A phenomenographical study of instructor perceptions and practices in multilingual computing classrooms in India

Anon Author
University of Anon
Anonville, Rep. of Anon
anon@anonuni.edu

Anon Author
University of Anon
Anonville, Rep. of Anon
anon@anonuni.edu

Anon Author
University of Anon
Anonville, Rep. of Anon
anon@anonuni.edu

Anon Author
University of Anon
Anonville, Rep. of Anon
anon@anonuni.edu

Anon Author
University of Anon
Anonville, Rep. of Anon
anon@anonuni.edu

Anon Author
University of Anon
Anonville, Rep. of Anon
anon@anonuni.edu

Abstract

Background:
Objective:
Method:
Results:

CCS Concepts

• **Social and professional topics → Computing education**; • **Computing methodologies → Artificial intelligence**.

ACM Reference Format:
Anon Author, Anon Author, Anon Author, Anon Author, Anon Author, and Anon Author. 2024. A phenomenographical study of instructor perceptions and practices in multilingual computing classrooms in India. In *Proceedings of the 2024 Conference on Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2024), July 8–10, 2024, Milan, Italy*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/XXXXXX.XXXXXX

1 Introduction

The rise of generative AI (GenAI) tools is fundamentally transforming education and how it is practiced across various disciplines [5, 25] and leading institutions to integrate it into all elements of their educational models [31]. In computing education, there has been a focus on how the use of these tools may cause writing code from scratch to take a backseat to This shift brings with it both opportunities—namely the ability to enable student to engage inmore personally meaningful and complex projects [6, 24]–and challenges related to the importance of teaching students to clearly articulate requests and requirements in natural language [3, 11, 27, 29]. Given the existing prevelance of English as the dominant medium of instruction in computing education [15], the multilingual nature of the state-of-the-art GenAI models presents both opportunities to overcome existing inequities as well as challenges that may exacerbate them [26, 30, 35].

ITiCSE 2024, Milan, Italy

ACM ISBN 979-8-4007-XXXX-X/XX/XX
https://doi.org/10.1145/XXXXXX.XXXXXX

Recent work by Jacob et al. [16] and Cheung and Ng [9] notes there is little work on the experiences of multilingual students in computing education—and those that do appear to focus on the North American contex. Additionally, Cheung and Ng [9] notes that the majority of work on translanguaging is focused on science and math education, with only a small fraction of work (6 studies) focused on engineering disciplines. As computing education continues to shift towards natural language tools for the dearth of work informing how to best support multilingual students in computing education becomes more pressing.

India presents a unique context for examing these dynamics. The country simultaneously represents one of the world's largest technical education systems [12] and one of the most linguistically diverse, with 22 official languages and over 700 documentated languages—and many more dialects that vary in differences of degree—that are spoken across its states [7]. This linguistic diversity, in the computing education context, embodies what Mohanty and Panda [22] refered to as a "double divide" between English and dominant regional languages, as well as between dominant and minority regional languages. The context is further complicated by the paradox of the prevelance of English in programming languages—along with the resources to learn them—and the emphasis on mother-tongue instruction that both India's National Education Policy (NEP) 2020 and the UNESCO's longstanding guidance on education policy have advocated for [21, 34].

Though this gap should always have been addressed, the rise and prevelance of GenAI tools in computing education brings with it a new found urgency to address the issue of linguistic diversity in computing education. The primary means of interacting with GenAI models is through natural language. As such, this could deepen existing inquities between those who are fluent in English—or other dominant languages of instruction in India—and those who are not. Further more, there is the overarching issue of the performance of these models in languages that are not as well represented in their training data, which could further exacerbate existing inequities.

To address these gaps, this paper presents a phenomenographics study of computing instructors' perceptions of and experiences with linguistic diversity in Indian higher education. Drawing on semi-structured interviews with XX instructors across XX institutions in XX states, we analyze how educators navigate the realities of linguistic diversity in their classrooms. Our analysis employs

theoretical frameworks from translanguaging pedagogy, culturally relevant education, and critical language awareness to interpret instructor practices and their potential for creating equitable learning environments. This study aims to advance the theoretical understanding of linguistic diversity in computing and technology education while offering practical insights for more equitable pedagogy in an era where natural language mediates the act of computing itself.

2 Background

To situate the work presented in this paper, we provide an overview of three key areas of literature. First, in Section 2.1, we examine the Indian higher education system, focusing on how multilingualism shapes both educational policy and classroom practice. Second, in Section 2.2, we explore the evolving relationship between language and computing education, particularly the shift towards natural language programming interfaces and its implications for non-English speakers in India. Finally, in Section 2.3, we review theoretical frameworks and pedagogical approaches for leveraging linguistic diversity in educational contexts, including translanguaging, culturally relevant pedagogy, and critical language awareness.

2.1 Indian Higher Education & Language

Here, we provide a broad overview of Indian higher education both from the perspectives of policy and practice. In Sections 2.1.1 and 2.1.2, we discuss the multilingual nature of India and Indian education, respectively. We follow this with an overview of empirical work on language and Indian higher education in Section 2.1.3.

2.1.1 Multilingual Landscape. The linguistic landscape of India is incredibly diverse, to such an extent that it can be difficult to exactly quantify the number of languages spoken in the country. India has 22 scheduled languages—meaning they are officially recognized by the government and have a special status in the constitution—and many hundreds of other languages and dialects that are spoken across its 28 states and 8 union territories [7]. TODO: Add some more of Mohanty and alls work on characterizing the landscape.

2.1.2 Language and Education Policy. India's National Education Policy (NEP) of 2020 [21] emphasizes the importance of of equity and the centrality of language in achieving equitable education.

> "*The aim must be for India to have an education system by 2040 that is second to none, with equitable access to the highest quality education for all learners regardless of social or economic background*"(p. 3)

In previous editions of the NEP, which took places in 1968 and 1980, the three-language formula has been a key policy in Indian education system—though it has undergone considerable evolution in its most recent iteration (Table 1). This policy mandates that students learn three languages where the specific languages differed based primarily on whether a student lived in a Hindi or non-Hindi speaking state. The 2020 edition of the NEP, though continuing to emphasize the importance of learning three languages introduces greater flexibility in which languages fit into each category and leaves it to the state or region to make these policy decision.

TODO: Add some context regarding praise and criticism of the NEP 2020. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.1.3 Empirical Work. TODO: Fill in he K-12 education stuff I've found. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

TODO: Fill in Gerald's work on Tamil in higher education. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.2 Language and Computing Education

Programming, once the domain of structured syntax and rigid semantics, is increasingly shifting towards natural language interfaces for program specification [8, 20, 23, 25]—courtesy of the advanced program synthesis capabilities of large language models [1, 18]. The prevalence of English vernacular in programming language keywords and documentation has long been a barrier to entry for non-English speakers [15]. However, the shift towards natural language prompts—brought about courtesy of multilingual GenAI models, presents both new opportunities as well as new challenges for non-English speakers. Below, we highlight some of the key challenges non-English speaking novice programmers have faced historically as well as some of the opprotunities that may arise as programming moves towards natural language interfaces.

NEP Edition	**1968 & 1986** [13, 14]	**2020** [21]
Hindi-Speaking States	Hindi, English, and a modern Indian language (preferably from Southern India)	Three languages of the student's or state's choice (at least two must be native to India)
Non-Hindi Speaking States	Regional language, Hindi, and English	Three languages of the student's or state's choice (at least two must be native to India)
Key Change	1968: First formal introduction of the formula to promote national integration 1986: Reiterated the 1968 formula without changes	Introduced flexibility, allowing students to choose the languages, with the condition that two of the three languages must be Indian

Table 1: Evolution of the Three-Language Formula in National Education Policies

2.2.1 Potential Challenges for non-English Speakers. As noted, by Guo [15], the vast majority of programming languages—and almost certainly all widely adopted programming languages—rely on English keywords. Stefik and Siebert [32] compared error rates by novice programmers among a variety of programming languages (i.e., Quorum, Python, Java, C, Pearl) to that of a language that used random keywords (i.e., Randomo). They found that languages that used syntax that more naturally resembled natural language (i.e., Quorum and Python) had lower error rates and that students had similar error rates for Randomo and languages like Java. This not only highlights the role that syntax plays but, as hypothesized by Becker [2], that non-English speakers may not be able to share in the affordances of languages such as Quorum and Python. This hypothesis was, in part, confirmed by the work of Dasgupta and Hill [10] which found that Scratch users coding in their localized native language demonstrate new programming concepts at a faster rate than users from the same countries using English interfaces, controlling for activity levels and socialization.

As noted—somewhat prophetically in retrospect given it was published before widely avail be GenAI models—by Becker [2] as programming language design moves away from random keywords, to keywords that invoke natural language, towards fully natural language interface, the difficulty gap in acquiring that language will become steadily larger and favor the user who is already familiar with the natural language which the interface is based on.

> "However, imagine if perfect natural (English) language programming was achievable today."

Imagine indeed. However, we are fortunate in that large language models are trained on a wide variety of languages, suggesting that interfacing with these models in non-English languages may be possible and thus lower the barrier to entry for non-English speakers.

2.2.2 Opportunities for non-English Speakers. Perhaps the immediate, and most obvious, opportunity is the ability to interact with programming environments in one's native language [26, 30]. As noted by

2.2.3 India Specific Opportunities and Challenges. However, even prior to the advent of GenAI models, it is important to understand that Natural Language Processing (NLP) technologies have long faced challenges in effectively supporting Indic languages. Bhattacharyya et al. [4] characterized a wide variety of challenges relating to Indic languages and computing. These included,

(1) **Scale and Diversity** Indic languages encompass a vast array of languages and dialects, belonging to multiple linguistic families and written in numerous distinct scripts.
(2) **Longer Utterances** Sentences in Indic languages are often longer and more complex than in English, complicating tasks like parsing and speech recognition.
(3) **Code Mixing** The frequent mixing of multiple languages in a single sentence or conversation is a common challenge in computational linguistics for the region.
(4) **Resource Scarcity** Many Indic languages lack sufficient annotated datasets for building robust NLP and speech tools.
(5) **Absence of linguistics knowledge** A limited understanding of the linguistic structure of many regional languages hinders the development of computational models.
(6) **Script complexity and non-standard input mechanisms** The diversity of scripts and their associated vowel and consonant combinations make input systems slower and less intuitive.
(7) **Non-standard transliteration** Roman transliteration of Indic languages lacks standardization, leading to multiple ways of representing the same word.

Recent work indicates some of these issues are far from resolved. Jordan et al. [19] found that problem statements for programming problems generated in Tamil by GPT-3.5 were often non-sensical and hypothesized this may be due to the non-latin alphabet or insufficient training data. Smith IV et al. [30] had similar findings with Tamil translations of correct English descriptions of code having far less success in generating correct code—using GPT-4o—than 10 other of he most commonly spoken languages in India. These findings highlight that, though the multilingual nature of the current state of the art GenAI models presents new opportunities for non-English speakers, these opportunities are not necessarily equitably distributed across all languages. This is of particular concern in the Indian context where minority and tribal languages are already at a systemic disadvantage in the education system [22] and are almost certainly underrepresented in the training data of these models.

2.3 Pedagogy and Linguistic Diversity

To understand how instructors navigate multilingual computing classrooms, we examine pedagogical frameworks that reconceptualize linguistic diversity as a resource rather than a barrier in educational settings (Sections 2.3.1–2.3.4). These theoretical perspectives—including translanguaging, code-switching, culturally relevant pedagogy, and critical language awareness—provide lenses for interpreting instructor practices and their potential to create more equitable learning environments in India's linguistically diverse computing education landscape.

2.3.1 Theories of Translanguaging. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Though analysis of translanguaging in programming education is limited, there is one notable exception. Tai [33] introduced a theory of *transprogramming*. They expand on the idea of translanguaging to describe the process by which teachers and students navigate multilingual, multimodal, and programming language resources to accomplish programming tasks and develop mental models of of computational concepts.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.3.2 Code-switching and Code-meshing. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.3.3 Funds of Knowledge and Culturally Relevant Pedagogy. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2.3.4 Critical Language Awareness. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

3 Methods

We developed a semi-structured interview protocol to explore how educators across a variety of tertiary insitutions. In Section **??**, we describe the interview questions that we developed and in Subsection 3.2 we describe the taxonomy of interview probes we empoloyed to guide consistent follow-up questions across interviews. Following this, in Section 3.3 we describe our recruitment and data collection procedures. Finally, in Section 3.4 we describe our approach to analyzing the data collected through these interviews.

3.1 Interview Questions

3.1.1 Persona and Positionality. Given the phenomenographic nature of this work, it is important that we understand the context and positionality of each participant. These questions are designed to provide this context—with particular focus on the educational, instructional, and linguistic backgrounds of each participant. The questions, in this regard, are as follows:

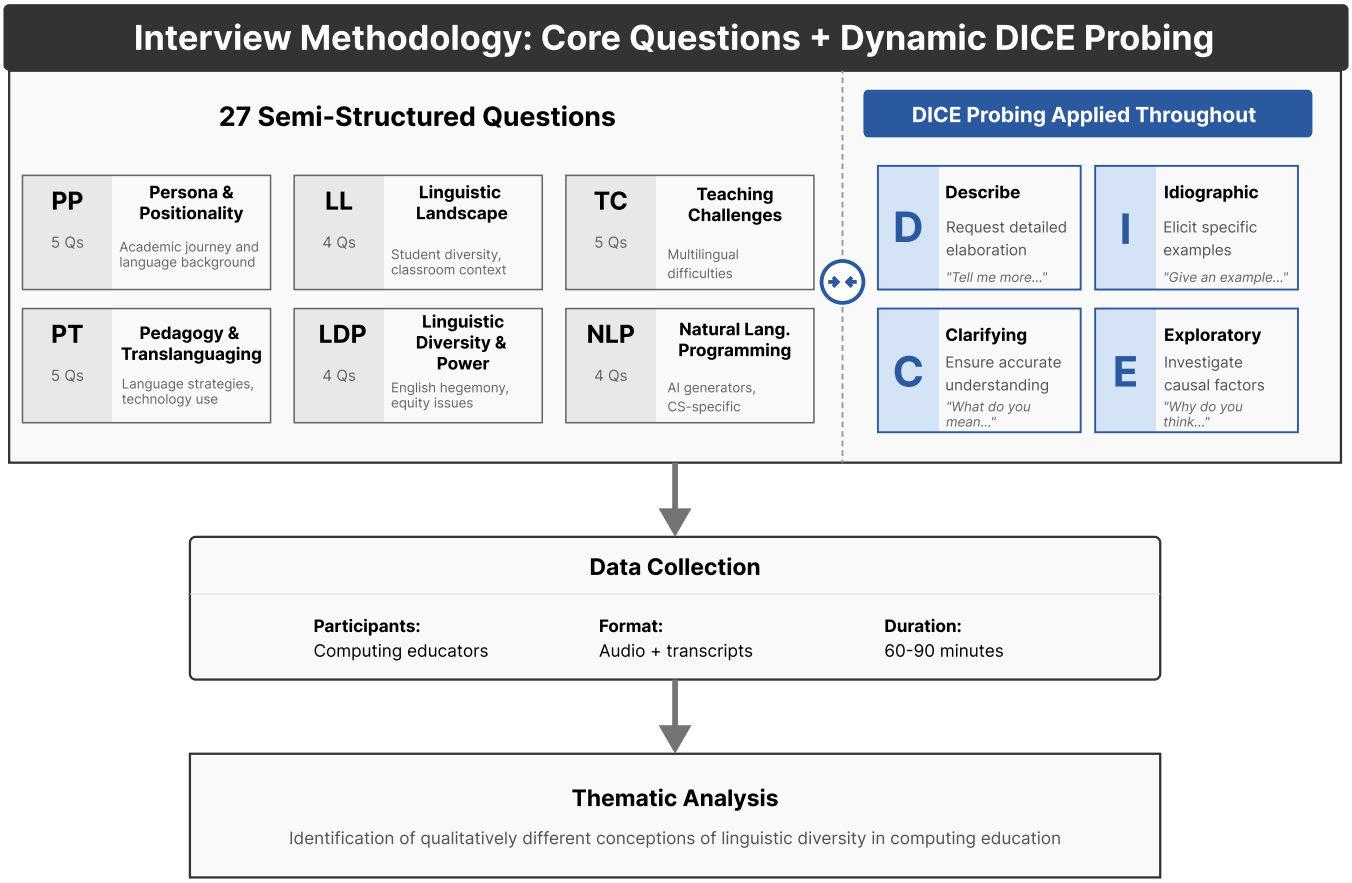

Figure 1: Your caption here.

PP.1 Could you briefly describe your academic journey and how you came to teach in your current position?
PP.2 What subjects do you currently teach, what levels (e.g., undergraduate, graduate) are these courses, and in what contexts (e.g., in-person, online, hybrid) are they delivered?
PP.3 What languages do you speak, and in what contexts do you typically use each?
PP.4 In your own education, what language(s) were used for instruction?
PP.5 What are your general thoughts and beliefs about the role of language in education, particularly in technical fields like computer science?

Probes during this section of the interview focused on understanding TBD.

3.1.2 Linguistic Landscape. Related to understanding the persona of each participant, it is essential to underastand the linguistic landscape in which they operate. This includes not only the languages spoken by their students, but also the languages used in instruction

LL.1 How would you describe the linguistic diversity of your current students?
LL.2 Your second question here

3.1.3 Teaching Challenges.

TC.1 What are the biggest challenges you face teaching topics in computing to students who speak a variety of languages?
TC.2 Can you describe

3.1.4 Pedagogy and Translanguaging. In this section, we explore the intersection of translanguaging practices and pedagogical strategies employed by educators in multilingual settings. <Resaerch on india specifically and findings that inform these questions>.The questions are designed to uncover how educators navigate—or leverage—the linguistic diversity of their classrooms.

PT.1 How do you decide when to use a particular language for instruction or explanation during your classes and when, if at all, do to change languages?
PT.2 If you've adopted a teaching strategy that helps you navigate linguistic diversity, can you describe it and why you find it effective?
PT.3 How do you use technology, if at all, to support students with different language backgrounds?

3.1.5 Linguistic Diversity and Power. Central to language is the notion of power. The use of language carries with it implications of power, privilege, and access. As noted by Jhingran [17],

> "*English is seen as the language of power and the vehicle for getting better jobs. Even poor families in urban areas aspire to send their children to these private English-medium schools.*"

Similarly, as noted by Mohanty and Panda [22], beyond the 22 scheduled languages, there are little instruction at the primary and secondary levels of education which creates a systemic hierarchy of language: English and Hindi at the top, regional scheduled languages in the middle, and minority and tribal languages at the bottom. In this section, we explore how educators in computing education perceptions and experiences with the role of power and privilege in the context of linguistic diversity.

- LDP.1 How do you perceive the role of English in your institution and field of study?
- LDP.2 Have you observed any instances where language has influenced students' academic performance or participation in class? Can you describe one such instance?
- LDP.3 How do you address issues of linguistic equity in your classroom?
- LDP.4 What are your thoughts on the use of local languages versus English in higher education, particularly in technical fields like computer science?

3.1.6 Natural Language Programming. This section explores the perceptions and experiences of educators regarding the use of natural language programming tools, such as AI-driven code generators, in multilingual educational settings. The questions aim to uncover how these tools are integrated into teaching practices and their impact on students with diverse linguistic backgrounds.

- NLP.1 What concerns or difficulties have you encountered with respect to linguistic diversity that you feel are specific to teaching topics in computing?
- NLP.2 Have you integrated or do you have plans to integrate any natural language programming tools (e.g., AI code generators) into your teaching?

3.2 Interview Probes

In conducting these interviews we use probes aligned with the DICE (*"Describe, Idiographic, Clarifying, Exploratory"*) taxonomy as described in Robinson [28]. In more detail, the four types of probes are:

(1) **Describe Probe:** These probes ask that participants provide more detail about a specific situation or experience they mentioned. These probes often take the form of *"Tell me more about..."* or *"Do you recall what was happening when..."*.
(2) **Idiographic Probe:** These probes encourage participants to share a specific example from a specific period of time. These probes often take the form of *"Can you give me an example of..."* or *"Do you recall what was happening the week you..."*.
(3) **Clarifying Probe:** These probes are used to ensure that the interviewer accurately understands what the participant is saying. These probes often take the form of *"What do you mean by..."* or *"Can you expand on..."*.
(4) **Exploratory Probe:** These probes encourage participants to think about the causal factors or underlying reasons behind their thoughts, opinions and experiences. Such questions take the form of *"Why do you that happened?"* or *"What do you think led to..."*.

This taxonomy is designed to help interviewers structure follow-up questions in a way that encourages participants to provide rich, detailed responses.

3.3 Data Collection

3.4 Data Analysis

4 Results

5 Discussion

6 Limitations

7 Conclusion

References

[1] Jacob Austin, Augustus Odena, Maxwell Nye, Maarten Bosma, Henryk Michalewski, David Dohan, Ellen Jiang, Carrie Cai, Michael Terry, Quoc Le, et al. 2021. Program synthesis with large language models. *arXiv preprint arXiv:2108.07732* (2021).

[2] Brett A Becker. 2019. Parlez-vous Java? Bonjour La Monde!= Hello World: Barriers to Programming Language Acquisition for Non-Native English Speakers.. In *PPIG*.

[3] Brett A Becker, Paul Denny, James Finnie-Ansley, Andrew Luxton-Reilly, James Prather, and Eddie Antonio Santos. 2023. Programming is hard-or at least it used to be: Educational opportunities and challenges of ai code generation. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*. 500–506.

[4] Pushpak Bhattacharyya, Hema Murthy, Surangika Ranathunga, and Ranjiva Munasinghe. 2019. Indic language computing. *Commun. ACM* 62, 11 (2019), 70–75.

[5] Melissa Bond, Hassan Khosravi, Maarten De Laat, Nina Bergdahl, Violeta Negrea, Emily Oxley, Phuong Pham, Sin Wang Chong, and George Siemens. 2024. A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International journal of educational technology in higher education* 21, 1 (2024), 4.

[6] Axel Böttcher, Veronika Thurner, and Benedikt Zönnchen. 2025. Concepts for Teaching Software Development in the Age of AI-Tools. In *2025 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 1–10.

[7] C Chandramouli and Registrar General. 2011. Census of india. *Rural urban distribution of population, provisional population total. New Delhi: Office of the Registrar General and Census Commissioner, India* (2011).

[8] Xiang Echo Chen, Wenhan Zhu, Guoshuai Albert Shi, and Michael W Godfrey. 2025. An Empirical Study of GenAI Adoption in Open-Source Game Development: Tools, Tasks, and Developer Challenges. *arXiv preprint arXiv:2507.18029* (2025).

[9] Kason Ka Ching Cheung and Davy Tsz Kit Ng. 2025. A systematic review of research on translanguaging in STEM education. *International Journal of Multilingualism* (2025), 1–18.

[10] Sayamindu Dasgupta and Benjamin Mako Hill. 2017. Learning to code in localized programming languages. In *Proceedings of the fourth (2017) ACM conference on learning@ scale*. 33–39.

[11] Paul Denny, Juho Leinonen, James Prather, Andrew Luxton-Reilly, Thezyrie Amarouche, Brett A Becker, and Brent N Reeves. 2024. Prompt Problems: A new programming exercise for the generative AI era. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. 296–302.

[12] Dr Vijay P Goel. 2017. Echnical and vocational education and training (tvet) system\in india for sustainable development. (2017).

[13] Government of India. 1968. *National Policy on Education.* Policy Document. Ministry of Education, New Delhi.

[14] Government of India. 1986. *National Policy on Education 1986.* Policy Document. Ministry of Human Resource Development, New Delhi.

[15] Philip J Guo. 2018. Non-native english speakers learning computer programming: Barriers, desires, and design opportunities. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–14.

[16] Sharin Rawhiya Jacob, Jonathan Montoya, Ha Nguyen, Debra Richardson, and Mark Warschauer. 2022. Examining the what, why, and how of multilingual student identity development in computer science. *ACM Transactions on Computing Education (TOCE)* 22, 3 (2022), 1–33.

[17] Dhir Jhingran. 2009. Hundreds of home languages in the country and many in most classrooms: Coping with diversity in primary education in India. *Social justice through multilingual education* 250 (2009), 267.

[18] Carlos E Jimenez, John Yang, Alexander Wettig, Shunyu Yao, Kexin Pei, Ofir Press, and Karthik R Narasimhan. [n. d.]. SWE-bench: Can Language Models Resolve Real-world Github Issues?. In *The Twelfth International Conference on Learning Representations*.

[19] Mollie Jordan, Kevin Ly, and Adalbert Gerald Soosai Raj. 2024. Need a programming exercise generated in your native language? chatgpt's got your back: Automatic generation of non-english programming exercises using openai gpt-3.5. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. 618–624.

[20] Sam Lau and Philip Guo. 2023. From" Ban it till we understand it" to" Resistance is futile": How university programming instructors plan to adapt as more students use AI code generation and explanation tools such as ChatGPT and GitHub Copilot. In *Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 1*. 106–121.

[21] Ministry of Education, Government of India. 2020. National Education Policy 2020. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf. Accessed: 2025-09-20.

[22] Ajit K Mohanty and Minati Panda. 2017. Language policy and education in the Indian subcontinent. In *Language policy and political issues in education*. Springer, 507–518.

[23] Olga Petrovska, Lee Clift, Faron Moller, and Rebecca Pearsall. 2024. Incorporating generative AI into software development education. In *Proceedings of the 8th Conference on Computing Education Practice*. 37–40.

[24] Leo Porter and Daniel Zingaro. 2024. *Learn AI-assisted Python programming: with github copilot and ChatGPT*. Simon and Schuster.

[25] James Prather, Juho Leinonen, Natalie Kiesler, Jamie Gorson Benario, Sam Lau, Stephen MacNeil, Narges Norouzi, Simone Opel, Vee Pettit, Leo Porter, et al. 2025. Beyond the hype: A comprehensive review of current trends in generative AI research, teaching practices, and tools. *2024 Working Group Reports on Innovation and Technology in Computer Science Education* (2025), 300–338.

[26] James Prather, Brent N Reeves, Paul Denny, Juho Leinonen, Stephen MacNeil, Andrew Luxton-Reilly, João Orvalho, Amin Alipour, Ali Alfageeh, Thezyrie Amarouche, et al. 2025. Breaking the programming language barrier: Multilingual prompting to empower non-native English learners. In *Proceedings of the 27th Australasian Computing Education Conference*. 74–84.

[27] Brent N Reeves, James Prather, Paul Denny, Juho Leinonen, Stephen MacNeil, Brett A Becker, and Andrew Luxton-Reilly. 2024. Prompts first, finally. *arXiv preprint arXiv:2407.09231* (2024).

[28] Oliver C Robinson. 2023. Probing in qualitative research interviews: Theory and practice. *Qualitative Research in Psychology* 20, 3 (2023), 382–397.

[29] David H Smith IV, Paul Denny, and Max Fowler. 2024. Prompting for comprehension: Exploring the intersection of explain in plain english questions and prompt writing. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale*. 39–50.

[30] David H Smith IV, Viraj Kumar, and Paul Denny. 2024. Explain in plain language questions with Indic languages: Drawbacks, affordances, and opportunities. In *Annual ACM India Compute Conference*. Springer, 3–17.

[31] Jane Southworth, Kati Migliaccio, Joe Glover, Ja'Net Glover, David Reed, Christopher McCarty, Joel Brendemuhl, and Aaron Thomas. 2023. Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artificial Intelligence* 4 (2023), 100127.

[32] Andreas Stefik and Susanna Siebert. 2013. An empirical investigation into programming language syntax. *ACM Transactions on Computing Education (TOCE)* 13, 4 (2013), 1–40.

[33] Kevin WH Tai. 2024. Transprogramming in the primary-level programming lessons: reconceptualizing translanguaging in the era of artificial intelligence. *Applied Linguistics* (2024), amae076.

[34] UNESCO. 1953. The Use of Vernacular Languages in Education. 8 (1953). http://unesdoc.unesco.org/images/0000/000028/002897eb.pdf

[35] Zheng-Xin Yong, Ruochen Zhang, Jessica Zosa Forde, Skyler Wang, Arjun Subramonian, Holy Lovenia, Samuel Cahyawijaya, Genta Indra Winata, Lintang Sutawika, Jan Christian Blaise Cruz, et al. 2023. Prompting multilingual large language models to generate code-mixed texts: The case of south East Asian languages. *arXiv preprint arXiv:2303.13592* (2023).