2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
Paper ID #36798

A Scoping Review of Engineering Textbooks to Quantify the Teaching of Uncertainty

Katelyn Vo, Franklin W. Olin College of Engineering
A.J. Evans, Franklin W. Olin College of Engineering
Shreenithi Madan, Franklin W. Olin College of Engineering
Prof. Zachary Riggins del Rosario, Franklin W. Olin College of Engineering

Zachary del Rosario is an Assistant Professor of Engineering and Applied Statistics at Olin College. His goal is to help scientists and engineers reason under uncertainty. Zach uses a toolkit from data science and uncertainty quantification to address a diverse set of problems, including reliable aircraft design and AI-assisted discovery of novel materials.

A Scoping Review of Engineering Textbooks to Quantify the Teaching of Uncertainty

Abstract

Safe design requires that engineers consider sources of uncertainty. For instance, variability in loading conditions and material properties create a risk of structural failure. However, previous research suggests that engineers, and people more generally, often fail to recognize uncertainty. In this study, we aimed to quantify the degree to which engineering curricula cover topics related to uncertainty. Aligned with the preliminary nature of our research question, we employ scoping review methods.

We operationalized this study by studying the index section of textbooks used in engineering courses. By consolidating library reserve lists from a mix of public and private universities, we constructed a corpus of digitized index sections. Using programmatic tools, we found that the concept of "force" appeared in the corpus ~2.5x as frequently as "uncertainty," 5.6x as frequently as "tolerance," and 7.5x as frequently as "safety factor." This suggests that core ideas related to the practical treatment of uncertainty—and tools to guarantee safety—are considerably under-emphasized in engineering curricula.

Introduction

Safe design requires that engineers consider sources of uncertainty. For instance, variability in loading conditions and material properties create a risk of structural failure [1]. Prior empirical work in statistics education [2–4] and behavioral economics [5,6] has shown that people are highly biased in their treatment of uncertainty. Engineering as a discipline has developed sophisticated tools for identifying and reducing sources of uncertainty; for instance, the tools of statistical process control [7,8]. However, it is not clear how widely these tools are adopted in engineering practice, nor how widely the concepts of uncertainty are taught in engineering programs.

There is reason to believe that uncertainty is not emphasized in engineers' training. Modern engineering curricula heavily emphasize mathematics. For instance, the ABET criteria require 30 credit hours of "college-level mathematics and basic science" and 45 credit hours of engineering topics relevant to the discipline [9]. While mathematics is certainly important to engineering practice, mathematics as a language inherently emphasizes *certainty* [10]. A recent literature review of articles on mathematics within engineering-related disciplines found that only 2 out of 5466 articles discussed "uncertainty" or "error" [11]. While mathematics is fundamental to engineering practice, additional curricular content is necessary to ensure engineers are equipped to handle uncertainties.

Some disciplines require specific training in uncertainty; for instance, the ABET Civil criteria require that curricula train engineers to "apply probability and statistics to address uncertainty" [9]. However, the Aerospace and Mechanical requirements from ABET have no such requirement.

Given the background above, we set out to quantify the consideration of uncertainty in engineering curricula. Our hypothesis is that uncertainty-related topics are not broadly emphasized in engineering curricula. While the review paper of Hadley and Oyetunji [11] quantifies the scholarly discourse, a different approach is necessary to study engineering curricula. We chose to operationalize our study using the index section of textbooks known to be used in engineering classes.

Methods

Aligned with the nature of our research question we elected to perform a *scoping review*, as opposed to a systematic review. Systematic reviews are highly-formalized reviews of existing literature, often including preregistered methods, detailed risk of bias assessment, and critical evaluation of the evidence [12]. These methods are employed to answer highly-specific research questions, particularly when these questions have strong consequences, e.g., to help set evidence-based medical practices [12,13] In contrast, scoping reviews are used for "reconnaissance" [13]—to answer preliminary research questions. In our case, we did not preregister review methods nor evaluate the quality of the artifacts (textbooks) gathered. Given the nature of our research question—to quantify the degree to which concepts are taught, not to evaluate the way they are taught—these aspects of a systematic review did not serve our purposes. While we did not complete a detailed risk of bias assessment, we did design our sampling to ensure corpus alignment with our research question.

Our goal in this study was to quantify the degree to which different engineering concepts are considered important to engineering educators; particularly those concepts related to uncertainty. We chose to study textbooks as a surrogate for this content: Textbooks are peer-reviewed, written by qualified experts, and selected by engineering faculty to teach engineering content. The proper sampling frame for our hypothesis is the set of all textbooks used in teaching courses at ABET-accredited engineering programs. Sampling from this population is challenging, as the set of textbooks in publication is far larger than the set in actual use. Therefore, we limited our sampling frame to a smaller set, in order to guarantee textbook usage.

To ensure we identified only textbooks used to educate engineering students, we used library reserve lists. We obtained these lists from five institutions: Massachusetts Institute of Technology (MIT), University of California Los Angeles (UCLA), State University of New York Polytechnic (SUNY Poly), California Polytechnic State University San Luis Obispo (CalPoly SLO), and Carnegie Mellon University (CMU). These institutions were selected to ensure a mix of public and private institutions, and for the availability of their reserve lists—either by request or through publicly available information. Reserve list policies vary by institution; some institutions list permanent reserves, while other institutions list only those books for the current semester. In our sample, the MIT, CalPoly SLO, UCLA, and SUNY Poly lists are for the Fall 2022 semester only, while the CMU list includes Fall 2022 and permanent reserves.

Table 1. Summary of books identified from course reserve lists of the five institutions.

Institution	Relevant Reserve Books
MIT	23
CalPoly SLO	17
UCLA	17
CMU	7
SUNY Poly	5

From these 5 reserve lists, we identified all textbooks reserved for courses in Mechanical, Civil, and Aerospace engineering, and Material Science books. We chose these three engineering disciplines for their common content of structural design. Materials Science textbooks were included for their relevance to structural mechanics. For some course reserve lists (CMU), departmental codes were not listed. In this case we had to reference the course information to identify relevant books. To check that our assessment was consistent with other lists, we compare the proportion of relevant books to the full course reserve list for two institutions: relevant books from CMU (no departmental codes) included 4.76% of their full list, while relevant books from CalPoly SLO (departmental codes given) included 4.91% of their full list.

Due to the inclusion of identical books across courses & institutions, there are 65 unique books across these five lists. All books except one had an index section, resulting in 64 unique books with an index. For most books we were able to obtain the reserved edition, except for Moran, Shapiro, Boettner, and Bailey [14], for which we used the 6th edition. Table 1 lists the number of relevant books identified from each reserve list, including duplicates.

We attempted artifact recovery for the index section of each textbook, either via the publisher's website or through interlibrary loan (ILL). We sought digitized indexes—PDFs with machine-readable text. Some indexes were only available in non-digitized form: PDF scans of textbook pages with non-machine-readable text. We were able to obtain 45 digitized index sections and 16 indexes in non-digitized form. This left 3 identified books unobtained—an artifact recovery rate of 95.3%. Note that two of the identified books are different editions of

Moran, Shapiro, Boettner, and Bailey; we use the 6th edition as a close content proxy. The full list of identified textbooks, including artifact recovery status, is reported in Appendix A1. The obtained textbooks formed the corpus for this study.

As a surrogate for topics considered important, we chose to study the index section of textbooks used in engineering courses. The index section of a book is more than an outline or a simple concordance; it is a carefully selected list of important terms with locators and cross-references [15]. Terms that appear in the index of a textbook have been deliberately chosen for their relevance to the reader; in our case, engineering students.

Using the digitized artifacts in the corpus, we devised a set of index terms to search across textbooks. These index terms represent fundamental concepts relevant to engineering, such as "force." Furthermore, the terms were organized into one of a few classes; for instance, terms such as "force" and "load" are categorized as *physics* terms, while "error" and "probability" are categorized as *uncertainty* terms. Other terms are categorized as *design*, such as "design", "safety", and "cost".

Terms are categorized as *(mixed)* if their categorization is ambiguous without further manual inspection. The term "limit" may refer to a limit state (in the reliability sense) or a mathematical limit. The term "variation" may refer to variability in a statistical sense, or be part of the phrase "calculus of variations." Finally, the term "error" has widely diverging definitions and interpretations across disciplines, even within engineering [16]. While it would be possible to categorize the use of these terms on an individual-textbook case, this resolution of investigation was outside the scope of the present work.

We detect the presence of selected index terms per-book and use this to compute statistics; we do not provide any additional weight to multiple occurrences. To handle synonyms, we operationalize index terms as regular expressions. For instance, we search for both "safety factor" and "factor of safety." The full list of index terms, along with regular expressions, is given in Appendix A2. After we operationalized the index terms for the digitized artifacts, we performed manual detection on the non-digitized artifacts, recorded these in a digital spreadsheet, and merged these records with the results from the digitized artifacts.

Results

Figure 1 reports the percentage of the digitized corpus that contains the selected search terms. As hypothesized, a minority of engineering textbooks in the corpus list *uncertainty*-related terms in their index: 48% of the corpus mentions "probability," the most-frequently occurring term that is unambiguously *uncertainty*-related. In close second is the term "statistics" (38%).

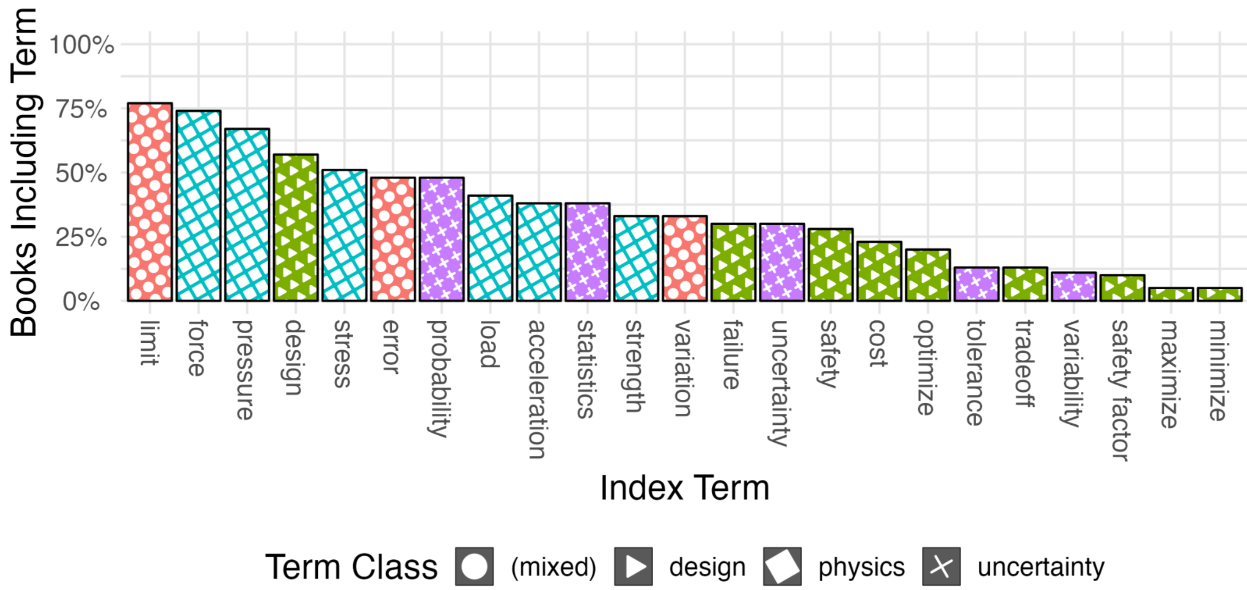

Figure 1. Fraction of books in digitized corpus that include selected search terms. The vertical axis depicts the fraction of books in the full corpus that contain a given index term, while the horizontal axis depicts each term considered in this study. Term categorization is depicted using a fill color and geometric pattern (for greyscale readability).

Other uncertainty-related terms are discussed at a much lower rate, such as "uncertainty" (30%) and "variability" (11%). Surprisingly, even "tolerance" is mentioned at a low rate (13%)—manufacturing tolerances are a key source of uncertainty in engineering. While uncertainty is occasionally discussed in mathematical terms ("probability"), it seems that this concrete source of uncertainty is not widely discussed.

Physics-related terms are broadly considered important; for instance, "force" is included in 74% of the corpus. Other physics-related terms also appear broadly; for instance, "pressure" (67%) and "stress" (51%). The term "acceleration" appears far less frequently (38%); this may be due to a focus on static analysis of structures, though dynamics is considered core to Aerospace engineering [9].

Design-related terms are mixed in their appearance; the term "design" appears in 57% of the corpus, but the next-highest occurring term is "cost" (23%). Note that, despite their ubiquitous use in engineering practice, the "safety factor" concept is mentioned in only 10% of textbooks in the corpus.

It is worth noting that 48% of the corpus mentions “error.” However, the term “error” is highly overloaded. In statistics, the term “error” refers to an “invented,” theoretical object: the difference between the population mean function and observed value [17]. In numerical analysis, the term “error” refers to the difference with some accepted “true” value [18]. In common usage, the term “error” refers to a human mistake, regardless of its nature [19]. Without a deeper inspection of each artifact, it is not possible to determine whether the term “error” refers to a statistical, mathematical, or common meaning.

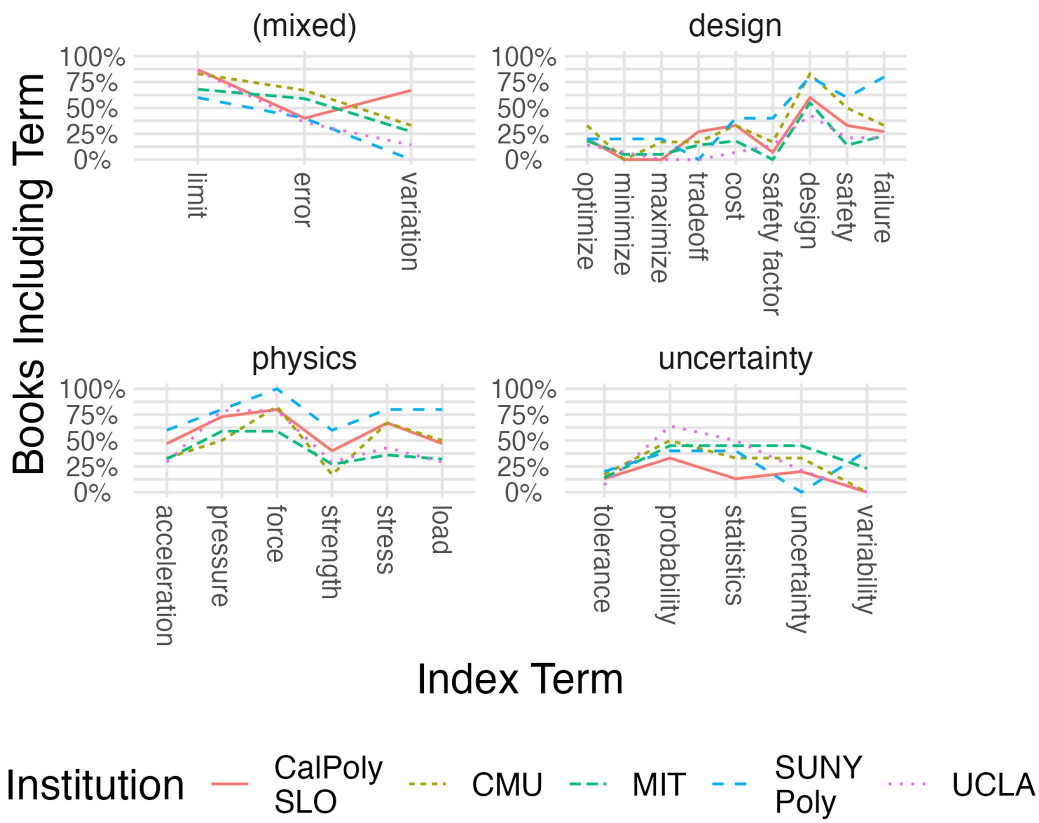

Figure 2. Fraction of books in corpus that include selected search terms, disaggregated by Institution and sorted by similarity of inclusion fraction. The vertical axis displays the fraction of textbooks (per institution) that include the relevant keyword, while the horizontal axis shows each keyword considered. Individual lines correspond to different institutions, depicted by color and linetype (for greyscale readability).

Our corpus also enables a comparative analysis across institutions. Figure 2 disaggregates the term counts by institution (including multiply-represented textbooks). This analysis provides a deeper understanding of curricular contents. For example, the results in Figure 2 illustrate that the course reserve lists at CalPoly SLO discuss *uncertainty*-related terms at a much lower rate than peer institutions. From the same figure, we can see that terms such as “safety” and “failure” are represented at drastically different rates at different institutions.

Finally, the same data can be used to analyze a single collection. This can be helpful for making book recommendations to students, and to help identify individual textbooks to complement the aggregate analyses carried out above. Figure 3 reports the detailed index term inclusion for all textbooks identified from the CalPoly SLO reserve list. This view of the data provides more details on how *uncertainty*-related terms are underrepresented in this collection. Only one book discusses four of these terms (ISBN 9781118651650, *Soil Strength and Slope Stability*, 2nd edition), which may be a useful reference for students interested in learning more about uncertainty.

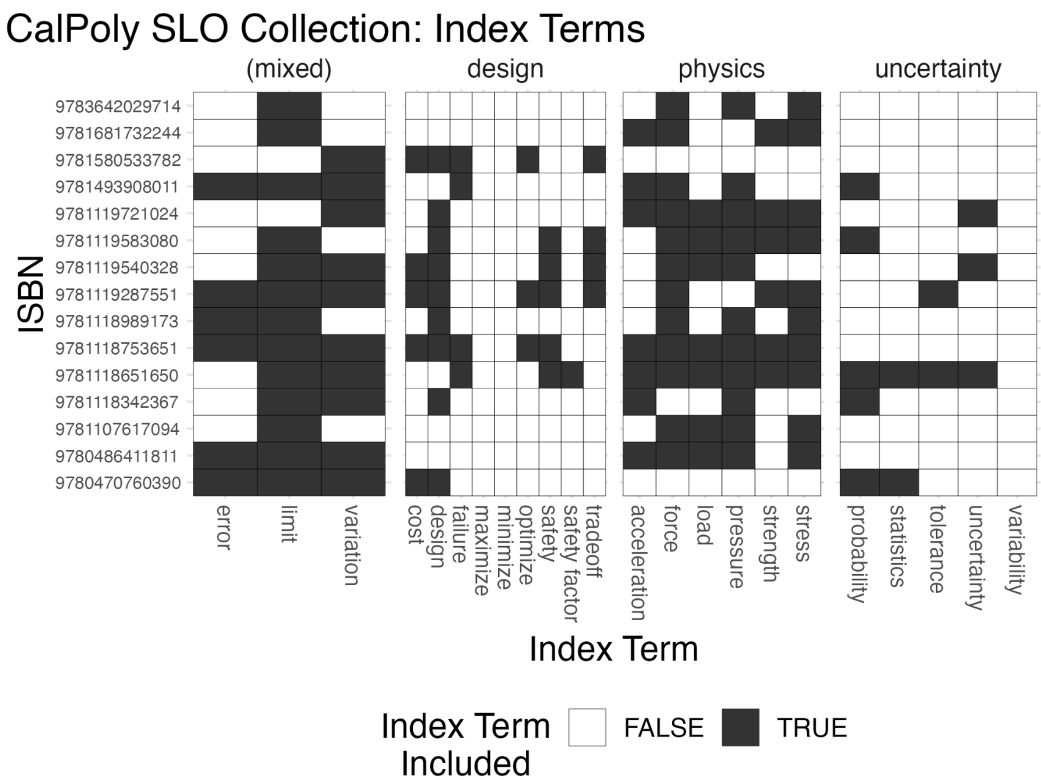

Figure 3. Detailed index term inclusion for all reserve books in our corpus from CalPoly SLO. The vertical axis enumerates all textbooks (by ISBN-13), while the horizontal axis lists all index terms considered. Collections of keywords are separated based on their categorization. A dark color is used to fill cells where the textbook index contains the relevant keyword.

Discussion

We conducted a scoping review of textbooks used to teach Aerospace, Civil, and Mechanical engineering courses. We guaranteed that all books in our corpus are actually used in teaching by consolidating library reserve lists, and operationalized the identification of important concepts by studying the index section of textbooks. We chose keywords that relate to core concepts in engineering physics, design, and uncertainty. By comparing the relative occurrence of these terms, we quantified the relative frequency at which uncertainty concepts are taught in engineering.

The results above support our hypothesis: Concepts related to uncertainty are not broadly represented in our corpus. However, concepts related to engineering physics reliably appear in engineering textbooks. The concept of "force" appears in our corpus ~2.5x as frequently as "uncertainty" and ~2.0x as frequently as "statistics." This suggests that concepts related to uncertainty are not considered core to engineering curricula, and highlights areas that could benefit from additional attention from engineering educators.

The concept of "tolerance" appeared 5.6x less frequently than that of "force," while "safety factor" appeared 7.5x less frequently than "force." Despite their ubiquity in engineering practice, the concepts of "tolerance" and "safety factor" are considerably under-emphasized in engineering textbooks.

Accessing course reserve lists was considerably more difficult for some institutions. Institutions that made this easy simply provided a spreadsheet with textbook (title, authors, ISBN) and course information (course code, department code). Institutions that made this difficult provided no consolidated list, and only allowed access to reserve lists through individual queries. Library staff interested in promoting this kind of research—particularly for internal curriculum development efforts—can provide reserve list information in a consolidated form, such as a spreadsheet. While producing these consolidated lists may be more work, we believe these enable useful analyses (such as the ones sketched above).

Limitations and Future Work

We used course reserve lists as a means to identify textbooks that are used in engineering curricula. This guarantees that our corpus represents actual engineering curricular content. However, this almost surely underrepresents the population of textbooks used in practice. If textbooks are placed on reserve differentially based on their content, this would challenge the external validity of our statistical findings. Future work could study the potential for this bias by studying patterns in library reserve decisions; however, this is outside the scope of the present work.

While our artifact recovery rate was high (95.3%), it was not perfect. We were unable to obtain three books. Of these, two were different editions of Moran et al., which we substituted using the 6th edition. The final unobtained book was a textbook on combustion by Turns (ISBN: 9781260477696). Given the small fraction of the corpus that this single book represents, this omission does not seriously threaten the external validity of our findings.

While reserve lists reflect an important aspect of curricula, textbooks are only an approximation of the curriculum as-taught and as-intended. Textbooks are often only partly aligned with the teaching goals of a particular class; in some cases, there is no textbook that reflects the content of a course. Given this approximate alignment, our results should be considered an approximation of the "true" representation of uncertainty in engineering curricula.

Our methods were aimed—by desire—at educators rather than students, in the sense that we aimed to study the curricular materials that instructors choose to present, rather than the materials that students choose to utilize. Put differently, we can state with confidence that engineering faculty chose to associate textbooks in our corpus with their courses, but this does not reliably indicate that students actually *read* those textbooks. While this does not threaten the validity of our study, it does raise interesting questions for future work: To what extent do the resources *that students seek out* consider uncertainty? Future work aimed at student resource utilization could consider materials beyond textbooks (e.g., open educational resources) by using different methods, such as student and instructor-facing surveys.

Our methods study the presence of certain keywords, but this does not quantify the importance or focus of terms to a particular textbook. For instance, Moran et al. [14] mention "cost," but this is in the context of "cost rate balance for turbine(s)"—hardly a major consideration of this textbook. Future work could further develop the methods used here to provide a finer resolution of topic focus across textbooks.

Reserve lists across institutions are potentially useful beyond the focus of the present study. As noted above, a reserve list reflects faculty decisions; thus, a reserve list reflects important aspects of a given curriculum. We chose to study reserve lists using keywords related to engineering and uncertainty. However, a similar approach could be used to study different aspects of a curriculum; for instance, using a different set of predefined keywords, or even inductively determining keywords from a corpus. Qualitative study of faculty reserve list decision-making may also provide deeper insight into the ways a reserve list does—and does not—reflect curricula. Answers to these questions can provide deeper insight into library collections.

Acknowledgements

We thank the Olin College library staff for their help obtaining many of the artifacts used in this study: Callan Bignoli, Maggie Anderson, and Katy Kania. We also thank three anonymous reviewers whose comments led to improvements in this work. In particular, engaging with their comments helped us improve the readability of this manuscript, better articulate limitations in our study, and identify opportunities for future work. This material is based upon work supported by the National Science Foundation under Grant No. 2138463.

References

(1) del Rosario, Z.; Fenrich, R. W.; Iaccarino, G. When Are Allowables Conservative? *AIAA J.* **2021**, *59* (5), 1760–1772. https://doi.org/10.2514/1.J059578.
(2) Mathews, D.; Pleasant, M.; Clark, J. M. Successful Students' Conceptions of Mean, Standard Deviation, and The Central Limit Theorem. **2007**, 12.
(3) Reading, C.; Pegg, J. Exploring Understanding of Data Reduction. In *Procedings of the Conference of the International Group for the Psychology of Mathematics Education*; International Group for the Psychology of Mathematics Education, 1996; Vol. 4, pp 187--194.
(4) Zieffler, A.; Garfield, J.; Alt, S.; Dupuis, D.; Holleque, K.; Chang, B. What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature. *J. Stat. Educ.* **2008**, *16* (2), 8. https://doi.org/10.1080/10691898.2008.11889566.
(5) Kahneman, D.; Tversky, A. Subjective Probability: A Judgment of Representativeness. *Cognit. Psychol.* **1972**, *3* (3), 430–454.
(6) Konold, C. Informal Conceptions of Probability. *Cogn. Instr.* **1989**, *6* (1), 59–98.
(7) Shewhart, W. A. *Economic Control of Quality of Manufactured Product*; D. Van Nostrand Company, Inc., 1931.
(8) Deming, E. Quality, Productivity, and Competitive Position; Quality Enhancement Seminars, Inc., 1991.
(9) ABET. Criteria for Accrediting Engineering Programs, 2022 – 2023, 2022. https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2022-2023/ (accessed 2022-12-27).
(10) Makar, K.; Rubin, A. A Framework for Thinking about Informal Statistical Inference. *Stat. Educ. Res. J.* **2009**, *8* (1).
(11) Hadley, K.; Oyetunji, W. Extending the Theoretical Framework of Numeracy to Engineers. *J. Eng. Educ.* **2022**, *111* (2), 376–399. https://doi.org/10.1002/jee.20453.
(12) Munn, Z.; Peters, M. D. J.; Stern, C.; Tufanaru, C.; McArthur, A.; Aromataris, E. Systematic Review or Scoping Review? Guidance for Authors When Choosing between a Systematic or Scoping Review Approach. *BMC Med. Res. Methodol.* **2018**, *18* (1), 143. https://doi.org/10.1186/s12874-018-0611-x.
(13) Peters, M. D. J.; Godfrey, C. M.; Khalil, H.; McInerney, P.; Parker, D.; Soares, C. B. Guidance for Conducting Systematic Scoping Reviews. *Int. J. Evid. Based Healthc.* **2015**, *13* (3), 141–146. https://doi.org/10.1097/XEB.0000000000000050.
(14) Moran, M. J.; Shapiro, H. N. *Fundamentals of Engineering Thermodynamics*, 6th ed.; Wiley: Hoboken, NJ, 2008.
(15) Index Evaluation Checklist, 2020. https://www.asindexing.org/about-indexing/index-evaluation-checklist/ (accessed 2022-12-26).
(16) Thunnissen, D. Uncertainty Classification for the Design and Development of Complex Systems. **2003**.
(17) Weisberg, S. *Applied Linear Regression*, 3rd ed.; Wiley series in probability and statistics; Wiley-Interscience: Hoboken, N.J, 2005.
(18) Higham, N. J. *Accuracy and Stability of Numerical Algorithms*, 2nd ed.; Society for Industrial and Applied Mathematics: Philadelphia, 2002.
(19) Error. *Merriam-Webster.com*; 2022.

(20) Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J.; Kuhn, M.; Pedersen, T.; Miller, E.; Bache, S.; Müller, K.; Ooms, J.; Robinson, D.; Seidel, D.; Spinu, V.; Takahashi, K.; Vaughan, D.; Wilke, C.; Woo, K.; Yutani, H. Welcome to the Tidyverse. *J. Open Source Softw.* **2019**, *4* (43), 1686. https://doi.org/10.21105/joss.01686.

Appendices

A1. Full corpus

Table 2 lists all textbooks identified by this study, as well as study metadata (institution course reserve list, whether the book has an index, and artifact recovery status).

Table 2. Full corpus with metadata.

ISBN	Institution	No Index	Recovered
9781118342367	CalPoly	FALSE	TRUE
9781118753651	CalPoly	FALSE	TRUE
9781118753651	CalPoly	FALSE	TRUE
9781107617094	CalPoly	FALSE	TRUE
9781580533782	CalPoly	FALSE	TRUE
9781493908011	CalPoly	FALSE	TRUE
9781119287551	CalPoly	FALSE	TRUE
9781119583080	CalPoly	FALSE	TRUE
9780470760390	CalPoly	FALSE	TRUE
9781118651650	CalPoly	FALSE	TRUE
9781119721437	CalPoly	FALSE	FALSE
9781681732244	CalPoly	FALSE	TRUE
9781119721024	CalPoly	FALSE	TRUE
9781118989173	CalPoly	FALSE	TRUE
9780486411811	CalPoly	FALSE	TRUE
9781119540328	CalPoly	FALSE	TRUE
9783642029714	CalPoly	FALSE	TRUE
9781118989173	CMU	FALSE	TRUE
9780131433564	CMU	FALSE	TRUE
9780521845878	CMU	FALSE	TRUE

9781119723196	CMU	FALSE	TRUE
9780470495902	CMU	FALSE	FALSE
9780199766970	CMU	FALSE	TRUE
9780073529288	CMU	FALSE	TRUE
9780262111621	MIT	FALSE	TRUE
9780521883030	MIT	FALSE	TRUE
9780262016230	MIT	FALSE	TRUE
9781482229561	MIT	FALSE	TRUE
9780262035354	MIT	FALSE	TRUE
9780471862567	MIT	FALSE	TRUE
9780486652429	MIT	FALSE	TRUE
9780486652429	MIT	FALSE	TRUE
9781891389153	MIT	FALSE	TRUE
9780072472271	MIT	FALSE	TRUE
9781118146927	MIT	FALSE	TRUE
9781107617063	MIT	FALSE	TRUE
9780738204536	MIT	FALSE	TRUE
9780471720645	MIT	FALSE	TRUE
9781119186847	MIT	FALSE	TRUE
9781420068610	MIT	FALSE	TRUE
9780300169720	MIT	FALSE	TRUE
9780273016045	MIT	FALSE	TRUE
9780471742999	MIT	FALSE	TRUE
9780471457282	MIT	FALSE	TRUE
9780486837352	MIT	FALSE	TRUE
9781119494966	MIT	FALSE	TRUE
9780073529349	MIT	FALSE	TRUE
9780073398242	SUNY Poly	FALSE	TRUE
9780784415863	SUNY Poly	FALSE	TRUE
9780128150733	SUNY Poly	FALSE	TRUE

9780134441184	SUNY Poly	FALSE	TRUE
9780133840544	SUNY Poly	FALSE	TRUE
9780134859286	UCLA	FALSE	TRUE
9781260471441	UCLA	FALSE	TRUE
9780415413527	UCLA	TRUE	NA
9780521239295	UCLA	FALSE	TRUE
9780471046899	UCLA	FALSE	TRUE
9780521870528	UCLA	FALSE	TRUE
9781260477696	UCLA	FALSE	FALSE
9780805398014	UCLA	FALSE	TRUE
9781461435228	UCLA	FALSE	TRUE
9781119721437	UCLA	FALSE	FALSE
9780486428659	UCLA	FALSE	TRUE
9780691201894	UCLA	FALSE	TRUE
9780716710882	UCLA	FALSE	TRUE
9788131718360	UCLA	FALSE	TRUE
9781498757003	UCLA	FALSE	TRUE
9780132496346	UCLA	FALSE	TRUE
9780898716559	UCLA	FALSE	TRUE

A2. Index terms

The following R code lists all search terms used to parse the index section of textbooks in the corpus. This is a code snippet from the same analysis code used to generate the results above. The code is written using Tidyverse data-wrangling tools, such as the substring detection utility str_detect() [20].

```
```{r define-search-terms}
# Define search terms through helper functions
term_summaries <- list(
  ## Physics
  "acceleration" = ~max(str_detect(.x, "acceleration")),
  "force" = ~max(str_detect(.x, "force")),
  "load" = ~max(str_detect(.x, "load")),
  "pressure" = ~max(str_detect(.x, "pressure")),
```
```

```
  "strength" = ~max(str_detect(.x, "strength")),
  "stress" = ~max(str_detect(.x, "stress")),
  ## Engineering design
  "cost" = ~max(str_detect(.x, "cost")),
  "design" = ~max(str_detect(.x, "design")),
  "failure" = ~max(str_detect(.x, "failure|fail")),
  "maximize" = ~max(str_detect(.x, "maximize|maximization")),
  "minimize" = ~max(str_detect(.x, "minimize|minimization")),
  "optimize" = ~max(str_detect(.x, "optimize|optimization")),
  "safety" = ~max(str_detect(.x, "safety")),
  "safety factor" = ~max(str_detect(.x, "safety factor|factor of safety")),
  "tradeoff" = ~max(str_detect(.x, "tradeoff|trade")),
  ## Uncertainty
  "error" = ~max(str_detect(.x, "error")),
  "probability" = ~max(str_detect(.x, "probability|probabilities")),
  "statistics" = ~max(str_detect(.x, "statistic")),
  "tolerance" = ~max(str_detect(.x, "tolerance")),
  "uncertainty" = ~max(str_detect(.x, "uncertainty|uncertain")),
  "variability" = ~max(str_detect(.x, "variability")),
  ## Mixed / ambiguous
  "variation" = ~max(str_detect(.x, "variation")),
  "limit" = ~max(str_detect(.x, "limit"))
)

```

Note that we selected active terms such as “maximize|maximization” rather than “maximum” to emphasize the process of design, rather than mathematical objects.