Feasibility of Using the CAPE Framework to Identify Gaps in Equity-focused CS Education Research

Joey Reyes*
Knox College & CSEdResearch.org
jmreyes@knox.edu

Monica M. McGill*
CSEdResearch.org
monica@csedresearch.org

Abstract

As computer science (CS) quickly gains ground in K-12 classrooms, CS education research (CSER) is also rapidly growing. It remains unclear, however, the extent to which this growth in research captures key equity-focused areas. In this article, we describe a pilot test to determine the feasibility of using the CAPE theoretical framework to identify coverage of equity-focused CSER. The Capacity, Access, Participation and Experience (CAPE) framework examines the capacity to offer CS education, learner access to CS education, learner participation in CS education (enrollment) and experiences learners have when learning CS. We started with one primary research question: *How feasible is it to use the CAPE framework for identify coverage gaps in K-12 CS education research?*

We then created a secondary research question for narrowing down the set of articles examined and testing its feasiblity: *What are the gaps in research focused on K-12 CS education in which girls are participants in the studies?*. We chose to use an existing, publicly-available dataset of 800+ articles and examined studies in which only girls were participants (n=51), then examined each of the 51 articles to determine which key CAPE component(s) each covers. Our pilot results show that CSER among girls covers areas related to Experience (92%) and Capacity (59%), but little to no coverage in the areas of Access (0%) and Participation (2%) of girls.

To answer the primary research question and determine the feasibility of using CAPE for analyzing the entire corpus of 800+ articles (which is the next step in our research plan), we evaluated feasibility across two key areas, implementation and practicality, and found both to be satisfactory. This study is important in understanding how to identify areas of equity covered in K-12 CS education research and areas that need more attention in order to build a broader set of research knowledge for identifying promising practices for all learners.

1 Introduction

K-12 computer science (CS) education is rapidly expanding, as well as the need for more education research to investigate promising practices–particularly those that study across the educational ecosystem and the broad range of student populations and groups. With the different subgroups of students that are only becoming more diverse, particularly in the United States, there is a need to understand how well our existing body of research expands across student subgroups and whether or not gaps exist. Further, there have been calls to be more inclusive in research and previous studies have also shown gaps in equity-focused research [2, 3, 8, 13]. By identifying such equity gaps, the broader research community can start to be more deliberate about conducting research in areas where there may be gaps [14].

In 2020, Fletcher and Warner developed the equity-focused CAPE theoretical framework to better understand the related components that impact all student learning of and growth in computer science. This includes the capacity to offer CS education, access to and participation in CS education, and experiences of those students receiving CS education. After working with CAPE for several months across various projects, we hypothesized that we could use it to help identify gaps in equity-focused CSER given its intent and focus as a framework specifically created for K-12 CS education. Our primary research question became: *How feasible is it to use the CAPE framework for identify coverage gaps in K-12 CS education research?*

In this paper, we present two methodologies–one for investigating the feasibility of using CAPE to identify gaps across a dataset focused on girls (specifically focusing on implementation and practicality of using this framework) and the other running the CAPE framework through its paces against a limited dataset as a pilot test. That is, we conducted a limited study to investigate whether or not we could use CAPE to identify the gaps in research in which girls were identified as participants in the study. Therefore, our secondary research question became: *What are the gaps in research focused on K-12 CS education in which girls are participants in the studies?*

This study is a precursor to understanding the broader landscape of equity-focused research in the computing education research literature. By understanding the areas in which the research evidence is excelling and where additional research is needed, the broader research community and stakeholders who understand the importance of filling such gaps can start to address them.

2 Background

2.1 Previous CSER Research Gaps

Multiple studies have investigated topics of research to identify knowledge gaps. Saha explores gaps in local governments' sustainability efforts to drive future research [21]. Macintosh et al. investigate research gaps as they pertain to eParticipation, finding that one of the six gaps is in equity [15]. With respect to gaps specifically in equity-focused research in general, Lubienski recognize the importance of exploring these gaps, stating that "Research on gaps between underserved groups and their more advantaged peers are important for shaping public opinion and informing education policy." [14, abstract]. Recently, Ayalew et al. used a sociological framework with an equity lens to identify the types of equity gaps in research [4]. Ab Rahim et al. conducted an interesting equity-related gap analysis that considered whether there was a research gap in the introductions of CS research articles by authors who were non-native English writers [1].

Closer to CSER, there has been a history of considering where gaps are in the research to improve our body of knowledge. Several authors (Randolph et al. [20], Heckman et al. [11], McGill and Decker [17] and Sanders et al. [22]) have considered gaps in data and methodological analysis. McGill et al. conducted an analysis to uncover gaps in instrumentation that is used to study the effectiveness of interventions [19]. Further, McGill et al. considered reporting of data and how gaps in reporting could be addressed [18].

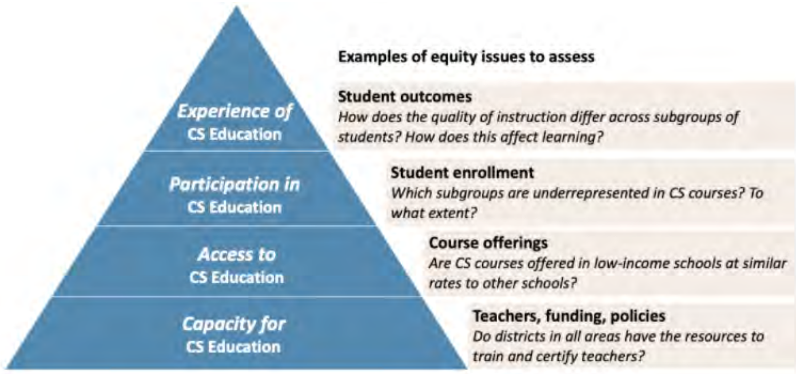

Figure 1: CAPE Framework as defined by Fletcher and Warner [10].

With respect to equity-focused CSER, Smith considers the gaps in participant diversity in CS education research by analyzing "...the gender composition of the subject pool, the extent to which cognitive learning results are disaggregated by gender, and whether there are gender differences in outcomes" [23, abstract]. The author found that only about one-third of studies indicated the gender composition of the participants, with only one-third of the studies that identified gender composition being female. This process aligns with participant demographics found in an earlier study [25][ANON].

Bianchini et al. considers three key areas with respect to science education research:

- Researchers organizing and sharing their work in ways that align with the same theories of teaching and learning that we promote and study,
- Researchers' framing and research methods that more directly address issues of power, voice, and even impact and
- Researchers' ability to impact funding, evaluation and policy that is equity-centered [6].

Their position is that in order to move the "equity agenda" forward, the landscape of research, practice and policy need to shift to all be equity-focused. Although these questions are differently focused, the center the researcher in the process of driving equity-focused research forward. This further supports the need to begin to understand the landscape of equity-focused research in our field, to identify the gaps, and then to address them.

2.2 The CAPE Framework

The CAPE framework is a relative newcomer to understanding and disaggregating the complexities of the CS educational ecosystem, with a particular focus on K-12 [10]. The Capacity, Access, Participation and Experience framework examines the capacity to offer CS education, learner access to CS education, learner participation in CS education (enrollment) and experiences learners have when learning CS. The framework, as shown in Figure 1, shows some leading questions that can be asked to analyze CS education at a systems-level approach. A systems-level approach considers student or teacher level outcomes and how those outcomes are situated within a larger initiative and policy level environment [5].

The CAPE framework provides a lens from which to view educational systems and whether or not they are equitable. As such, it may be able to provide researchers a way of understanding coverage of research that falls within each of these categories and highlighting where important gaps may exist.

Given its newness, there are not currently any existing studies that look at gaps in K-12 CS education research using the CAPE framework. Though there are other frameworks that might be used to investigate those gaps, we chose to use the CAPE framework because of its particular focus on K-12 computing education and the unique insight it provides for the emergent K-12 subject area of computing education.

3 Methodology

In this section, we present the methodology for both research questions.

3.1 Primary Research Question: Feasibility

To answer our primary research question, *How feasible is it to use the CAPE framework for identify coverage gaps in K-12 CS education research?*, we adapted an existing design methodology to evaluate the feasibility of our classification process [7]. Bowen et al. provide eight successive areas of focus for such studies that include: Acceptability, Demand, Implementation, Practicality, Adaptation, Integration, Expansion and Limited Efficacy Testing. Acceptability and Demand are two areas in which we made judgment calls–our future research plans call for a way to identify gaps in equity-focused research in CS education. Therefore, we believe using the CAPE framework may be an acceptable way to identify those gaps and the demand was our own based on our future research needs. The latter areas of Adaptation, Integration, Expansion, and Limited Efficacy Testing are used to test feasibility of those who use and adapt the process and integrate it in their research, integrating into their practices, expanding the usage of it beyond its original intent and test among a broader population for efficacy.These latter four were beyond the scope of what we wanted to do in this pilot study.

Implementation, however, considers how feasible it is to implement the process. It considers four primary areas: the degree of execution of the process, its success or failure, the resources needed to implement, and the factors affecting implementation ease or difficulty [7]. Practicality is a step that follows implementation and it considers how practical the process is for both the researchers studying the process and others. It considers the efficiency, speed, or quality of implementation, positive/negative effects of using the process, ability to administer the process and a cost analysis of using the process. Based on these two definitions, these most aligned with what we wanted to explore for feasibility (see Table 1).

Though our original overarching question was *How feasible is it to use the CAPE framework for identify coverage gaps in K-12 CS education research?*, we further clarified this to two primary questions about feasibility:

- To what extent can the process of classifying articles against the CAPE framework be performed successfully?
- To what extent can classifying articles against the CAPE framework be carried out using existing means, resources, and circumstances and without outside assistance?

These two questions serve as the basis for our discussion in Section 6.

Table 1: Two areas of focus (Implementation and Practicality) adapted from Bowen et al. [7].

Area of focus	**Our feasibility study asks . . .**	**Sample outcomes of interest**
Implementation	To what extent can the process of classifying articles against the CAPE framework be performed successfully?	Degree of execution; Success or failure of execution; Amount, type of resources needed to implement; Factors affecting implementation ease or difficulty
Practicality	To what extent can classifying articles against the CAPE framework be carried out using existing means, resources, and circumstances and without outside assistance?	Efficiency, speed, or quality of implementation; Positive/negative effects of using the process; Ability to administer the process; Cost analysis

3.2 Secondary Research Question: Girls

To conduct this pilot study and answer the secondary research question, *What are the gaps in research focused on K-12 CS education in which girls are participants in the studies?*, we needed to gather a proper dataset, develop inclusion and exclusion criteria, and then classify each article according to which category of CAPE it fell under: Capacity, Access, Participation, or Experience. We also needed to identify a methodology for determining if our process was feasible for broader studies that we or others may conduct.

3.2.1 Dataset

Our dataset is drawn from an existing, publicly-available pool of over 800 articles from the K-12 CS Education Research Resource Center [16]. This dataset was created by the authors' curating data from articles related to K-12 CS education from each of the following journals and conference publication venues (2012-2020): ACM International Computing Education Research, ACM Innovation and Technology in Computer Science Education, ACM SIGCSE Technical Symposium on Computer Science Education, ACM Transactions on Computing Education, Frontiers in Education, IEEE Global Engineering Education Conference, IEEE Transactions on Education, Journal of Educational Computing Research, Koli Calling, Taylor & Francis' Computer Science Education and Workshop in Primary and Secondary Computing Education (WIPSCE).

Using an SQL query on the dataset provided to us, we limited our dataset to articles that only used girls as participants. This allowed us to use an appropriate sample size (54 articles) for a pilot study while also focusing on an underrepresented group in computing fields. The SQL query captured the relevant articles, their abstracts, and what was measured in the study and then we placed the data in a spreadsheet and added four columns for each component of CAPE. We then downloaded all of the articles and placed them in a shared folder for further analysis.

3.2.2 Classification and Criteria

In order to classify each article under a component of CAPE, we developed inclusion and exclusion criteria for each. The inclusion/exclusion criteria was based on the CAPE definitions as follows:

- Capacity: Investigates resources (e.g. faculty, funding, curriculum and policies) in regards to CS education

- Access: Investigates CS education offerings and/or barriers of entry (e.g. course requirements)
- Participation: Investigates enrollment in CS courses
- Experience: Investigates the student outcomes of participation in CS (e.g. content learned, attitudes)

In order to draw a clear line between what the research says and our own projection of how the research might impact a component of CAPE, we decided to only focus on what the instruments measured. For instance, a study might use positive attitudes towards CS from student surveys to show that a workshop has the potential to improve CS enrollment among girls. However, since the research is measuring attitudes rather than actual enrollment numbers, it would not be categorized under Participation.

Using this criteria, we read through the articles and specifically looked at whether the instruments measured Capacity, Access, Participation, or Experience and tagged them accordingly. We also made note of which specific constructs were measured and identified each as being *directly* investigated (direct) where the research was intended to measure this component and *indirectly* investigated where the research was not intended to measure this component, but evidence for this component was present in the research findings. For example, one of the articles classified under Participation did not measure enrollment as that was not the aim of the study, but reported on future camp enrollment numbers in the discussion section. Thus, the paper indirectly investigated participation. This is further explained in Section 4.2.

3.2.3 Process

After we established the inclusion criteria, both authors reviewed the first two papers and classified them against the CAPE components. After this was completed, the authors met via Zoom to discuss how and why they classified each study under the specific components. From this, we both classified these as falling into the Experience category. However, this also prompted a deeper discussion about how the experience was measured as part of an intervention, and the intervention was related to the Capacity component. We decided that in cases like this, we would categorize the intervention where it belonged (Capacity, Access, or Participation). Therefore, we learned very quickly that the classification process required much more thought when interventions were being classified.

Over the course of the next several weeks, the first author read through approximately 10 articles per week and categorized them under the CAPE framework. At the end of each week, both researchers met via Zoom to discuss questions, comments, and concerns about the process that the first author had. During these meetings, the second author reviewed the classifications of each article and changes were made, if necessary. This process continued until every article had been properly categorized according to our criteria.

Table 2: Articles that were classified as Capacity, with the focus area

Focus Area	Count	% of all articles
Pedagogy	16	31%
Resource/Tool	12	23%
Assessment	3	6%
Ecosystems	3	6%
Curriculum	2	4%

4 Results: Secondary Research Question

4.1 Capacity

As we reviewed the articles and classified those that met the definition as Capacity, we also captured the major area of focus that the intervention was targeting. As we did this, we came up with five focus areas: Pedagogy, Assessment, Resource/Tool, Curriculum, and Ecosystems (broader interventions across Capacity) (see Table 2).

Of the 51 articles, 30 (59%) of them directly measured Capacity. The majority of the articles about Capacity dealt with Pedagogy and Resources/Tools (e.g., Scratch, Google Blockly).

4.2 Access and Participation

None of the articles measured Access and only one article (indirectly) measured Participation. The one article relating to participation was indirect since the research was not intended to measure enrollment but reported on it in the article. The author aimed to find the impact of an outreach camp on computing confidence, intent to persist, social support, and computing outcome expectations. The survey instruments measured these constructs rather than enrollment. However, in the Discussion section the author makes note of camp enrollment numbers. Thus, participation is indirectly investigated.

4.3 Experience

Similar to Capacity, as we started classifying Experience reports, we also identified the construct that was investigated (e.g., self-efficacy, perceptions about CS). Once we tagged each with the area of investigation for the research, we then cross-referenced this with categorizations created in [ANON] and [12] and used previously defined construct terminology. We learned that 23 constructs were measured across these articles (see Table 3).

Of the articles, 47 (88%) directly measured the Experience of girls in CS. The most frequent construct measured Enjoyment followed by Confidence and Content Knowledge, respectively.

5 Discussion: Secondary Research Question

When considering the results in total (Table 4), we see that the majority (92%) of the articles that specifically have girls as participants are focused on the student experience. This is a significant number, yet also not surprising–centering education research on the learner experience is vital to understanding interventions that specifically target students.

However, as Fletcher and Warner indicate via the CAPE framework, Capacity is the foundation of creating an ecosystem that supports student learning. By only focusing on the learner experience,

Table 3: Articles that were classified as Experience, with the construct measured.

Construct	**Count**	**%**
Enjoyment	20	40%
Confidence	16	31%
Content Knowledge	14	27%
Intent to Pursue (CS Related Career)	11	22%
Engagement	11	22%
Intent to Pursue (CS Related Major)	11	22%
Self-efficacy	10	20%
Usefulness	10	20%
Motivation	9	18%
Intent to Pursue (CS Related Courses)	6	12%
Satisfaction	6	12%
Beliefs/Perceptions	5	10%
Belonging	3	6%
Collaboration	3	6%
Support	3	6%
Identity	3	6%
Problem Solving Strategies	3	6%
Comfort with Peers	1	2%
Perceptions about CS (general)	1	2%
Perceptions about Topic being Taught (in intervention)	1	2%
Self Concept	1	2%
Self Regulation	1	2%
Persistence	1	2%

it directs our attention away from the structural needs of a healthy education ecosystem in which all students can thrive. Organizations that support the professional development of teachers and the growth of a school's, district's, state's or country's ability to offer CS courses and extracurricular activities would all fall into the Capacity component of CAPE–yet this area is not measured nearly as frequently as Experience. Capacity's relationship to Access and Participation also cannot be adequately studied without studies that explicitly study these two components.

Further, by understanding the importance of Access and Participation, one can start to also understand the importance of research focused on both [24]. For example, if a school only offers CS courses that conflict with a physical education class that is popular with volleyball or softball players (girls), this limits the ability for these students to *participate* in formal CS education because their *access* to it is restricted. Or, if the requirement for CS is Algebra II, yet only college-oriented students typically take Algebra II, then students who may not take Algebra II would not be able to participate due to limited access. Research that studies barriers to access and participation (which must be addressed at the capacity level) are critical if we are to design and promote CS for all students.

When taking a closer look at the focus areas of Capacity research as related to offering girls CS

Table 4: Overall results of the process. Papers that studied Capacity and Experience were all directly measuring these two. Participation was only indirectly measured as a byproduct of a study that measured both Capacity and Experience. To highlight the overlaps with Capacity, we also provide Capacity with each other component.

	Count	%
Capacity	30	59%
Access	0	0%
Participation	1	2%
Experience	47	92%
Capacity + Access only	0	0%
Capacity + Participation only	0	0%
Capacity + Experience only	29	57%
Capacity + Participation + Experience	29	57%

education, we see that Pedagogy (31%) is the most frequently examined focus area of these papers, whiles Resources/Tools are not too far behind (23%). However, Capacity covers so much more–administrator support, teacher factors (e.g., experience, equity training, growth mindset), guidance counselor training and experience, classroom factors, etc. These are all vital areas to study as they affect girls' access to CS education, their participation in the courses, and, of course, their experiences in the classroom. The focus areas seem to overlook an important aspect of Capacity that the CAPE framework asks: Who has the human and financial resources to offer CS? While many of the articles in our dataset can be classified as Capacity, none of the articles investigated the financial resources that schools have to implement the pedagogy and incorporate the resources/tools being studied.

When taking a closer look at the constructs investigated under student experience, we see that the most researched constructs are Enjoyment (40%), Confidence (31%), and Content knowledge (27%). While Enjoyment is an important construct, it is not always the most accurate measure. For instance, external factors can influence the extent to which participants enjoy an experience (e.g. time of day of the camp, prior relationships with other participants, food/refreshments being offered, etc) [9, 12]. Also, some of the other constructs can be an indicator of enjoyment. For example, if a student is intending to pursue CS, whether as a major or as a career, it stands to reason that the student likely had some positive experience with the subject already. Additionally, if the course was high quality and the students could relate to the material, their enjoyment levels would rise [9]. Without a way to discern why the ”enjoyment” is occurring, we cannot easily relate this to its primary cause. Content knowledge, on the other hand, can be measured as a more direct result of the topic being taught in the course.

While it is important to have diversity amongst components of CAPE, it is also important to have variety amongst the constructs measured within each component. For example, Farrington et al. state that Study Skills is a core area that researchers should study given its impact on student academic achievement and growth [9]. However, Study Skills was not an area that researchers studied among girls-only participant studies, thus indicating a gap in the research. The authors also mention that other academic behaviors can influence grades and achievement test scores. These include class attendance and homework completion are important constructs to measure (as well as their barriers). As important as they have been noted in prior research findings, neither of these appear as a measurement in the articles we reviewed.

6 Results & Discussion: Primary Research Question

To measure the primary research question *How feasible is it to use the CAPE framework for identify coverage gaps in K-12 CS education research?*, we examined each of our subquestions related to feasbility, one for implementation and one for practicality.

6.1 Implementation

In regards to Implementation, we set out to reflect on and answer the feasibility question: *To what extent can the process of classifying articles against the CAPE framework be performed successfully?*

6.1.1 Degree of execution

For the most part we were successful in the execution of this process. Since the dataset was relatively small, we looked at both research articles and experience reports. We were able to successfully categorize all of the articles in our dataset according to the CAPE framework and the results showed us gaps in the research for this particular set of data. The high degree of execution in this pilot study showed us that it is possible to implement this process on a larger scale. Implementation on a larger dataset is likely to yield greater results by highlighting more significant gaps in the literature.

6.1.2 Success or failure of execution

The high degree of execution was possible. After we resolved the initial challenges, we were able to classify each paper as measuring one or more of the CAPE components. Before narrowing the definitions, we initially came to different categorizations of the studies. Several articles mentioned various aspects of Capacity, Access, Participation, and Experience and there was too much room for interpretation without clarifying definitions. We needed to limit subjectivity involved in the process by creating stricter definitions of what should be included in each category. Once we decided to only look at what the researchers measured rather than discussed in the paper (e.g., future goals of the project or future predictions of the impact on participants), we were successful in our execution.

6.1.3 Amount, type of resources needed to implement

The entire process took approximately 72 hours to complete. Around half of our time was spent building and clarifying the methodology, while the other half was spent reviewing the articles. It took less time to categorize a paper as we became more familiar with the process and the specifics of what we were looking for.

This also took less time due to the fact that our dataset was essentially at our fingertips. Using an existing dataset rather than having to construct one from scratch greatly reduced the resources needed for this research study.

6.1.4 Factors affecting implementation ease or difficulty

There were a number of factors that made it easier and more challenging for us. Since researchers themselves do not classify their papers as belonging to one or more of the CAPE components and we were specifically looking at what was measured, we could not use a keyword search or automate the process. To classifying the papers, we were required to read through each individual article. The beginning of the categorization process was harder as we were still working with loose definitions of CAPE, but it became easier as our definitions became clearer.

As mentioned in the previous section, one factor that made the process easier is that we already had the dataset at hand and this simplified the process of gathering relevant articles. Rather than searching for articles in the existing literature, we were able to obtain relevant articles within a few minutes through a query.

6.2 Practicality

For the practicality of this process, we set out to reflect on and answer the question: *To what extent can classifying articles against the CAPE framework be carried out using existing means, resources, and circumstances and without outside assistance?*

6.2.1 Efficiency, speed, or quality of implementation

The process we used is fairly efficient as we were able to successfully categorize the dataset under the CAPE components relatively quickly- in part because of the already existing dataset. We were also able to produce high quality results with limited subjectivity since we were strictly looking at what was measured.

6.2.2 Positive/negative effects of using the process

Although we did not yield earth-shattering results, we can start to see that classifying articles along the CAPE framework to look for gaps in equity-focused research is viable and produces positive effects. For instance, we can start to see the gaps in the literature in our own limited dataset. For example, there was a large gap in research that investigated access and participation. When applied to larger datasets, this categorization process will show researchers where attention needs to be focused.

6.2.3 Ability to administer the process

Once we had developed the inclusion/criteria for each category of CAPE, we were able to categorize each article with relative ease. With our formal definitions that we provide here, we anticipate that others will also be able to adapt this process and that we will be able to also use these definitions in a larger, more complex dataset.

6.2.4 Cost analysis

Few resources were necessary to conduct this analysis. Since we had easy access to the articles [ANON], there was no pay wall or cost incurred reviewing the articles (other than our time and effort costs).

7 Limitations

We specifically focused on papers that focused on girls and had no participants who were boys for this pilot study. In this regard, our dataset is limited–however, intentionally so. Still, papers that include participants who were only boys or had participants of both genders were not reviewed. As we expand our process, it may illustrate other differences between these types of studies.

As it is in our study, after the first three articles were used to baseline the analysis process, one researcher classified the papers and met with the other researcher to discuss the classifications weekly. Additional researchers might be able to provide broader or more nuanced perspective on how classifications can be performed. More researchers could also provide us with more resources to have multiple researchers independently categorize the studies and then conduct interrater reliability on the classifications to provide greater assurance that they were classified appropriately.

There may be other frameworks for analyzing the body of research to determine where the equity gaps exist in research. In the future, we plan to investigate other frameworks for scoping the literature as well to see if they might offer more information about research gaps. However, the CAPE framework holds great promise and we are more likely to expand upon this by building out further subcomponents of each of the CAPE components in our future scoping reviews–since it has been specifically created for CAPE.

Lastly, it would also be ideal to compare these gaps with what is considered to be important factors that impact student achievement and growth. Though we did not go as far as to conduct that analysis in this paper, we plan on doing so in our future research.

8 Conclusion

In this paper we have shown that it is possible to classify K-12 CS Education research papers along the CAPE framework. We collected and classified 51 CS education research papers and our analysis has revealed gaps in areas of participation and access. We also evaluated the feasibility of this process by looking at implementation and practicality. In regards to implementation, we found that the process of classifying articles against the CAPE framework can be performed successfully and with a high degree of execution. In terms of practicality, we found this process relatively efficient due to the few resources needed and the speed at which we were able to gather our dataset and classify the articles.

While these results are not representative of the entire current CS education literature, our pilot study demonstrates that it is feasible to implement this process on a larger dataset, which is the next step in our research. This would provide greater insight into where our attention needs to be focused in regards to equity driven research.

Acknowledgements

This material is based upon work supported by the U.S. National Science Foundation under Grant Nos. 2122212, 1757402, and 1933671.

References

[1] Ina Suryani Ab Rahim, Aizan Yaacob, and Noor Hashima Abd Aziz. 2015. "Indicating a research gap" in computer science research article introductions by non-native English writers. *Asian Social Science* 11, 28 (2015), 293–302.

[2] Julia Aguirre, Beth Herbel-Eisenmann, Sylvia Celedon-Pattichis, Marta Civil, Trena Wilkerson, Michelle Stephan, Stephen Pape, and Douglas H Clements. 2017. Equity within mathematics education research as a political act: Moving from choice to intentional

collective professional responsibility. *Journal for Research in Mathematics Education* 48, 2 (2017), 124–147.

[3] Alfredo J Artiles. 2019. Fourteenth annual Brown lecture in education research: Reenvisioning equity research: Disability identification disparities as a case in point. *Educational Researcher* 48, 6 (2019), 325–335.

[4] Betlihem Ayalew, Elizabeth Dawson-Hahn, Rushina Cholera, Olanrewaju Falusi, Tamar Magarik Haro, Diana Montoya-Williams, and Julie M Linton. 2021. The health of children in immigrant families: Key drivers and research gaps through an equity lens. *Academic Pediatrics* (2021).

[5] Frank Betts. 1992. How systems thinking applies to education. *Educational leadership* 50, 3 (1992), 38–41.

[6] Julie A Bianchini, Valarie L Akerson, Angela Calabrese Barton, Okhee Lee, and Alberto J Rodriguez. 2012. *Moving the equity agenda forward: Equity research, practice, and policy in science education*. Vol. 5. Springer Science & Business Media.

[7] Deborah J Bowen, Matthew Kreuter, Bonnie Spring, Ludmila Cofta-Woerpel, Laura Linnan, Diane Weiner, Suzanne Bakken, Cecilia Patrick Kaplan, Linda Squiers, Cecilia Fabrizio, et al. 2009. How we design feasibility studies. *American journal of preventive medicine* 36, 5 (2009), 452–457.

[8] Christopher Chapman and Mel Ainscow. 2019. Using research to promote equity within education systems: Possibilities and barriers. *British Educational Research Journal* 45, 5 (2019), 899–917.

[9] Camille A Farrington, Melissa Roderick, Elaine Allensworth, Jenny Nagaoka, Tasha Seneca Keyes, David W Johnson, and Nicole O Beechum. 2012. *Teaching Adolescents to Become Learners: The Role of Noncognitive Factors in Shaping School Performance–A Critical Literature Review.* ERIC.

[10] Carol L Fletcher and Jayce R Warner. 2021. CAPE: a framework for assessing equity throughout the computer science education ecosystem. *Commun. ACM* 64, 2 (2021), 23–25.

[11] Sarah Heckman, Jeffrey C Carver, Mark Sherriff, and Ahmed Al-Zubidy. 2021. A Systematic Literature Review of Empiricism and Norms of Reporting in Computing Education Research Literature. *arXiv preprint arXiv:2107.01984* (2021).

[12] Jihyun Lee and Valerie J Shute. 2010. Personal and social-contextual factors in K–12 academic performance: An integrative perspective on student learning. *Educational psychologist* 45, 3 (2010), 185–202.

[13] Okhee Lee. 2003. Equity for culturally and linguistically diverse students in science education: Recommendations for a research agenda. *Teachers College Record* 105, 3 (2003), 465–489.

[14] Sarah Theule Lubienski. 2008. Research Commentary: On Gap Gazing in Mathematics Education: The Need for Gaps Analyses. *Journal for Research in Mathematics Education* 39, 4 (2008), 350–356.

[15] Ann Macintosh, Stephen Coleman, and Agnes Schneeberger. 2009. eParticipation: The research gaps. In *International Conference on Electronic Participation*. Springer, 1–11.

[16] Monica M. McGill and Adrienne Decker. 2017. Computer Science Education Resource Center. https://csedresearch.org

[17] Monica M McGill and Adrienne Decker. 2020. A Gap Analysis of Statistical Data Reporting in K-12 Computing Education Research: Recommendations for Improvement. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 591–597.

[18] Monica M McGill, Adrienne Decker, and Zachary Abbott. 2018. Improving research and experience reports of pre-college computing activities: A gap analysis. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. 964–969.

[19] Monica M McGill, Adrienne Decker, Tom McKlin, and Kathy Haynie. 2019. A gap analysis of noncognitive constructs in evaluation instruments designed for computing education. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. 706–712.

[20] Justus J Randolph, George Julnes, Erkki Sutinen, and Steve Lehman. 2008. A methodological review of computer science education research. *Journal of Information Technology Education: Research* 7, 1 (2008), 135–162.

[21] Devashree Saha. 2009. Empirical research on local government sustainability efforts in the USA: gaps in the current literature. *Local Environment* 14, 1 (2009), 17–30.

[22] Kate Sanders, Judy Sheard, Brett A Becker, Anna Eckerdal, and Sally Hamouda. 2019. Inferential Statistics in Computing Education Research: A Methodological Review. In *Proceedings of the 2019 ACM Conference on International Computing Education Research*. 177–185.

[23] Julie M Smith. 2020. The Data Gap: A Potential Barrier to Gender Equity in Computer Science Education. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 1426–1426.

[24] Ryan Torbey, Nicole D Martin, Jayce R Warner, and Carol L Fletcher. 2020. Algebra I Before High School as a Gatekeeper to Computer Science Participation. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 839–844.

[25] Bishakha Upadhyaya, Monica M McGill, and Adrienne Decker. 2020. A longitudinal analysis of k-12 computing education research in the United States: Implications and recommendations for change. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 605–611.