Osuna, J. M. (2011). How Brianna became bossy and Kofi came out smart: Understanding the trajectories of identity and engagement for two group leaders in a project-based mathematics classroom. *Canadian Journal of Science, Mathematics and Technology Education*, *11*, 207–225. https://doi.org/10.1080/14926156.2011.595881

Langer-Osuna, J. M. (2016). The social construction of authority among peers and its implications for collaborative mathematics problem solving. *Mathematical Thinking and Learning*, *18*(2), 107–124. https://doi.org/10.1080/10986065.2016.1148529

Mulryan, C. M. (1994). Perceptions of intermediate students' cooperative small-group work in mathematics. *Journal of Educational Research, 87*(5), 280–291. https://doi.org/10.1080/00220671.1994.9941255

Oppland-Cordell, S., & Martin, D. B. (2015). Identity, power, and shifting participation in a mathematics workshop: Latin@ students' negotiation of self and success. *Mathematics Education Research Journal*, 27(1), 21–49. https://doi.org/10.1007/s13394-014-0127-6

Reinholz, D. L. (2024). Disrupting gender marginalisation in undergraduate mathematics: a case of assigning competence. *Intercultural Education*, *36*(1), 96–110. https://doi.org/10.1080/14675986.2024.2425561

Reinholz, D., Johnson, E., Andrews-Larson, C., Stone-Johnstone, A., Smith, J., Mullins, B., Fortune, N., Keene, K., & Shah, N. (2022). When Active Learning Is Inequitable: Women's Participation Predicts Gender Inequities in Mathematical Performance. *Journal for Research in Mathematics Education*, *53*(3), 204-226. https://doi.org/10.5951/jresematheduc-2020-0143

Reinholz, D. L., Moschetti M. G., Camacho J. T., Fuentes-Lopez E., Wilkes II C., et al. (2024) Equity in practice: Assigning competence to shape STEM student participation. *PLOS ONE* 19(4): e0299984. https://doi.org/10.1371/journal.pone.0299984

Salleh, N., Mendes, E. & Grundy, J. (2011). Empirical studies of pair programming for CS/SE teaching in higher education: A systematic literature review. Software Engineering, IEEE Transactions, 37(4), 509-525. https://doi.org/10.1109/TSE.2010.59

Wickham H., Averick M., Bryan J., Chang W., McGowan L. D., François R., Grolemund G., Hayes A., Henry L., Hester J., Kuhn M., Pedersen T. L., Miller E., Bache S. M., Müller K., Ooms J., Robinson D., Seidel D. P., Spinu V., Takahashi K., Vaughan D., Wilke C., Woo K., and Yutani H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). *R for Data Science (2nd ed.)*. O'Reilly Media. https://r4ds.hadley.nz/