Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
ASEE
Paper ID #47545

# Exploring the Cognitive Complexity of K-12 CS Standards (Fundamental)

**Dr. Julie M. Smith, Institute for Advancing Computing Education**

Dr. Julie M. Smith is a senior education researcher at the Institute for Advancing Computing Education. She holds degrees in Curriculum & Instruction and Software Development. She also completed a doctoral program in Learning Technologies. Her research focus is computer science education, particularly the intersection of learning analytics, learning theory, and equity and excellence. She was a research assistant at MIT's Teaching Systems Lab, working on a program aimed at improving equity in high school computer science programs; she is also co-editor of the SIGCSE Bulletin.

**Jacob Koressel**
**Sofia De Jesus, Carnegie Mellon University**
**Joseph W Kmoch**
**Bryan Twarek**

# Exploring the Cognitive Complexity of K-12 CS Standards (Fundamental)


**Abstract**

Introduction: State and national learning standards play an important role in articulating and standardizing K-12 computer science education. However, these standards have not been extensively researched, especially in terms of their cognitive complexity. Analyses of cognitive complexity, accomplished via comparison of standards to a taxonomy of learning, can provide an important data point for understanding the prevalence of higher-order versus lower-order thinking skills in a set of standards.

Objective: The objective of this study is to answer the research question: *How do state and national K-12 computer science standards compare in terms of their cognitive complexity?*

Methods: We used Bloom's Revised Taxonomy in order to assess the cognitive complexity of a dataset consisting of state ($n = 9695$) computer science standards and the 2017 Computer Science Teachers Association (CSTA) standards ($n = 120$). To enable a quantitative comparison of the standards, we assigned numbers to the Bloom's levels.

Results: The CSTA standards had a higher average level of cognitive complexity than most states' standards. States were more likely to have standards at the lowest Bloom's level than the CSTA standards. There was wide variety of cognitive complexity by state and, within a state, there was variation by grade band. For the states, standards at the *evaluate* level were least common; in the CSTA standards, the *remember* level was least common.

Discussion: While there are legitimate critiques of Bloom's Revised Taxonomy, it may nonetheless be a useful tool for assessing learning standards, especially comparatively. Our results point to differences between and within state and national standards. Recognition of these differences and their implications can be leveraged by future standards writers, curriculum developers, and computing education researchers to craft standards that best meet the needs of all learners.


## 1 Introduction and Background

State and national learning standards play an important role in articulating and standardizing K-12 computer science (CS) education [1]. However, these standards have not been extensively researched, especially in terms of their cognitive complexity [2]. Learning that incorporates more cognitive complexity (i.e., higher-order thinking skills) is crucial for CS students [3]. CS is a key component of engineering education, especially given the quickly growing and highly remunerated positions in CS-related fields [4], including in cybersecurity, robotics, and artificial intelligence.

Thus, the objective of this study is to answer the research question: *How do state and national K-12 computer science standards compare in terms of their cognitive complexity?*

Because there is little prior research on the optimal level and distribution of cognitive complexity in CS learning standards, an exploration of state and national standards is a first step to better understanding the current landscape of cognitive complexity and considering what changes may be warranted in future iterations of learning standards.

One way to assess cognitive complexity is via taxonomies of learning, which are tools commonly used to determine learning objectives and standards, articulate course content, and assess student performance [5]. Bloom's Revised Taxonomy [6] is the most commonly used [5] such taxonomy. This taxonomy uses the following levels, from the least to the greatest cognitive complexity: *remember*, *understand*, *apply*, *analyze*, *evaluate*, and *create*. Bloom's has been aligned with the activities that are specific to learning CS by the ACM Committee for Computing Education in Community Colleges [7]; for example, *code* and *randomize* are part of the *apply* level while *debug* and *optimize* are at the *evaluate* level.

We note that there are substantial and important critiques of Bloom's taxonomy, including its reductionism of learning into quantifiable, hierarchical, atomized content devoid of creativity [8]. However in specific contexts such as this one, Bloom's can nonetheless be a useful tool for assessing learning standards [9]. Thus, while other learning taxonomies have a distinct set of strengths, Bloom's was chosen for this project since it is the most commonly used and has been aligned with distinct CS tasks.

Despite the potential to highlight important aspects – and gaps – within CS learning standards, little prior research has analyzed CS standards in terms of their cognitive complexity. However, one such effort is found in the work of Ardito, who explored the Bloom's levels of the New York CS standards [2] and the CSTA standards [10] via an analysis of frequently used words in the standards, finding that the former tended toward the lower levels of Bloom's but the latter did not. This project extends that work by analyzing the standards of all states and more directly assessing each standard's Bloom's level, as described below.

## 2 Methodology

Our dataset consisted of the standards from states with CS standards ($n = 42$) and the CSTA standards. (This dataset did not include career and technical education standards unless they were the only high school CS standards articulated by the state.) We did not include standards that had language such as "continued growth" or "this standard is not specifically required until . . ."

We determined the level of Bloom's taxonomy for the state ($n = 9695$) and CSTA ($n = 120$) standards. To do this, we assigned the first verb in each standard to a Bloom's level. This assignment was made using the mapping of verbs to Bloom's levels found in *Bloom's for Computing: Enhancing Bloom's Revised Taxonomy with Verbs for Computing Disciplines* [7].

Other common verbs (defined as those that occurred more than 30 times in the state standards) that were not included in this source were assigned the Bloom's level suggested by (1) the context of the standard and/or (2) similar verbs in *Bloom's for Computing*. For example, *seek* commonly occurs in standards in the context of seeking and incorporating feedback, so it was assigned to the *evaluate* level. And *utilize* was assigned the same level as *use* since the meaning of these verbs is similar.

Then, to enable a quantitative comparison of the standards, we mapped numbers to the Bloom's levels, so that 1 = *remember*, 2 = *understand*, 3 = *apply*, 4 = *analyze*, 5 = *evaluate*, and 6 = *create*. This assignment process was accomplished programatically, using a Python script, which is publicly available at tinyurl.com/2ermdxhv. This process resulted in the assignment of a Bloom's level to nearly all standards (the average percent of unincluded standards by state was 6.6%, due to the exclusion of verbs that occurred fewer than 31 times).

For analysis purposes, every state standard was assigned to a grade band (K - 2nd, 3rd - 5th, 6th - 8th, 9th - 12th). This assignment does not always match how the state assigns its standards to grade bands and/or to grade levels. For example, some state standards are assigned to a 9th - 10th grade band or to a 9th grade level; we reassigned those standards to the 9th - 12th grade band to create a uniform assignment system to enable our analysis. We then conducted an exploratory data analysis comparing the state standards (as a whole, and on a state-by-state level) to the CSTA standards.

## 3 Results

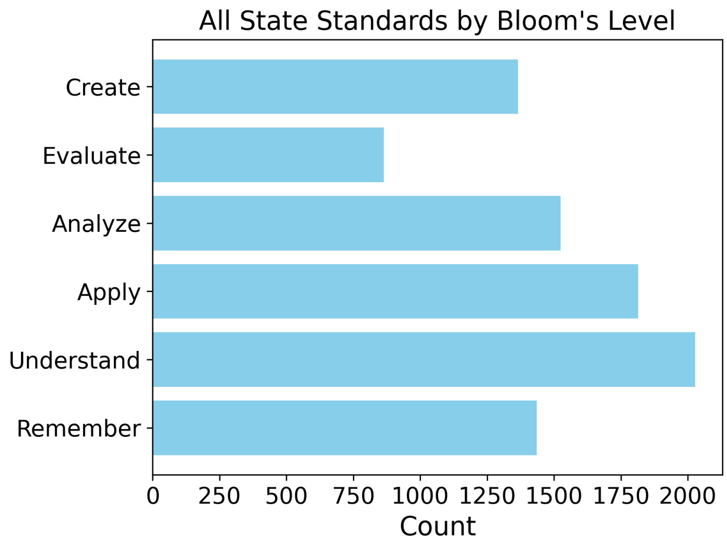


Figure 1: Count of State Standards by Bloom's Level

Figure 1 shows the count of standards for each Bloom's level for the state. The most common Bloom's level is *understand* and the least common level is *evaluate*, with a skew toward the lower levels of the taxonomy.

Figure 2 shows the count of standards for the CSTA standards; there is rough parity in the set of CSTA standards between the five highest levels of Bloom's, with very few standards at the lowest level, *remember*.

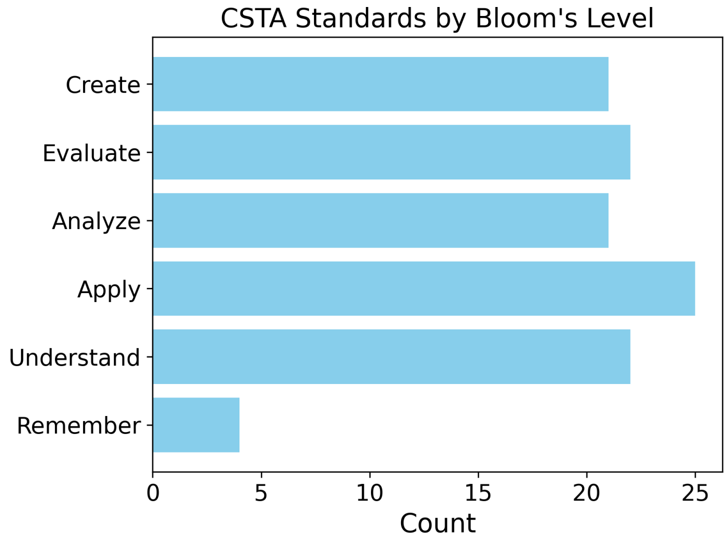


Figure 2: Count of CSTA Standards by Bloom's Level

Thus, the primary difference in cognitive complexity between the state and the CSTA standards is that the CSTA standards largely avoid the lowest level of complexity, while the states have a tendency (albeit more modest) of avoiding the second highest level of complexity. The result is that the CSTA standards have a higher average Bloom's level than the aggregated state standards, with the states averaging 3.27 and CSTA averaging 3.85. We also analyzed the standards by grade band; the results largely mirrored the overall results; see Table 1.

Table 1: Average Bloom's level for state and CSTA standards by grade band. Note that higher numbers refer to higher levels of Bloom's, so that 1 = *remember*, 2 = *understand*, 3 = *apply*, 4 = *analyze*, 5 = *evaluate*, and 6 = *create*.

| **Grade Band** | **State Standards** | **CSTA Standards** |
|---|---|---|
| K - 2nd | 2.7 | 3.3 |
| 3rd - 5th | 3.1 | 3.6 |
| 6th - 8th | 3.3 | 3.8 |
| 9th - 12th | 3.6 | 4.1 |

We then compared the average Bloom's level for each state to the average for the CSTA standards. As shown in Figure 3, the CSTA standards had a higher Bloom's level than all states except for New Jersey (3.90) and Kentucky (3.89). Several states (Iowa, New Hampshire, Connecticut, Washington, Hawaii, New Mexico, and Michigan) have state standards that are identical (or nearly identical) to CSTA's standards and therefore had the same average Bloom's level.

We then analyzed the average Bloom's level for each state and for CSTA by grade band. As Figure 4 shows, there are often differences in the Bloom's level by grade band within a given state, with a common pattern being that the average Bloom's level increases from *apply* in K-2nd to *analyze* in 9th - 12th.

Finally, we grouped the state and CSTA standards by Bloom's level and determined which words

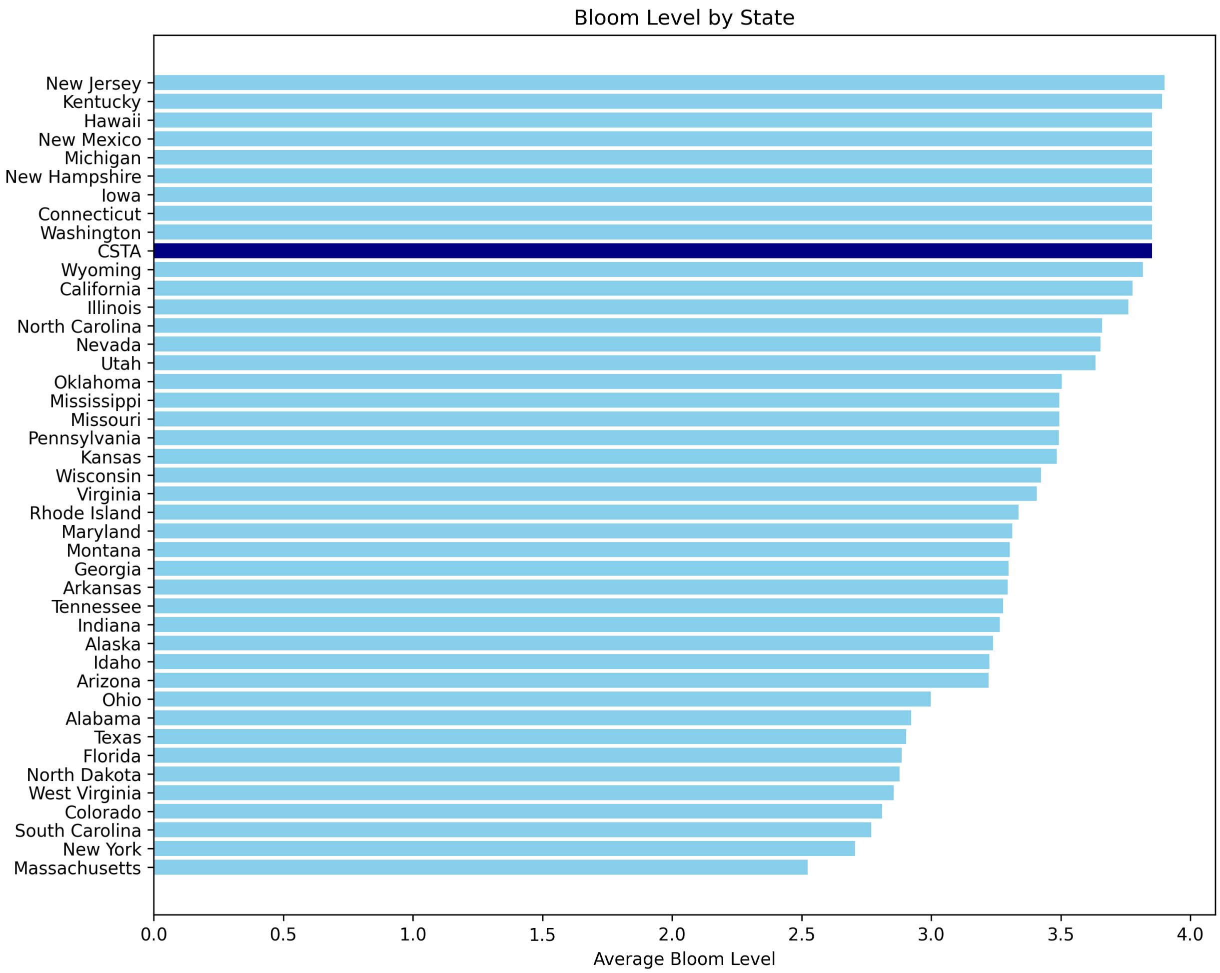


Figure 3: Average Bloom's Level by State

had the highest relative frequency for each level (omitting the previously-analyzed verbs, which would by definition differ by Bloom's level). In general, the most common word lists were largely the same for each Bloom's level, as Table 2 shows. However, some words appeared in two or fewer levels, and these rarer words concentrate at the higher levels of Bloom's.

# 4 Discussion

## 4.1 Overall Cognitive Complexity

To test whether differences in cognitive complexity might be attributable to differences in the topics of various standards, we determined the relative word frequencies by Bloom's level. Our analysis showed that there was substantial overlap in the most frequent words for each level (see Figure 2). For example, *data* was a common term at all levels and *information* appeared on the lists for all four of the lowest levels. Thus, with a few exceptions, it does not appear that differences in topic explain differences in Bloom's level for the set of standards.

Our analysis uncovered some differences in the overall level of cognitive complexity for the state standards (in aggregate) relative to the CSTA standards, with the primary differences being that

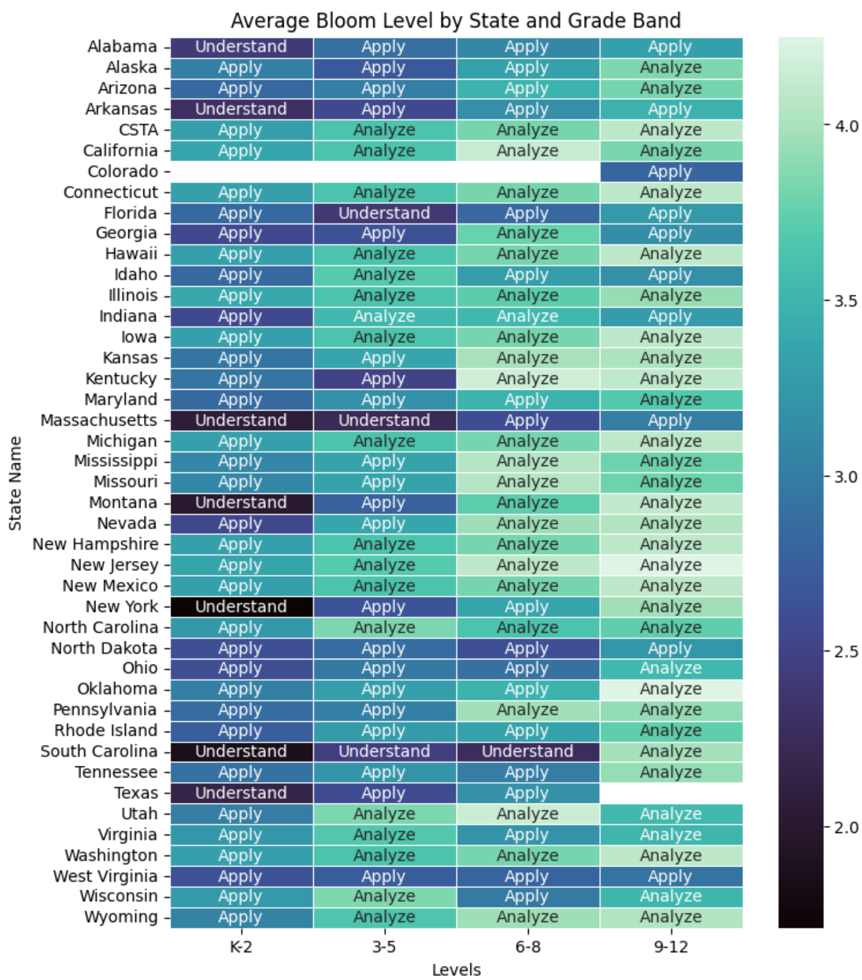


Figure 4: Bloom Level for CSTA and for each State and by Grade Band

Table 2: Most frequent words for each Bloom's level, in descending order (i.e., the most common word is at the top). Words that appear on the lists for two or fewer levels are in bold.

| Remember | Understand | Apply | Analyze | Evaluate | Create |
|---|---|---|---|---|---|
| computing | eg | data | data | computing | using |
| eg | data | using | eg | computational | **programs** |
| data | information | **tools** | problems | data | data |
| devices | computing | eg | software | **artifacts** | computational |
| information | using | **programs** | information | using | eg |
| problems | software | problems | computing | **algorithms** | **algorithms** |
| software | devices | computing | using | **impact** | problems |
| **digital** | **digital** | **development** | **multiple** | eg | **artifacts** |
| **different** | **systems** | information | devices | software | **problem** |
| **hardware** | **computer** | computational | **problem** | **users** | **variables** |

(1) CSTA has very few standards at the *remember* level and (2) the states have the fewest

standards at the *evaluate* level. While there is a strong case to be made that all K-12 students should be given opportunities to develop higher-order thinking skills, it is not necessarily the case that the emphasis on lower-level standards at the state level represents a problem. (Note that there are over 1300 state standards at the highest level, *create*.)

For example, one CSTA standard (in the 6th - 8th grade band) at the *analyze* level is "model the role of protocols in transmitting data across networks and the Internet" [11]. Implicit in this standard is that the student *remembers*, *understands*, and can *apply* prior knowledge of protocols, data, and networks. Presumably, then, the student has already engaged with these lower-order thinking skills before tackling the CSTA standard. Thus, it appears that lower-order thinking skills are *implicit* in the CSTA standards.

The relative lack of state standards at the *evaluate* level presents a somewhat different situation, particularly since some core CS skills – including *debug*, *optimize*, *test*, *validate*, and *verify* – are included in this level [7]. Our analysis did not suggest any rationale for precisely why the *evaluate* level is the least common in the state standards; we believe this would be an interesting avenue for future research that might provide important insight into the process of developing state standards and, potentially, ideas for improving future iterations of learning standards.

### 4.2 Cognitive Complexity by State

Figure 3 shows each state's average Bloom's level and the average level for the CSTA standards. The CSTA standards have a higher average level than most states. This portion of the analysis does suggest that there is a wide variety in the cognitive complexity of state standards as written, ranging from a high of nearly 4.0 (or the *analyze* level) to a low of about 2.5 (the *understand* level). Our analysis did not pinpoint the underlying cause(s) of these differences. Future teams of state standards writers may, therefore, find it useful to assess whether their standards reflect the cognitive complexity level that they deem most appropriate. To date, there is little to no research indicating what average level of cognitive complexity – or distribution of cognitive complexity in a set of standards – is most likely to meet students' needs.

### 4.3 Cognitive Complexity by State and by Grade Band

As shown in Figure 4, there is a range of cognitive complexity by grade band and state. Most states have an average Bloom's level of *apply* for their K-2nd standards and *analyze* for their 9th - 12th standards, showing a gradual increase in cognitive complexity as the grade level increases. Standards writers may want to consider whether their average level of cognitive complexity – as well as whether and how it changes by grade band – reflect best practices for CS education.

### 4.4 Limitations

Our methodological approach was able to classify the vast majority of state standards according to their Bloom's level. However, we did not classify verbs used fewer than 30 times. We also ignored all but the first verb in a standard. We also did not assess career and technical education standards (except in cases where these were a state's only high school CS standards). There are also some limitations related to the use of Bloom's Revised Taxonomy as a proxy for cognitive complexity. For example, an assessment task that is very similar to examples that have previously been provided to students will generally be of a lower cognitive complexity than a task that involves a greater transfer distance.

## 5 Conclusions

We believe this study is one of the first to analyze the cognitive complexity of all state CS standards. This analysis shows important similarities and differences across states, and in comparison to the current CSTA standards. While Bloom's Revised Taxonomy is not a perfect metric, it can provide an indication of the cognitive complexity of a group of learning standards. A better understanding of the cognitive complexity of learning standards will likely be of benefit to standards writers, curriculum developers (especially as they develop learning activities and assessments), CS educators, and CS education researchers. Analyses such as this one can provide insight into the cognitive complexity of standards and suggest directions for future iterations of CS learning standards and related materials.

## 6 Acknowledgments

This project is supported by the National Science Foundation (NSF) under Grant No. 2311746. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF. We thank Vicky Sedgwick for her support in collecting and tagging the state standards and her contributions to our analyses and Monica M. McGill for supportive leadership throughout the process.

## References

[1] Meize Guo and Anne Ottenbreit-Leftwich. Exploring the K-12 computer science curriculum standards in the U.S. In *Proceedings of the 15th Workshop on Primary and Secondary Computing Education*, WiPSCE '20, pages 1–6, New York, NY, USA, October 2020. Association for Computing Machinery. ISBN 978-1-4503-8759-0. doi: 10.1145/3421590.3421594. URL `https://doi.org/10.1145/3421590.3421594`.

[2] Gerald Ardito. A close reading and analysis of the New York state computer science learning standards. *International Journal on Integrating Technology in Education*, 11(1):55–69, 2022. URL `https://www.academia.edu/download/87230040/11122ijite04.pdf`.

[3] Shekhar Kalra, Charles Thevathayan, and Margaret Hamilton. Developing Industry-Relevant Higher Order Thinking Skills in Computing Students. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '20, pages 294–299, New York, NY, USA, June 2020. Association for Computing Machinery. ISBN 978-1-4503-6874-2. doi: 10.1145/3341525.3387381. URL `https://doi.org/10.1145/3341525.3387381`.

[4] Bureau of Labor Statistics. Employment Projections: 2022-2032 Summary - 2022 A01 Results. https://www.bls.gov/news.release/ecopro.nr0.htm, 2023.

[5] Ursula Fuller, Colin G. Johnson, Tuukka Ahoniemi, Diana Cukierman, Isidoro Hernán-Losada, Jana Jackova, Essi Lahtinen, Tracy L. Lewis, Donna McGee Thompson, Charles Riedesel, and Errol Thompson. Developing a computer science-specific learning taxonomy. *ACM SIGCSE Bulletin*, 39(4):152–170, December 2007. ISSN 0097-8418. doi: 10.1145/1345375.1345438. URL `https://dl.acm.org/doi/10.1145/1345375.1345438`.

[6] David R. Krathwohl. A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4):212–218, November 2002. ISSN 0040-5841, 1543-0421. doi: 10.1207/s15430421tip4104$_2.URL$