ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022

Paper ID #37970

Developing Computational Thinking skills and STEM+C Career Interest through Adaptive Content Curation for Middle School Students

Emmanuel Johnson (Postdoctoral Research Associate)

Post Doctoral Research Associate USC Information Sciences Institute

Teresa M Ober (Assistant Research Professor) (University of Notre Dame)

Philip Gonsalves

Mayank Kakodkar (Purdue University at West Lafayette (COE))

Janice Zdankus (Vice President, Office of CTO)

Developing Computational Thinking skills and STEM+C Career Interest through Adaptive Content Curation for Middle School Students

Abstract

There is a growing interest in developing students' interest in computer science and computational thinking as computing has become ubiquitous in various fields. This has led to a plethora of online educational platforms offering computer science content for learning. There are several advantages and disadvantages of existing platforms. To address the current limitation of both the paid and free online learning activities, we present the Curated Pathways to Innovation (CPI), an innovative platform that utilizes machine learning techniques to curate and personalize readily available computational learning material online. Initial results suggest that students benefit from using CPI. Not only are students exposed to developmentally appropriate and personally salient computer science-related content aligned with educational standards, but they also experience a significant increase in their attitudes towards computer science activities, particularly those involving computer programming and computational thinking. In this way, CPI is highly scalable, having the potential to reach a broad audience of learners by curating content from an integrated set of educational resources and thus also orienting a future generation of students towards careers in computer science and related fields. The strengths of this approach, as well as opportunities for future platform, content, and curriculum development, are considered.

Introduction

There is a growing interest in developing students' interest in computer science, programming, and computational thinking as computing has become ubiquitous in various fields.[1] Within engineering fields, there is a growing recognition of the need to provide an undergraduate education that supports the development of computational and mathematical modeling skills.[2] Figure 1 shows examples of STEM+C programs available across the lifespan. Such interest has led to several online educational platforms offering computer science content for learning. While there are many widely available platforms promoting computer science content, and more general content in science, technology, engineering, mathematics, and computing (STEM+C), there are several disadvantages of the wide variety of current platforms. There are a number of ways to categorize currently available platforms. Though a considerable number of intelligent tutoring platforms exist to promote STEM+C achievement,[3 4 5 6] surprisingly no such integrated technology exists to promote curated learning opportunities for learning computing and computer programming skills.[7] The sections that follow provide an overview of some of the limitations of currently available STEM+C educational content.

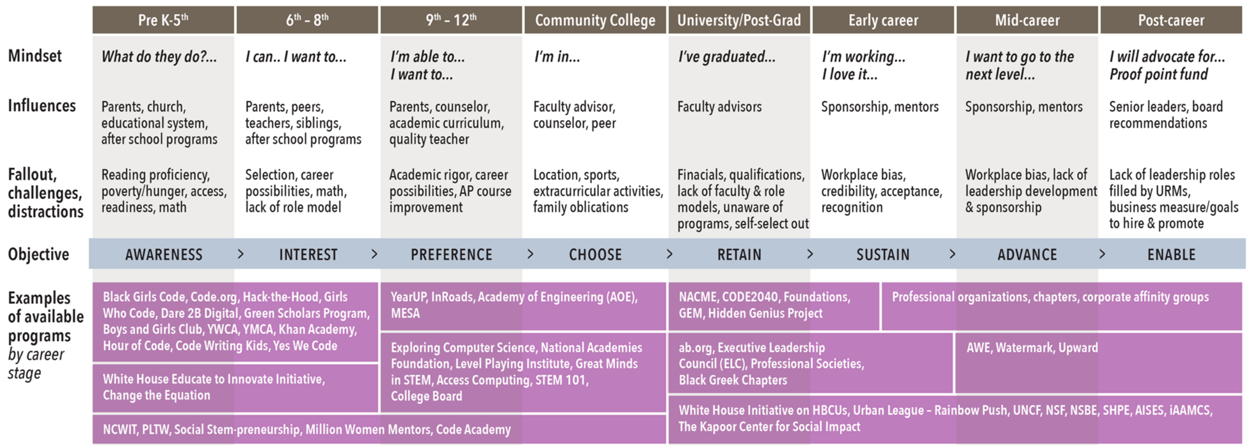

Figure 1: Key STEM+C development phases, programs, and tools

Limitations of Current Approaches

Cost of Paid Platforms

The advantage of many paid platforms is that they offer high-quality content. Such platforms can be expensive for students seeking an introduction to STEM+C and may be difficult for caregivers or school administrators to justify. Given such content is paywall protected, it is likely to require a log-in to access which may be difficult for younger students to manage.

Barriers to Accessing Free High-Quality Content

For those looking for a less costly alternative, there are many free STEM+C learning platforms available online. Much of this content is high quality; however, some may not be as rigorously developed or not aligned with educational standards as more costly options. Though not always, due to limited resources available for maintaining the content, some freely available platforms may suffer from "URL drift" whereby links no longer direct to the intended content. Furthermore, while some content available may be offered on a no-cost basis to users, as users progress through the content, they may soon meet a paywall thus also limiting access.

Inconvenience and Risks Associated with Multiple Log-ins

Most of these platforms are individual entities focused on specific topics that do not share content or data with each other. Thus students are required to create multiple logins and user accounts to gain access to a variety of learning content. This is problematic for several reasons. First, there are obvious concerns over the privacy and security of youth having multiple log-ins in an online environment. Each log-in account credentials created could expose students to different vendors who may have varying data use agreements and intents for using their data. While such data use agreements and intents may not be ill-guided, for minors especially the lack of uncertainty is concerning. Second, there are practical concerns in that students may struggle to manage their accounts, for example by not remembering their log-ins or account credentials. Such an experience can be deeply frustrating for learners and may deter them from using the online learning platform entirely.

Content Lacks Diverse Representation

Past research on visual representations of scientists via the "draw a scientist" task has found that even older children tend to carry certain stereotypes that may deter them from pursuing STEM+C.[8] Such representations are likely conveyed through media and appear to influence adolescents as they form career identities concerning STEM+C.[9] Providing underserved and underrepresented students in STEM+C with salient visual representations through the content available in a platform is therefore important.

Lack of Personalization

Instructional practices that are tailored specifically to students' individual interests, knowledge, and skills are more effective than those that do not take individual factors into account.[10] Yet, in nearly any learning environment - formal or informal, computer-mediated or not, providing instruction that is individualized and differentiated for each learner poses obvious challenges, particularly in deterring efforts to bring the quality learning experiences to scale. Even when personalized learning is recognized as a goal, many educational resources are designed in a more conventional manner that does not take into accounts learners' interests and motivations.[11] Within currently available STEM+C online learning platforms, particularly those which are free and openly available, there is typically not only limited content but also a lack of personalization within the design framework[12]

Innovations in technology, however, may offer some solutions for the paradox of making personalized learning experiences at scale more feasible. In particular, there is growing interest in the applications of artificial intelligence (AI) and machine learning (ML) in educational settings. Though there has been a recognition of the applications of AI/ML-enabled technologies in education for quite some time,[13] the topic has only recently garnered growing interest.[14][15][16] Despite the growing interest, a recent scan of available platforms suggested less than 15% appeared to leverage AI/ML for the sake of personalized learning. Such technology has the potential to provide students personalized learning experiences, support teachers in conducting targeted interventions through teaching, and assist educational administrators with insight in making decisions in the best interest of their students and school community.[17] One application of AI/ML technology for personalized learning is the use of recommender algorithms to "suggest" content to learners.[18][19][20]

Curated Pathways to Innovation

Given the current landscape of available platforms for teaching STEM+C, there is presently a need for a more integrated approach. Despite a relatively vast number of intelligent tutoring platforms for promoting STEM+C achievement, surprisingly, few integrated technologies exist to promote curated learning opportunities in STEM+C subjects. Curated Pathways to Innovation (CPI) is distinguished in this landscape as a freely available platform designed to encourage and prepare individuals from underserved and underrepresented groups in STEM+C. This is done by leveraging artificial intelligence to promote a personalized, educational standards-based, and structured experience using culturally responsive content. In particular, the CPI platform is an online platform that teaches computer programming and other STEM+C skills to middle and high school students by curating relevant content that is developmentally appropriate and interesting.[21] The platform delivers individualized learning possibilities for students who have traditionally had fewer options leading to careers in computer science by delivering curated content aligned with learners' interests. The benefits of CPI are described in further detail in the sections that follow. Table 1 provides a summary of the current limitations in the landscape of online STEM+C content and the strengths of the CPI platform in addressing those limitations.

Table 1: Summary of strengths of the CPI platform

Limitations of Current STEM+C Online Content	Strengths of the CPI Platform
- Often costly	- Free and openly available
- Barriers to accessing high-quality free content	- Aggregates content for multiple providers
- Multiple log-ins	- One log-in for multiple accounts
- Content often lacks diverse representation	- Vetted content reflects characters and settings showcasing diverse STEM+C talent
- Lack of personalization	- Recommender algorithm suggests personalized and developmentally appropriate content
- No pathway or clear progression	- Structures an individualized pathway for progress spanning
	Plus ...
	- Aligned content with CS educational standards
	- Scalable, requires minimal time invested in training educators

CPI is an innovative personalized learning application that leverages artificial intelligence and gamification to guide and incentivize students as they engage with computing and other STEM+C education content. It is unique because it connects previously disparate opportunities to learn computing skills available from other resources within a single platform to help students navigate a rich – but often inaccessible – set of STEM+C education programs. CPI is thus distinguished in this landscape as a freely available platform designed to address the needs of individuals from underserved groups in STEM+C by leveraging artificial intelligence to promote personalized, educational standards-based, and curated experience using culturally responsive content in computing. It is also mapped to Computer Science Teachers Association (CSTA) standards, Next Generation Science Standards (NGSS), and Common Core Mathematics Standards.

Developing CS Knowledge, Skills, and Career Interest

CPI pedagogy is based on the notion of "learn by doing."[21] As students engage with CPI, they are guided through three levels: Awareness, Cultivating Interest, and Preparation. Awareness activities are focused on engagement and exposing students to STEM+C careers and may include videos of diverse role models speaking about their career choices, and infographics about opportunities and salaries for individuals from underserved groups in STEM+C. Cultivating Interest activities help students see themselves as belonging to STEM+C as they gain confidence and skills, such as learning basic programming. Preparation activities are more challenging, requiring teaching concrete skills like Python and advanced mathematics while supporting the development of mindsets conducive to pursuing STEM+C. The activities recommended to students are sourced from local and national offerings that are vetted and mapped to national standards to form the basis of an ecosystem and continuum of program interventions. CPI is free to students and schools, easy to deploy, and does not require that teachers or facilitators have any STEM+C expertise. With limited content knowledge, teachers and facilitators can glean information from CPI about student progress. With such information, teachers can provide individualized guidance to students in using CPI.

CPI focuses on creating awareness, enhancing self-efficacy, building interest, and promoting career aspirations in STEM+C with an emphasis on computing fields (Figure 2). Within the broader mission of CPI, there are two long-term education and technology goals. The first long-term goal is to perform a methodologically rigorous assessment of underserved students' socioemotional development with respect to computing, particularly focusing on how mobile technology-driven personalized learning can broaden and sustain students' engagement and aspirations in computing and other STEM+C careers. The second goal is to increase the representation of historically underserved and underrepresented students in STEM+C disciplines. The third long-term goal is to leverage advanced mobile/web, machine learning, and data federation technologies to evolve the methodologies in CPI for deployment through under-resourced school districts.

Personalized Learning through AI-enabled Curation of Content

Through a machine learning algorithm, the CPI platform is designed to recommend activities based on learners' interests and ratings of previous activities. The recommender, in particular, learns matrix-factorization[22] based representations for learners and activities at the same time. A Bayesian model[23] is then used to describe how each student assessed each activity using these representations and other factors.

Learners have the option of selecting a badge after logging into the web-based portal. On a page that represents a pathway, learners are shown their progress in completing tasks toward badge completion. A badge signifies that a learner has successfully completed specified exercises aimed at teaching a common computer programming skill. They return to CPI after finishing the activity to rate it and take a short comprehension quiz to validate completion. Learners may be shown a gif or short animation relevant to an area they've indicated an interest in if they're successful. They are then returned to the pathway page to begin another action toward badge completion. The learner obtains credit for the badge after completing a certain number of activities in a badge sequence. The completion of badges often comes with bonuses, such as the ability to customize an avatar or access to additional activities. Some badges have prerequisites, thus earning one unlocks the possibility of earning more.

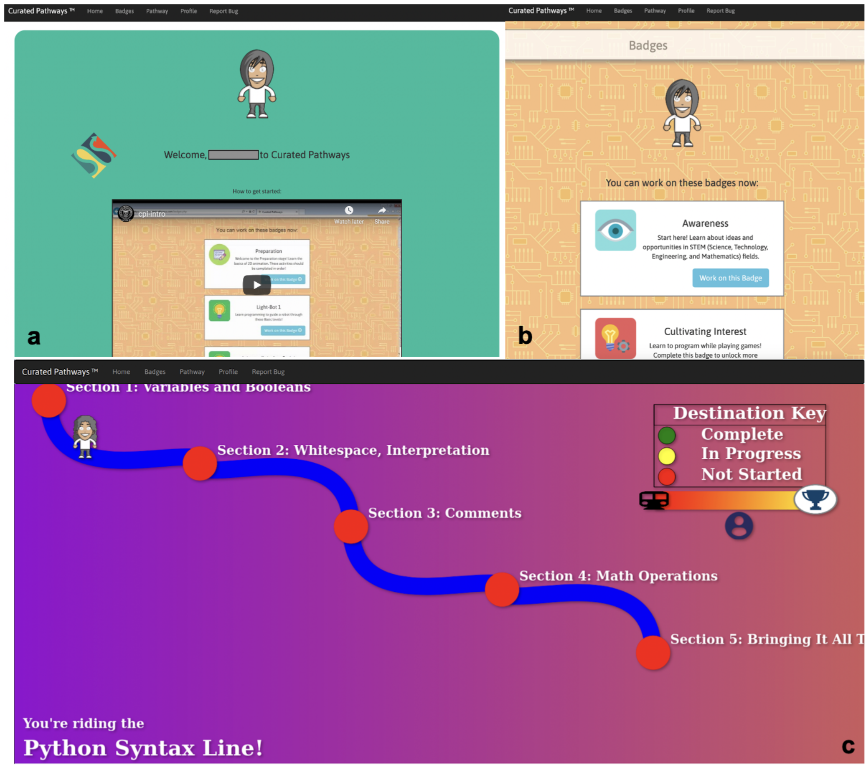

Figure 2: After logging in, students are greeted by their Avatar and land on the homepage where a tutorial video is available (a). Students can choose from a set of possible badges to complete (b). Students are shown their progress through the badge's activities (c). After completing the activity, they return to CPI to rate the activity and take a quiz to certify completion. If successful, they're shown a gif related to an area they have expressed interest in and sent back to the Pathway page (c) to continue the badge. After completing all activities, the student gets credit for the badge.

Current Impact

CPI has so far engaged over 5,000 middle and high school students (40.8% female, 33.6% Hispanic/Latinx, 16.2% White/European American, 16.1% Asian/Asian American and Pacific Islander, 6.9% Black/African American, 21.1% Other, 6.1% Did not respond). Teachers who have used CPI in the past have praised the activities for sparking students' interest in STEM+C careers, with one recently saying, "My ultimate goal is to expose students to STEM+C so they begin to consider a STEM+C career at an early age; CPI helps me achieve this goal through the interactive activities." "It provides pupils an idea of paths they may go and occupations they are interested in all while having fun and collecting badges," another teacher said.

Previous research involving 610 middle school students (Mean_{Age} = 12.07 years, SD_{Age} = .77 years, %female = 37.2%; 68.3% Hispanic/Latinx, 16.3% Asian/Asian American, 4.9% Hawaiian or Pacific Islander, 4.7% White/European American, 3.5% American Indian or Alaskan Native, 2.0% Black/African American, 20.1% Other category) in the United States who used CPI for one academic year found that their attitudes toward computer programming improved immediately, particularly in terms of awareness, self-efficacy, interest, and aspirations for a career in computer programming.[24] Furthermore, the study found that improvements in attitudes toward computer programming were similar for female and male students, as well as students from underrepresented racial/ethnic groups.

We wanted to test how effective the recommendation system based on user self-report data was in light of these promising preliminary findings. With a sample of 78 CPI student users ($Mean_{Age}$ = 13.2 years, SD_{Age} = .84 years, %female = 37.2%; 60.3% White/European American, 27.6% Black/African American, 8.6% Hispanic/Latinx, 3.4% Asian/Asian American), we looked at whether a recommender score obtained by a ML algorithm that was used to identify a preferred badge (i.e., group of activities) for a specific student was positively and significantly associated the student's self-reported ratings of the badge. Importantly, the ratings of the badge provided by students were independent of those used by the algorithm to generate the recommender score. There was a significant positive relationship between overall badge approval ratings provided by learners and the recommender score after controlling for other factors such as the position of the recommendation on a list and the number of previous selections the learner had made, as well as certain learner demographics. These findings demonstrate the utility of the ML-generated recommendations of badges given that the learner tended to approve more of a recommended badge, even when other important parameters such as the location of the recommendation on the list and the number of badges selected by the learner are taken into account.

Future Directions

There has been considerable work on CPI towards improving the availability of content. However, there are still several critical means by which its potential wide-reaching impact could be improved.[25] First, little is presently understood about how different students perceive and benefit from certain types of content within CPI. Though it reflects a limitation of the content available to teach STEM+C, there is currently an underrepresentation of female and racially/ethnically diverse role models in popular STEM+C education media relative to the U.S. population, which may place certain students at a disadvantage in learning content.[26] Second, further work is needed to understand if and how the sequencing of STEM+C educational content (e.g., beginning with more or less abstract concepts) may impact students learning and interest development. Content sequencing has also been found to have differential effects on students based on whether they prefer experiential learning.[27] Third, there is a need to further refine and assess certain aspects of methodologies in CPI that are likely to contribute to its efficacy (e.g., incentivization, personalization, strengthening data reporting for teacher-student feedback processes, presentation of recommended activities, integration of hands-on and hybrid activities, clear pathways to STEM+C careers, etc.) to enhance students' experience using the platform and the potential for its long-term impact. Finally, there is a need to consider the ethical and long-term consequences of this particular application of AI in education[13,14] to ensure that it is achieving its aim of supporting underserved and underrepresented students in STEM+C.[28] For example, we should strive to ensure that AI-enabled platforms such as CPI minimize the negative impact of algorithmic bias, whether it is apparent or not[29]. Inviting input from educators and students from diverse backgrounds in the development of such systems may help to prevent the negative impact of algorithmic bias.[30] Further work in each of these areas has the potential to greatly improve the impact CPI platform on student users' knowledge of STEM+C as well as their future academic and career interests development at scale. These aims are one step closer towards increased capacity and greater diversity in the STEM+C workforce.

Conclusions

There is a growing interest in providing STEM+C online educational experiences to K-12 learners which promote not only knowledge and skills in the area, but also facilitate interest development. However, there are several limitations of currently available online educational content. These limitations pertain to the high cost of paid platforms, the barriers to accessing high-quality content openly available, the need for multiple log-ins to access a variety of educational content, as well as the lack of diverse representation and personalization. The CPI platform addresses these limitations by offering an integrated approach that leverages AI to provide learners not only access to quality educational STEM+C content but also curates content and provides personalized pathways to navigate through it. Ultimately, the CPI platform supports students' knowledge, skills, and long-term interest in STEM+C, thus motivating young learners to pursue an academic and professional future in those areas. Further work to develop the CPI platform will continue to consider the users' experience and the ethical implications of AI in an applied educational context.

Acknowledgements

This research was funded in part by the Office of Naval Research under award N00014-21-1-2437.The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation hereon.

References

[1] Engineering National Academies of Sciences, Medicine, et al. *Assessing and responding to the growth of computer science undergraduate enrollments*. National Academies Press, 2018.

[2] Joseph A Lyon and Alejandra J Magana. A review of mathematical modeling in engineering education. *Int. J. Eng. Educ*, 36:101–116, 2020.

[3] Tsai Chen, Aida MdYunus, Wan Wan Ali, and AbRahim Bakar. Utilization of intelligent tutoring system (its) in mathematics learning. *International Journal of Education and Development using ICT*, 4(4):50–63, 2008.

[4] Xudong Huang, Scotty D Craig, Jun Xie, Arthur Graesser, and Xiangen Hu. Intelligent tutoring systems work as a math gap reducer in 6th grade after-school program. *Learning and Individual Differences*, 47:258–265, 2016.

[5] Walter L Leite, Dee D Cetin-Berber, Anne C Huggins-Manley, Zachary K Collier, and Carole R Beal. The relationship between algebra nation usage and high-stakes test performance for struggling students. *Journal of Computer Assisted Learning*, 35(5):569–581, 2019.

[6] Maria Ofelia Z San Pedro, Ryan SJ d Baker, Ma Rodrigo, and T Mercedes. Carelessness and affect in an intelligent tutoring system for mathematics. *International Journal of Artificial Intelligence in Education*, 24(2): 189–210, 2014.

[7] Brian R Belland, Andrew E Walker, Nam Ju Kim, and Mason Lefler. Synthesizing results from empirical research on computer-based scaffolding in stem education: A meta-analysis. *Review of Educational Research*, 87(2):309–344, 2017.

[8] David I Miller, Kyle M Nolla, Alice H Eagly, and David H Uttal. The development of children's gender-science stereotypes: a meta-analysis of 5 decades of us draw-a-scientist studies. *Child development*, 89(6):1943–1955, 2018.

[9] Jocelyn Steinke. Adolescent girls' stem identity formation and media images of stem professionals: Considering the influence of contextual cues. *Frontiers in psychology*, 8:716, 2017.

[10] Jay McTighe and John L Brown. Differentiated instruction and educational standards: Is détente possible? *Theory into practice*, 44(3):234–244, 2005.

[11] Matthew L Bernacki, Meghan J Greene, and Nikki G Lobczowski. A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose (s)? *Educational Psychology Review*, 33(4):1675–1715, 2021.

[12] Candace Walkington and Matthew L Bernacki. Appraising research on personalized learning: Definitions, theoretical alignment, advancements, and future directions, 2020.

[13] Robert M Aiken and Richard G Epstein. Ethical guidelines for ai in education: Starting a conversation. *International Journal of Artificial Intelligence in Education*, 11:163–176, 2000.

[14] Wayne Holmes, Kaska Porayska-Pomsta, Ken Holstein, Emma Sutherland, Toby Baker, Simon Buckingham Shum, Olga C Santos, Mercedes T Rodrigo, Mutlu Cukurova, Ig Ibert Bittencourt, et al. Ethics of ai in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, pages 1–23, 2021.

[15] Yu-Shan Lin and Ying-Hsun Lai. Analysis of ai precision education strategy for small private online courses. *Frontiers in Psychology*, 12:749629–749629, 2021.

[16] Stephen JH Yang, Hiroaki Ogata, Tatsunori Matsui, and Nian-Shing Chen. Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers and Education: Artificial Intelligence*, 2:100008, 2021.

[17] Huiling Qin and Guan Wang. Benefits, challenges and solutions of artificial intelligence applied in education. In *2022 11th International Conference on Educational and Information Technology (ICEIT)*, pages 62–66. IEEE, 2022.

[18] Maria-Iuliana Dascalu, Constanta-Nicoleta Bodea, Monica Nastasia Mihailescu, Elena Alice Tanase, and Patricia Ordoñez de Pablos. Educational recommender systems and their application in lifelong learning. *Behaviour & information technology*, 35(4):290–297, 2016.

[19] Mojisola Erdt, Alejandro Fernández, and Christoph Rensing. Evaluating recommender systems for technology enhanced learning: a quantitative survey. *IEEE Transactions on Learning Technologies*, 8(4):326–344, 2015.

[20] John K Tarus, Zhendong Niu, and Abdallah Yousif. A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. *Future Generation Computer Systems*, 72:37–48, 2017.

[21] Natalie Linnel, Alankrita Dayal, Phil Gonsalves, Mayank Kakodkar, Bruno Ribiero, Ariel Starr, Tim Urdan, and Janice Zdankus. Curated pathways to innovation: personalized cs education to promote diversity. *Journal of Computing Sciences in Colleges*, 35(10):39–45, 2020.

[22] Simon Jackman. *Bayesian analysis for the social sciences*. John Wiley & Sons, 2009.

[23] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[24] Teresa M Ober, Ying Cheng, Meghan Coggins, Tim Urdan, Paul Brenner, Gonsalves Philip Zdankus, Janice, and Emmanuel Johnson. Evaluating longitudinal growth in middle school students' attitudes towards computer programming. *under review*.

[25] René F Kizilcec, Justin Reich, Michael Yeomans, Christoph Dann, Emma Brunskill, Glenn Lopez, Selen Turkay, Joseph Jay Williams, and Dustin Tingley. Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences*, 117(26):14900–14905, 2020.

[26] Fashina Aladé, Alexis Lauricella, Yannik Kumar, and Ellen Wartella. Who's modeling stem for kids? a character analysis of children's stem-focused television in the us. *Journal of Children and Media*, 15(3): 338–357, 2021.

[27] Kasee L Smith and John Rayfield. Stem knowledge, learning disabilities and experiential learning: Influences of sequencing instruction. *Journal of Agricultural Education*, 60(2):222–236, 2019.

[28] Guan Saw, Chi-Ning Chang, and Hsun-Yu Chan. Cross-sectional and longitudinal disparities in stem career aspirations at the intersection of gender, race/ethnicity, and socioeconomic status. *Educational Researcher*, 47 (8):525–531, 2018.

[29] René F Kizilcec and Hansol Lee. Algorithmic fairness in education. *arXiv preprint arXiv:2007.05443*, 2020.

[30] Olaf Zawacki-Richter, Victoria I Marín, Melissa Bond, and Franziska Gouverneur. Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1):1–27, 2019.