Testing Aids and Their Impacts in Engineering Education: An Early Look at a Scoping Review

David P. O'Neill
Department of Biomedical Engineering
Northwestern University
Evanston, IL, USA
david.oneill@northwestern.edu

Abstract— **This research work-in-progress paper describes the initial steps towards a scoping review of knowledge-based testing aids in engineering education. The goal of this scoping review is to map the existing evidence and identify gaps in the limited and disjointed literature on this topic. In engineering education, the use of testing aids is commonplace, with even industry-standard examinations like the Fundamentals of Engineering exam utilizing a dedicated handbook. Despite the prevalent use of testing aids, a lack of consensus exists on best practices in this area. Two reviews exist in the literature, but whilst informative, they fall short by either focusing solely on the outcome of test performance or only comparing open- and closed-note conditions. Further, no literature from the last decade is considered in either review, and both reviews include a range of non-engineering disciplines with a heavy representation of medical education and psychology courses.**

Using an *a priori* protocol based on PRISMA-ScR guidelines, 7 academic databases were searched, including ERIC, Scopus, Web of Science, Engineering Village, and IEEE Xplore. The search yielded 236 journal articles and conference proceedings, of which 93 were identified as duplicates, 65 were excluded during title and abstract review, and 49 were excluded during full-text review, resulting in 29 relevant papers.

This paper addresses 4 research questions: two are purely bibliometric, one explores the geographical distribution of studies, and the fourth examines the terminology used to describe knowledge-based testing aids. Analysis showed that there has been an increase in literature published in recent years (40% were published since 2019), and that publications were evenly split between journal articles and conference proceedings (14 vs. 15). A majority of the papers (72%) explored exams with open-resource use, and most of those (21 out of 26) used the term "open book".

The findings from this paper will help refine a revised review protocol for the full scoping review. The full review will seek to answer two overarching research questions. These are 1) In what course contexts and examination conditions, and for which outcomes have testing aids been studied in the engineering education literature? and 2) What gaps in the engineering education literature exist?

Keywords—scoping review, testing aid, cheat sheet, open-note, open-book, exam

I. Introduction

"Cheat sheets", "reference sheets", and "formula booklets" are amongst many terms used, often interchangeably, to describe knowledge-based tools that students use to aid in their test-taking. Despite origins as forbidden documents, cheat sheets, or "crib sheets" in the British Anglosphere, are mainly considered in their "authorized" (permitted) form. Some are created by individual students themselves whilst others might be supplied by the instructor. Different formats and constraints exist (e.g. 1 page of 8.5x11 paper, double-sided, handwritten) but such specifications are at the discretion of the instructor and, as such, have little consistency from course to course. Open-note and open-book exams are both also of variable definition, with differing expectations of how many books, or which specific texts are permitted.

The use of reference sheets and open-book exams in engineering education dates back to at least the early 20th century. The concept of open-book exams was formally discussed by Thomas in 1922 [1] and in the academic literature in 1951 with Tussing's "A Consideration of the Open Book Examination" [2]. Early experiments on the impact of open-book exams on student performance were conducted by [3], and student opinions of the format were explored by [4].

The use of testing aids saw significant growth in engineering education in the 1980s–90s, paralleling their increased adoption in medical education [5], [6]. Pedagogical rationale for incorporating testing-aids was often that the fundamental skill being tested was not rote memorization but rather the ability to quickly locate and apply relevant information to solve problems [7]. Similarly, unique characteristics of engineering education, where assignments and exams often mimic real-world scenarios in which engineers have access to reference materials and are expected to apply their knowledge to analyze problems and develop solutions [8] provides justification for testing-aid usage in exams. Indeed, the Principals of Engineering exam was open-note until 1993 when the exam become a reference-supplied format with the change to computer-based testing [9].

Some reviews and meta-analyses do exist [10], [11], [12] but none specifically focused on engineering education. Most studies have been small-scale experiments comparing one or two offerings of a single course. Meta-results have been mixed, with some showing improved exam scores with for both student-prepared testing aids, and for open-textbook exams [11] while others found no significant difference compared to traditional closed-book tests [10]. Some have tried to dig deeper into how students use testing aids. Beyond assessment, research topics explored include learning through the process of cheat sheet creation [13], the effects of open-book exams on classroom behavior (e.g. attendance) [14], exam-related anxiety [15], long-term retention and transfer of knowledge [14], and the opinions of both instructors [16], [17], [18] and students [14], [16], [18].

Testing aids, such as cheat sheets and open-note exams, are widely used in engineering programs, but their usage varies considerably. Considering this, along with the distinctive

characteristics specific to engineering education, and the relative lack of conclusive research evidence, further investigation within the engineering education field is warranted.

This project contributes towards addressing the limited and disjointed research on testing aids within engineering education. It characterizes the landscape of existing research on authorized, knowledge-based testing aids in engineering education and explores where gaps exist in the literature. The project seeks to address the following high-level project research questions:

P-RQ1. In what course contexts and examination conditions, and for which outcomes have testing aids been studied in the engineering education literature?

P-RQ2. What gaps in the engineering education literature exist?

As a first step in this project, this work-in-progress paper bibliographically describes the extent of the engineering education literature on authorized, knowledge-based testing aids with the following work-in-progress research questions:

WIP-RQ1. How have publication numbers for this topic changed over time?

WIP-RQ2. Through which publication venues have studies been published?

WIP-RQ3. Where (countries) were studies conducted?

WIP-RQ4. What terminology has been used to describe knowledge-based testing aids?

II. Positionality Statement

The author is instructional faculty in a biomedical engineering department at a mid-sized private R1 in the USA. He primarily teaches undergraduate courses variously with, and without examinations, and has implemented examinations using testing-aid formats across the gamut of common options. For every exam during his own tertiary education, he was provided with a single standardized 160+ page handbook of tables, data, and formulae. The wide variety of approaches to, and opinions of, testing aids amongst his colleagues prompted him to undertake this scoping review project.

III. Methods

This study follows the methodological framework proposed by Arksey & O'Malley [19] and uses an *a priori* protocol based on the PRISMA Extension for Scoping Reviews (PRISMA-ScR) as detailed by [20].

Inclusion criteria for this protocol include: 1. journal articles or conference papers, 2. full text in English, 3. a focus in higher education engineering education, and 4. has at least one research question (explicit or implicit) or hypothesis pertaining to testing aids. The inclusion of conference papers reflects the practice within engineering education scholarship of presenting works-in-progress or full papers at conferences and that these full texts are published [21], [22]. Higher education engineering will be considered to include courses in engineering programs or offered by engineering departments at tertiary-level institutions.

For this WIP, the following databases were queried: ERIC, Scopus, Web of Science, and IEEEXplore. Additionally, Compendex, INSPEC and KNOVEL were jointly queried through the engineering village interface. Space available in this WIP precludes reporting the full search query due to the large number of terms for which separate-word or hyphenated options were both included (e.g. "testing aid" and "testing-aid"). For databases other than ERIC (which was accessed through its API interface) the query included the terms: test aid, testing aid, cheat sheet, crib sheet, crib note, exam aid, support sheet, help sheet, memory aid, reference sheet, formula sheet, equation sheet, open note, open book, all of which were combined with Boolean OR operators. Additionally, the terms "education" and "engineering" were each incorporated with the above using Boolean AND operators. For website searches, all metadata were queried; for ERIC, titles and abstracts were queried.

The search was conducted on 5th April, 2024. Query results were exported, inclusive of abstracts, in formats compatible with the web-based literature review application, Covidence [23].

Covidence; was used to systematically record the process; no algorithmic features (e.g. latent semantic analysis) were used. First, multiplicates were first identified automatically, evaluated manually, and removed as appropriate. Next, titles and abstracts were reviewed in tandem and excluded through a conservatively-inclusive approach; i.e. publications were not removed unless they clearly met one or more exclusion criteria. Example #1: an abstract that makes no mention of the subject for the course studied. Example #2: an abstract focuses on a form of lab-report assignments and only mentions the use of open-book exams in passing; because it is not clear from the abstract if there is any investigation of the open-book exams or it is simply part of the course description, this paper would also not be removed. Next, remaining papers were re-examined with the full-text assessed. Finally, a custom extraction template was used to log data to address this paper's research questions.

IV. Initial Results

The searches returned 236 total results with the following distribution: ERIC (9), Scopus (56), Web of Science (43), Engineering Village combined (105), and IEEE Xplore (23).

A. PRISMA Flow Diagram

The automatic duplicate detection algorithm in Covidence identified 91 duplicates; each duplicate was confirmed manually by the author. The author manually identified two additional duplicates at the abstract-review stage. Of the 143 studies screened by title and abstract, 65 were determined to meet no inclusion criteria, with 78 progressing to full-text review. A further 49 were excluded during full text screening; most exclusory reasons were either that the course(s) investigated were not engineering courses, or that there was no research question relating to using testing aids (e.g. the course description mentioned open-book exams, but there was no assessment of this as an exam format).

B. Bibliographic Data of Included Papers

Work-in-progress research questions 1 & 2 are answered by the extracted bibliographic data. Fig. 2 addresses WIP-RQ1: no relevant engineering education literature was found prior to 1994, ~40% of included articles were published between 2009 and 2017, and ~40% of the included articles were published since 2019

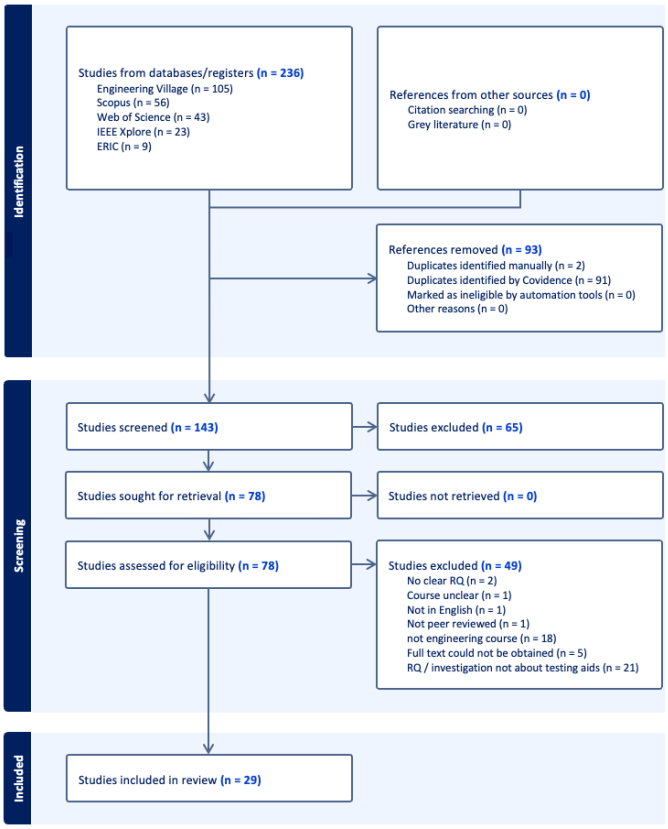

Fig. 1. PRISMA flow diagram for this work. Generated by Covidence [23]

Publication venues (Table 1) were split approximately evenly between conference proceedings and journals. ASEE conference proceedings accounted for nearly half of the final conference papers. The International Journal of Engineering Education published 21% of the final journal articles.

C. Study characteristics

Studies were conducted in 14 different countries. With one exception, studies were conducted in a single country. The most common location was the USA which accounted for 34% of studies. Fig. 3 shows the distribution of all study locations, answering WIP-RQ3.

D. Terminology Used

Terms used in the papers to describe knowledge-based testing aids are shown in Fig. 4. The majority of studies (21/29) explore "open-book" exams. Seven studies used two terms

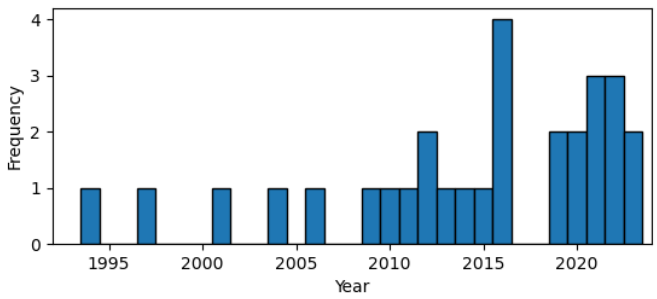

Fig. 2. Number of publications by year

TABLE I. PUBLICATION VENUES OF PAPERS

Publication Venue	Number of papers
ASEE	7
IEEE Global Engineering Education	2
ACM International	1
IEEE Delhi Section Flagship	1
IEEE Frontiers in Education	1
IEEE International conference on Engineering, Technology & Education	1
International Conference on Advance Computing and Innovative Technologies in Engineering	1
Nirma Univeristy International Conference on Engineering	1
Total publications in conference proceedings	**15**
International Journal of Engineering Education	3
Advances in Engineering Education	1
American Institute of Physics	1
Assessment & Evaluation in Higher Education	1
Engineering Education	1
Engineering Science and Education Journal	1
International Journal of Information and Education Technology	1
International Journal of Mechanical Engineering Education	1
Journal of Education and Learning	1
Journal of Materials Education	1
PLoS ONE	1
Sustainability	1
Total publications in journals	**14**

either through exploring two different versions of knowledge-based testing aids or by relating a study-specific term (e.g. "survival cards") to more common terminology (e.g. "open-note").

V. DISCUSSION

The WIP nature of this paper enables community feedback (from reviewers and conference attendees) that refines the planned full-scale scoping review. Methodological revisions could include: inclusory and exclusory criteria, databases searched and search terms, and research questions pursued.

The number of papers identified (236) by the search protocol was smaller than first expected. It is likely that additional papers exist but were not discovered during the search. Reasons might

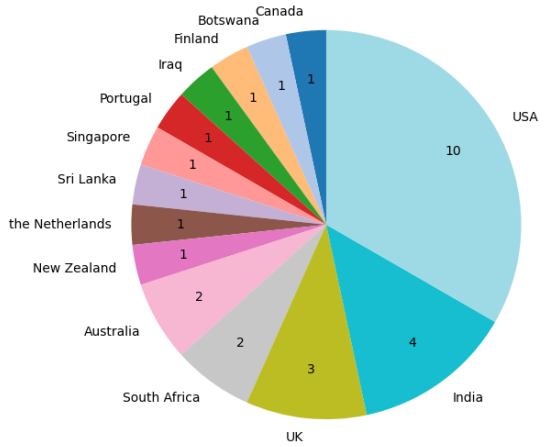

Fig. 3. Country locations of studies. There are 30 locations because one study was partially conducted in Botswana and partially conducted in New Zealand.

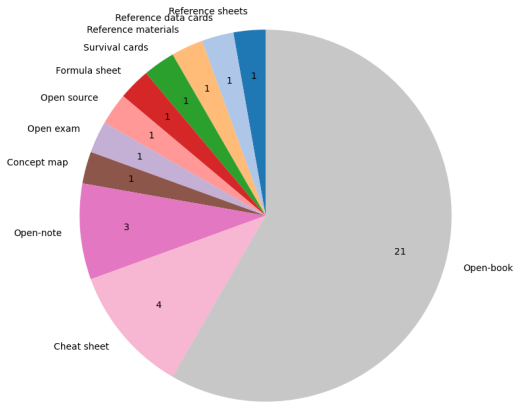

Fig. 4. Terminology used in extracted papers. Seven of the 29 papers used two terms.

include: 1) the breadth of terminology used in the literature, 2) varying definitions of terminology used, 3) additional repositories exist that might uniquely contain literature, 4) not every paper reports the subject for the course; and 5) it can be ambiguous if a course is for engineering students or not. Regarding the breadth of terminology used in the field, during the abstract review stage it was noticed that additional terms exist in the literature that were not used in the search queries, e.g. "mind dump" and "concept map". Notably, no literature was returned by the search protocol explored the use of the NCEES reference handbook for the FE or PE exams. This must be addressed so as to capture such papers as [24], [25] that explore how elements of the supplied-reference format of the fundamentals of engineering (FE) exam can be used in undergraduate engineering education. Unfortunately, this is no trivial task; a post hoc database query including "handbook" as an additional search term increased the number of returned responses by approximately tenfold (Eric: 9→150, Scopus: 56→401, Web of Science: 43→557, Engineering Village: 105→983, IEEE Xplore: 23→69).

A further complication comes from the varying usage of some terminology. There is no agreed-upon definition for each of "open book", "open note", "closed book" or "closed note". "Open book" and "open note" are sometimes used interchangeably, and other times a distinction is drawn where "open book" might permit use of only a single set textbook (as opposed to any and all relevant texts); similarly, "open note" is sometimes defined as permitting any and all relevant texts and other times is restricted to only notes the student has generated themselves during their participation in the course. Examination formats where students can use authorized, self-generated reference sheets (variously referred to as "cheat sheets", "crib sheets", "reference sheets" et cetera) are sometimes considered "open note" exams by researchers, and other times considered "closed note" exams, reflecting the restrictions imposed on the permitted testing aid (e.g. 1 page of 8.5"x11" paper); the discrepancy in terminology usage deriving from the connotation that the term "open" is equivalent to "unrestricted". Whilst Fig. 4 charts the terminology present in the final set of papers, it is not able to represent the variety of exam conditions studied; it likely that there are identical examination conditions and testing-aid usages described with conflicting terminology in the literature. A full scoping review and any systematic review should account for this likelihood in extraction and synthesis protocol design.

Additional literature repositories exist; however, more are searched in this study than in other published engineering-education scoping-review studies [21], [22]. Some gray literature (e.g. theses) likely exists on this topic, however, the expectation within engineering education to publish, and the relatively low threshold for successfully publishing in conference proceedings enabled by the work-in-progress format means that fully-unique studies are unlikely present in the gray literature that are not also represented in the non-gray literature.

The final two points relating to number of papers—that not every paper reports the course subject, and that the precise student population (engineering vs non-engineer) can be ambiguous—were dealt with through a conservatively inclusive practice at the abstract and title review stage. However, at the full-text review stage, inclusion and exclusion criteria were applied strictly. A notable number of papers were removed during the full-text review stage because of ambiguity or lack of detail. There is little that can be done to remedy this because this follows best practice for reviews.

The precise details of inclusion/exclusion criteria for this scoping regarding the definition of "engineering course" are open for community debate. Borderline cases that were classifiable by the protocol but may prompt discourse included: computer science courses (excluded), vector calculus for engineers (student population is engineers, included) [26], 'Materials in Today's World' (taught by an engineering department albeit as a natural science, general education course, included)[27].

The results for WIP-RQ2 (Fig. 2) show that a substantial portion (10/29) of the literature was published since 2020. It is worth noting that the author observed that recent papers focused on the transition to remote-teaching (and remote examining) caused by the Covid-19 pandemic. Many of these papers deal with more drastic examination format alterations beyond investigating the knowledge-based testing aids, and incorporate elements of remote proctoring, asynchronous exams, academic integrity, all in the context of course that have been taught remotely. This was not foreseen and consequently no data extraction relating to remote-examining was included in the *a priori* review protocol. The full scoping review will incorporate this into the data extraction such that this effect can be mapped.

Also not considered during the development of the *a priori* protocol, was the rapid advancement and widespread availability of generative AI tools like ChatGPT which introduces new considerations for open-book exams in engineering education. These AI systems can quickly generate responses to complex questions, potentially altering the landscape of knowledge-based testing aids. An *ad hoc* supplementary search of the literature suggests that much of the work exploring generative AI use in exams focuses on either non-authorized usage, i.e. academic integrity issues, or are benchmarking studies exploring the performance of generative AI with exam assessments; as such, these do not align with the underlying motivation and theme of this review project. The consideration of generative AI within the inclusion parameters

for subsequent stages in this review project remains an open question.

Despite the final number of retained papers (29) being smaller than expected, the Work-In-Progress Research Questions were able to be addressed with the parameters of this scoping review, despite this limited number of studies. However, such paucity of literature may pose challenges in comprehensively addressing the Project Research Questions (see Introduction) to a satisfactory degree.

Further, it seems likely that the engineering education literature overall is too sparse in this area to warrant any form of systematic review at this time. Perusal of the literature (during the full-text review stage) made few, if any, references to the unique characteristics of engineering education as is sometimes argued [8]. However, whilst specific comparisons were rare, some papers assessed both engineering and non-engineering students in their studies [28]. Where systematic reviews focused within the engineering education field may not be possible, it is anticipated that this scoping review project will provide useful data in identifying knowledge gaps for future work in the field.

References

[1] F. W. Thomas, "The meaning, importance, and conditions of study," in *Training for effective study*, Houghton Mifflin, 1922, pp. 17–20.

[2] L. Tussing, "A Consideration of the Open Book Examination," *Educ. Psychol. Meas.*, vol. 11, no. 4–1, pp. 597–602, 1951, doi: 10.1177/001316445101100406.

[3] R. A. Kalish, "An experimental evaluation of the open book examination," *J. Educ. Psychol.*, vol. 49, no. 4, pp. 200–204, 1958, doi: 10.1037/h0043180.

[4] J. F. Feldhusen, "An Evaluation of College Students' Reactions to Open Book Examinations," *Educ. Psychol. Meas.*, vol. 21, no. 3, pp. 637–646, 1961, doi: 10.1177/001316446102100310.

[5] N. Naeraa and G. Huxham, "Criterion-based open book examinations in physiology," *Méd. Educ.*, vol. 14, no. 2, pp. 113–118, 1980, doi:10.1111/j.1365-2923.1980.tb02622.x.

[6] C. C. Tan, "Rethinking medical education.," *Singap. Med. J.*, vol. 39, no. 6, pp. 273–6, 1998.

[7] T. G. F. Gray, "Open book examinations," *The New Academic*, vol. 1, no. 3, pp. 6–9, 1993.

[8] R. M. Felder and R. Brent, "Designing and Teaching Courses to Satisfy the ABET Engineering Criteria," *J. Eng. Educ.*, vol. 92, no. 1, pp. 7–25, 2003, doi: 10.1002/j.2168-9830.2003.tb00734.x.

[9] NCEES (National Council of Examiners for Engineering and Surveying), "The history of the National Council of Examiners for Engineering and surveying: 1920–2004." 2004. [Online]. Available: https://ncees.org/wp-content/uploads/2012/11/The-History-of-NCEES-full.pdf

[10] S. J. Durning *et al.*, "Comparing Open-Book and Closed-Book Examinations," *Acad. Med.*, vol. 91, no. 4, pp. 583–599, 2016, doi: 10.1097/acm.0000000000000977.

[11] K. H. Larwin, J. Gorman, and D. A. Larwin, "Assessing the Impact of Testing Aids on Post-Secondary Student Performance: A Meta-Analytic Investigation," *Educ Psychol Rev*, vol. 25, pp. 429–443, May 2013, doi: 10.1007/s.

[12] B. Johanns, A. Dinkens, and J. Moore, "A systematic review comparing open-book and closed-book examinations: Evaluating effects on development of critical thinking skills," *Nurse Educ. Pr.*, vol. 27, pp. 89–94, 2017, doi: 10.1016/j.nepr.2017.08.018.

[13] K. L. Dickson and J. J. Bauer, "Do Students Learn Course Material during Crib Sheet Construction?," *Teach. Psychol.*, vol. 35, no. 2, pp. 117–120, 2008, doi: 10.1080/00986280801978343.

[14] R. Moore and P. A. Jensen, "Do Open-Book Exams Impede Long-Term Learning in Introductory Biology Courses?," *Journal of College Science Teaching*, vol. 36, no. 7, pp. 46–49, 2007.

[15] C. Theophilides and O. Dionysiou, "The major functions of the open-book examination at the university level: A factor analytic study," *Stud. Educ. Eval.*, vol. 22, no. 2, pp. 157–170, 1996, doi: 10.1016/0191-491x(96)00009-0.

[16] R. Smith and H. Lester, "Instructor and Student Perceptions of the Authorized, Self-prepared Reference Sheet for Examinations," *2019 ASEE Annu. Conf. Expo. Proc.*, doi: 10.18260/1-2--32977.

[17] D. S. Anwer, A. Mamhusseini, Y. B. Omar, and M. S. Othman, "Attitudes of Faculty and Students toward Open-book Examination as a Teaching Strategy in Nursing Education at Hawler Medical University," *Erbil J. Nurs. Midwifery*, vol. 3, no. 2, pp. 136–142, Nov. 2020, doi: 10.15218/ejnm.2020.16.

[18] B. Marcano, B. Ortega-Ruipérez, and A. Castellanos-Sánchez, "Higher education teachers' and students' perceptions of open-book and proctored examinations in the COVID-19 pandemic," *Educ. XX1*, vol. 26, no. 1, pp. 207–228, 2023, doi: 10.5944/educxx1.33514.

[19] H. Arksey and L. O'Malley, "Scoping studies: towards a methodological framework," *Int. J. Soc. Res. Methodol.*, vol. 8, no. 1, pp. 19–32, 2005, doi: 10.1080/1364557032000119616.

[20] A. C. Tricco *et al.*, "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation," *Ann. Intern. Med.*, vol. 169, no. 7, pp. 467–473, 2018, doi: 10.7326/m18-0850.

[21] J. Reed, M. Phillips, A. S. V. Epps, and D. Zwicky, "An Early Look at a Scoping Review of Systematic Review Methodologies in Engineering," *2020 IEEE Front. Educ. Conf. (FIE)*, vol. 00, pp. 1–4, 2020, doi: 10.1109/fie44824.2020.9274232.

[22] M. Phillips, J. B. Reed, D. Zwicky, and A. S. V. Epps, "A scoping review of engineering education systematic reviews," *J. Eng. Educ.*, 2023, doi: 10.1002/jee.20549.

[23] V. H. Innovation, "Covidence systematic review software~." [Online]. Available: www.covidence.org.

[24] A. Miller, "'FE Exam' The First 'Reality Show' Encounter For Engineering And Engineering Technology Undergraduates," in *2006 Annual Conference and Exposition*, 2006.

[25] S. Laughton and D. Nale, "Student Interaction and Perception of FE-Based Formula Sheet Use in Engineering Exams," in *ASEE Southeastern Section Conference Proceedings*, George Mason University in Fairfax, VA, 2023. doi: 10.18260/1-2--45045.

[26] T. S. Craig, F. Kula, and T. Akkaya, "Now what? Pedagogical implications of a shift to open book assessment of vector calculus," IEEE Computer Society, 2022, pp. 1903–1909. doi: 10.1109/educon52537.2022.9766568.

[27] P. Howell, "An overview of an e-education materials science course: Philosophy, design and assessment," pp. 351–363, 2007.

[28] L. Anaya, N. Evangelopoulos, and U. Lawani, "Open Book Vs. Closed Book Testing: An Experimental Comparison," in 2010 Annual Conference & Exposition Proceedings. American Society for Engineering Education, 2010, p. 15.929.1-15.929.11. doi: 10.18260/1-2--16901.