What are We Asking our Students?
A Literature Review of Student Surveys in Computer Science Education

Angela Zavaleta Bernuy

Department of Computer and Mathematical Sciences, Univeristy of Toronto Scarborough

UNIVERSITY OF TORONTO SCARBOROUGH

Abstract

Many CS education research papers pull data from student surveys. But are those surveys well designed? Are the questions used validated? Are the results comparable across studies? What exactly are we asking our students? In this work, we performed a systematic literature review of the past 15 years of papers in the three main conferences sponsored by the ACM Special Interest Group on Computer Science Education. Out of 1313 papers analyzed, 42 papers referred to general questions applicable to many or all computer science students. Our analysis showed that many papers were using surveys to extract similar types of information, such as demographics, prior experience or motivation to study computer science. However, the questions were being asked in different ways, using different scales, thus making it difficult or impossible to compare survey results between studies. We further found that while some studies based their questions on well-validated surveys, or shared their questions for possible later validation, approximately half of the papers retrieved neither validated their questions nor shared them to allow for post-hoc validation.

Problems and Motivation

- Surveys are usually used by researchers to analyze the demographics, background, opinions, and attitudes of participants in their studies.
- The way questions are stated and asked are crucial factors in any study that uses survey data, but not all surveys are validated.
- There is a great deal of literature in sociology and psychology dealing with survey biases that detail the myriad of ways that questions can lead respondents into giving skewed or even contradictory answers [7].
- These influences can often be subtle and pernicious and can lead to skewed results even in seemingly innocuous questions [5].
- Computer Science education has struggled with providing validated tools for the research community [6], and the use of unvalidated and biased tools could lead to a loss of confidence in the research community and ultimately a replication crisis [1].
- It is essential to have standardized and replicable questions in surveys so that results can be compared to existing literature, and shared in a way that can be easily utilized for future studies.

Background and Related Work

- There are several general-purpose surveys available to researchers that focus on the attitudes and motivations of computer science students like the Computing Research Association Evaluation (CERP) and the Computing Attitudes Survey (CAS).
- CERP started a project called Data Buddies to collect survey data from students in computing; however, the tool is not publicly available which does not facilitate its use for research purposes [2].
- CAS is based on the Colorado Learning Attitudes about Science Survey (CLASS) that measures novice to expert-like perceptions about Computer Science [3]. Despite being validated and providing a well-explained methodology for question creation, CAS is far from the standard in the computer science research community.
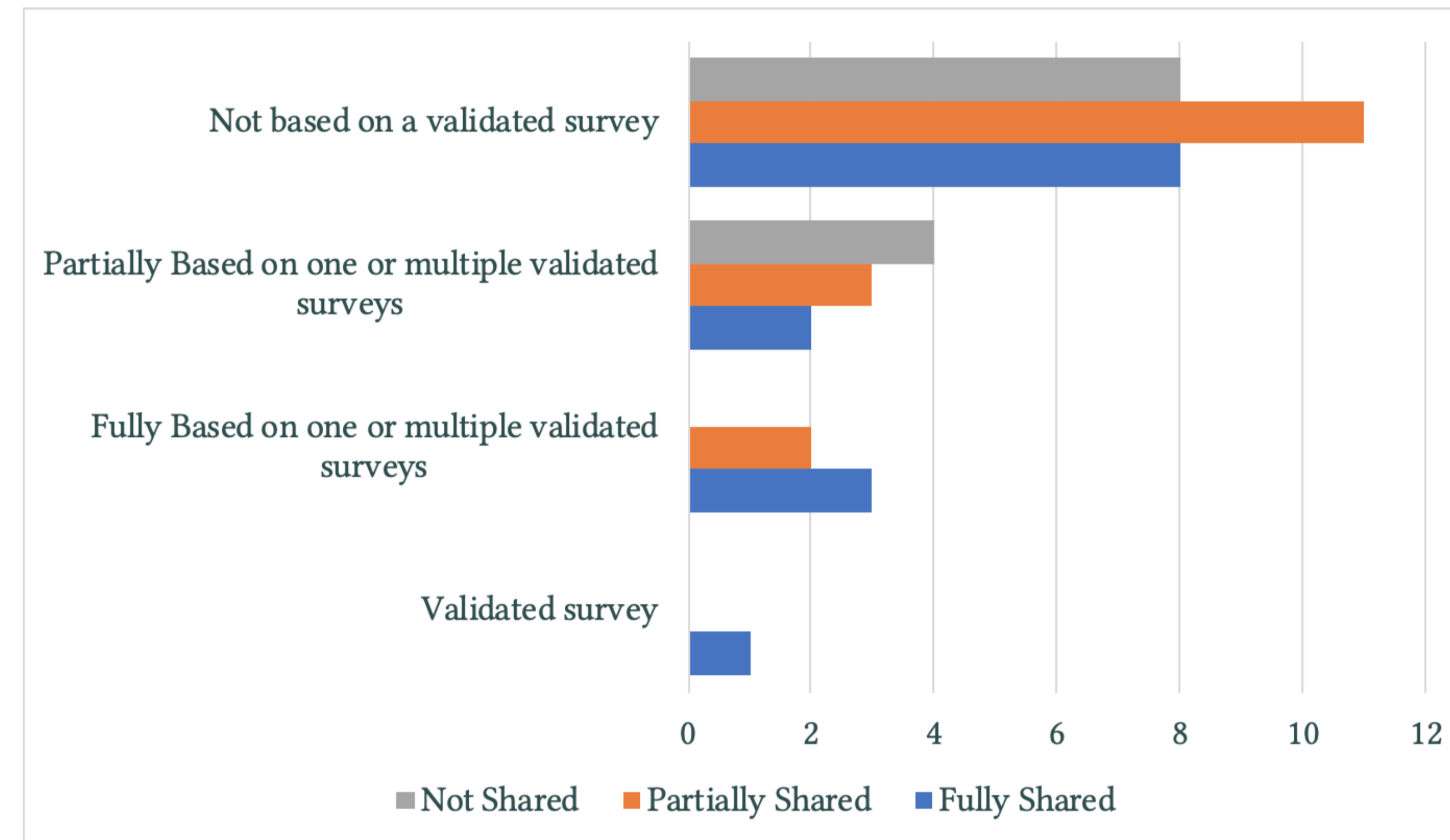

Figure 1: Papers categorized by validation and availability of survey instrument.

Approach and Uniqueness

- We performed a systematic literature search, in the model of Kitchenham and Charters [4], of all papers published in ICER, ITiCSE and SIGCSE in the past 15 years to identify common patterns and modalities of questions in surveys of student demographics, backgrounds and attitudes.
- We analyzed which papers used validated surveys, share survey questions, or asked questions in a manner that can be easily shared with and analyzed by the research community.
- To our knowledge, this is the first study to attempt to systematically review CS education literature and to analyze and quantify the pervasiveness of unsound research methodology concerning the design, implementation and reporting of survey data.
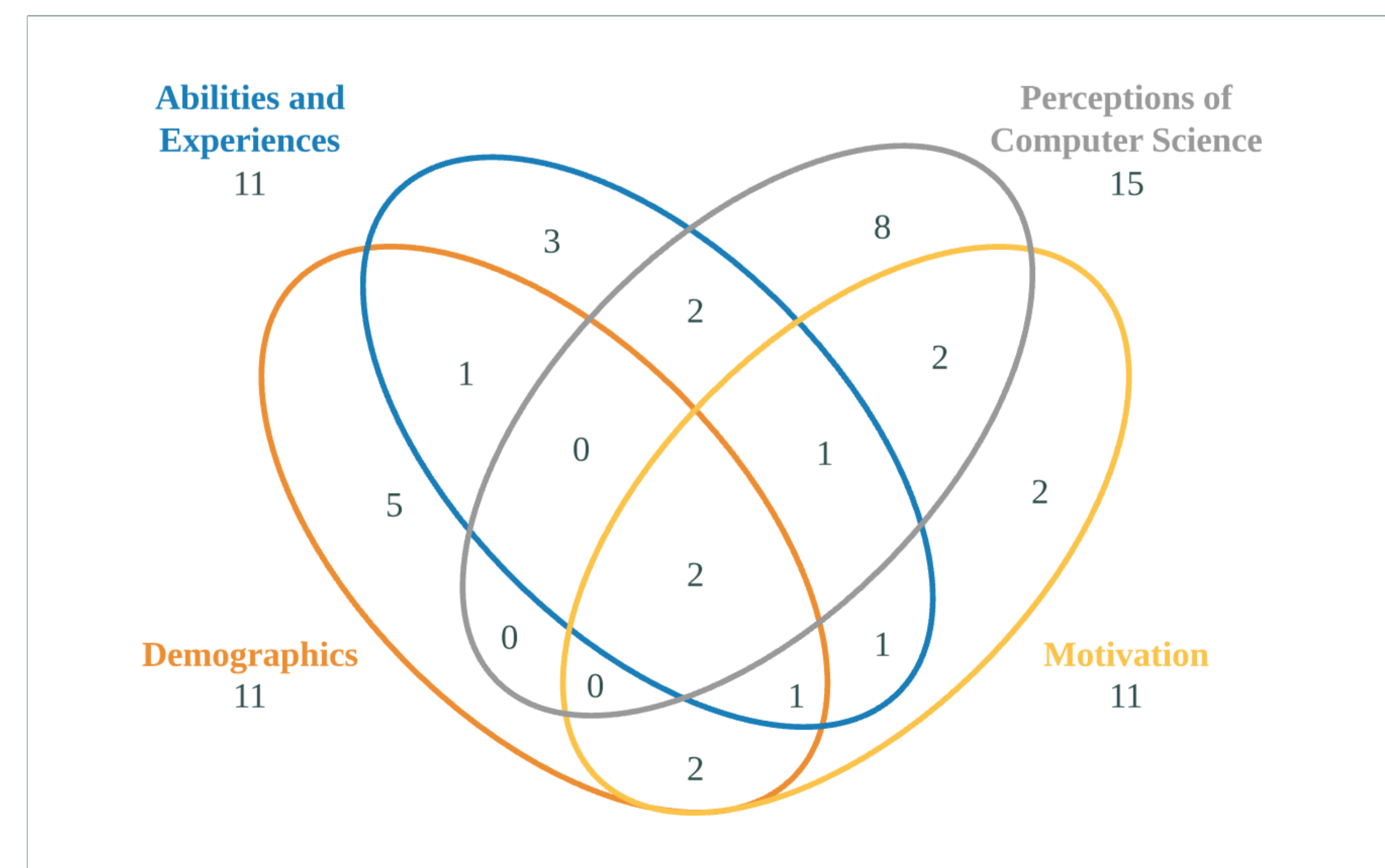
Figure 2: Papers categorized by survey topic(s) covered.

Question	Considerations
I have always wanted to **be in** computer science	Job prospects and future achievement
I have always wanted to **learn** computer science	Educational attainment and learning process

Figure 3: Example of subtle differences in Likert Scale questions.

Results and Contributions

- Out of the papers reviewed, 44 were found that used or referred to general surveys conducted to students.
- The majority of surveys found were unvalidated, failed to share their survey questions in a manner that would allow for post-hoc validation, did not discuss the design of the survey questions, and did not attempt to standardize their questions **(Figure 1)**.
- 30/42 papers shared their surveys partially or fully. We extracted the questions from the shared surveys and categorized them into similar topics.
- We found that most questions fell into four major topics of interest: Demographics, Abilities and Experiences, Perceptions of Computer Science, and Motivation **(Figure 2)**.
- Many of the surveys found solicited similar information, but between study data comparison was found to be difficult due to differences in wording and format of the questions.
- The example in **Figure 3** shows that very subtle differences in questions can get students to consider different factors that can lead to biases in responses.
- The lack of standardized, validated and publicly shared questions led to results that are not comparable across studies. The aforementioned reasonings also reveal subtle and pernicious biases.
- It is important that researchers attempt whenever possible to either use validated instruments or validate their own, and to share their survey questions to allow both post-hoc validation, easy replication, and meta-analysis.

References

[1] Alireza Ahadi, Arto Hellas, Petri Ihantola, Ari Korhonen, and Andrew Petersen. Replication in computing education research: researcher attitudes and experiences. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research*, pages 2–11. ACM, 2016.

[2] CERP. *Data Buddies*, 2018.

[3] Allison Elliott Tew, Brian Dorn, and Oliver Schneider. Toward a Validated Computing Attitudes Survey. In *Proceedings of the Ninth Annual International Conference on International Computing Education Research*, ICER '12, pages 135–142, New York, NY, USA, 2012. ACM. Auckland, New Zealand.

[4] Barbara Kitchenham and Stuart Charters. Guidelines for performing systematic literature reviews in software engineering. 2007.

[5] Diana Matovic and Joseph P Forgas. The answer is in the question? mood effects on processing verbal information and impression formation. *Journal of Language and Social Psychology*, 37(5):578–590, 2018.

[6] Miranda C Parker, Mark Guzdial, and Shelly Engleman. Replication, validation, and use of a language independent cs1 knowledge assessment. pages 93–101, 2016.

[7] Nora Cate Schaeffer and Stanley Presser. The science of asking questions. *Annual review of sociology*, 29(1):65–88, 2003.