# The Conceptual Fluency Approach for Introductory Thermodynamics

**Catherine Marie Hamel, University of Maryland**

Catherine "Cara" Hamel is a professional track faculty member with the Keystone Program at the University of Maryland. Within this role, Catherine focuses on effectively teaching fundamental engineering courses for first and second year students, teaching Thermodynamics, Introduction to Engineering Design among others. Catherine currently serves as the Course Lead for Thermodynamics, teaching alongside a group of professional-track and tenure-track instructors. Previously, she worked for AcuTech Group, Inc., a specialty consulting firm with a focus in process safety, chemical security and emergency management. Outside of UMD, Catherine can be found at her alma mater Elizabeth Seton High School, where she serves on the Strategic Planning Committee and helps with their blooming engineering program.

**Prof. W. Ethan Eagle, University of Maryland**

Dr. Eagle is a professor of practice in engineering and innovation design. His curriculum design for innovation, co-developed by and building on the research of Jeff and Staney DeGraff, was adopted by the University of Michigan 'Certified Professional Innovator' program in 2014, one of the first such certifications in the country. Now a faculty member in the Keystone Program at the University of Maryland, Dr. Eagle's current work is on the integration of diverse perspectives to discover unique engineering design spaces and on the development of multi-disciplinary courses that bring together students of multiple colleges and/or universities.

©American Society for Engineering Education, 2020

# Preliminary Results of the Conceptual Fluency Approach for Introductory Thermodynamics

## 1. Introduction

This work-in-progress (WIP) paper presents an overview and initial results of a novel assessment method for Thermodynamics that was recently adopted at the University of Maryland, labeled the "Conceptual Fluency Approach." As far as we know, this style of assessment approach has only been tried in the introductory mechanics courses at Michigan State University [1], thus the assessments were adapted to accommodate learning goals of an introductory thermodynamics course, which remained unchanged. This work also builds on other prior work in the area of thermodynamics instruction that has illuminated areas of conceptual difficulty and retention of concepts [2]-[7].

If successful, the Conceptual Fluency Approach achieves greater "fluency" in fundamental thermodynamics concepts and long-term retention of these concepts for future courses, such as heat transfer or fluid mechanics. Part of achieving this fluency is assessing students in a way that discourages students' behavior that seek points ahead of understanding. For example, it was not uncommon to assess a student learning thermodynamics prior to 2019 and see them providing multiple, perhaps conflicting, solution methods on exams in the hopes of receiving partial credit, or observing solution approaches based on 'pattern matching' to a specific problem type on exams without ever clearly demonstrating an understanding of the underlying concepts. Our motivation is rooted in a sense that we must continually adopt new practices to respond to changes in the student population to encourage critical thinking in thermodynamics. Additional motivation and research questions are elucidated in Averill, et al [1] but are beyond the scope of our WIP.

This WIP paper shares the outcomes of this piloted effort in Fall 2019, including how the method was implemented in three sections of Thermodynamics with over 100 total students. We share preliminary data analysis of exam results, faculty impressions of student behavior, and lessons learned from the faculty who implemented. Of considerable importance to the authors is that this change in approach does not result in additional bias [8]-[10]. Thus, we share a preliminary investigation of key metrics in the approach – partial credit requests – broken down by gender and race. We conclude with a summary and discussion of in progress changes made for spring 2020. The goal of this paper is not to prove conclusively the effectiveness of *this* implementation, but rather to share our motivations and implementation thus far, share evidence of what the instructional team observed during this piloted implementation, encourage discussion in the community about our motivations, and compare the levers we all use to elicit positive learning behaviors in our students.

## 2. Implementation of method

This pilot occurred in the Fall 2019 semester across three sections of an introduction to thermodynamics course. A different instructor lectured in each of three sections covering a total of 139 students. Thermodynamics was chosen rather than a freshman engineering introductory mechanics course because our faculty team had a smaller number of instructors, all of whom who

bought-in to making the methodology change. In our 2019 implementation we constructed an assessment method that included four parts, detailed below.

Three 'levels' for Exam questions

In order to attempt to promote conceptual fluency, we assigned three "levels" that students would see on an exam. Each exam included:

- Four Level 1 questions. These assess sense-making (conceptual) and simple (one-concept) problems and required no calculations. Following the language analogy, Level 1 questions are "vocabulary" questions that are the basic concepts of the thermodynamics concepts
- Two Level 2 questions. These assess moderate difficulty problems (single calculation.) Following the language analogy, Level 2 questions are like forming sentences: putting together the words of Level 1 questions
- One-or-Two Level 3 questions. These assess Fluency, the application of multiple concepts in a new way. Following the language analogy, Level 3 shows fluency of the thermodynamics language, able to write and speak about complex topics

A defined grading rubric delivered via online Learning Management System (LMS)

Conceptual Fluency moves away from generous and sometimes laborious partial credit and towards a binary grading technique using a strictly-defined rubric. We award full credit for a fully conceptually and numerically correct solution. For any answer with numeric or conceptual errors, we initially award no points. Defined partial-credit is discussed below.

We also adopted a new LMS to help us quickly operationalize and standardize the approach. It included rubric-based scoring and digitization of student submissions that allowed the grading of 139 exams to be completed and returned to students by TA and faculty in less than 12 hours.

"Flipped struggle" scheme to request partial credit

The LMS includes a feature where students may review their 'released' exams in a digital environment, see instructor comments, and make requests for partial credit on a per question basis. After grading, exam solutions were posted and students had 7 days to review their results and request partial credit. To be eligible, students had to identify the error in their initial solution and submit a statement documenting their understanding as a request through the LMS. For level 2 and 3 problems, we awarded 90% of points to students who reported a conceptually correct solution with one numeric error, and we awarded 85% of points for a conceptually correct solution with two or three numeric errors. Level 1 questions, any solution that demonstrated conceptual errors (including a lack of unit conversion) and solutions with 4 or more numeric errors were not awarded partial credit.

A two-attempt system for examinations.

Finally, to compensate for the additional pressure of the strict defined partial credit rubric and in the hopes of improving retention due to repetitive testing [11] the method allows the students a second attempt at each examination to showcase their understanding of the course content. The implementation utilized a "super-scoring" technique, where the higher of the two attempt scores, by question, was kept and summed together for a final "superscore." An example of how the final scores are calculated is shown in Table 1 below, where a "1" represents a conceptually correct solution to a problem and a "0" represents a conceptually incorrect solution.

Table 1: Mean exam scores.

| Question | Exam A | Exam B | Superscore |
|---|---|---|---|
| 1 | 1 | 1 | 1 |
| 2 | 1 | 0 | 1 |
| 3 | 0 | 1 | 1 |
| 4 | 0 | 0 | 0 |
| 5 | 1 | 0 | 1 |
| Total Score | 3 | 2 | 4 |

## 3. Exam Results

There were three exams, each with two attempts, given during the semester, and one final with two attempts. The results of each attempt, in addition to the "superscored" final score of each exam, is given in Table 2 and Figure 1. The reported scores are the mean score across all three sections, shown with 2 standard deviations of the mean.

Table 2: Mean exam scores and standard deviations after regrade requests.

| | Attempt A | Attempt B | Superscore |
|---|---|---|---|
| Exam 1 | 63 ± 22 | 52± 21 | 84 ± 14 |
| Exam 2 | 40 ± 20 | 56 ± 21 | 74 ± 21 |
| Exam 3 | 52 ± 21 | 60 ± 18 | 81 ± 16 |
| Final | 53 ± 21 | 45 ± 20 | 77 ± 19 |

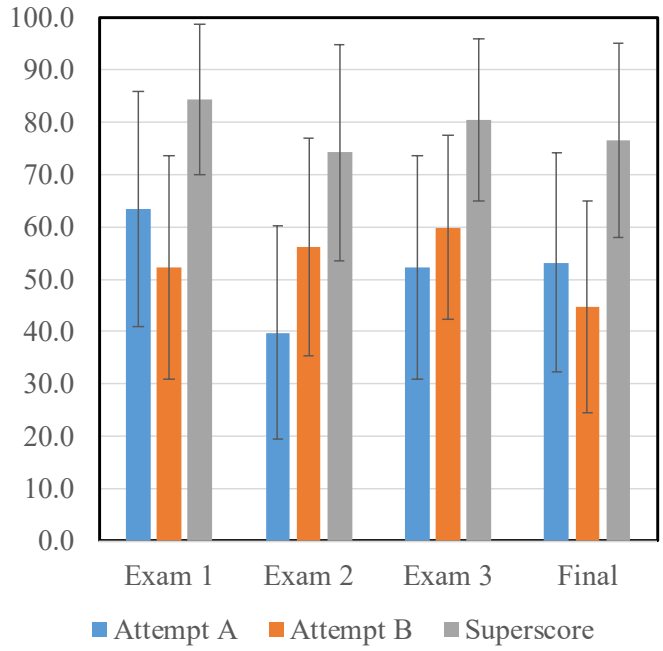


Figure 1: Mean exam scores shown with +/- standard deviations of the mean.

It should be noted that the exam scores shown in Table 2 and Figure 1 are the exam scores *after* points were given back for the numeric errors outlined in the defined partial credit rubric. On average across all exams, the exam mean increased by 4.1 points after these points were given back, with a maximum of 12.6-point increase for Exam 1A and a minimum of 1.2-point increase for Exam 3B.

In addition to exam scores, the number of requests for partial credit were traced and are given in Table 3. It should be noted that students could request partial credit for each question of the exam, so the reported number includes *all* requests rather than the number of students requesting partial credit.

One instructor from Fall of 2018 who made the switch wrote and assessed a problem on the final similar (though not exactly the same) to the previous year. This provides a single reference to compare the results for a subset of students holding 'instruction' nominally constant.

The selected problem was a multi-concept, multi-step Level 3 problem. The results for Fall 2018 and Fall 2019 are shown in Table 4. In Fall 2018, the mean score amongst all parts of the problem ("Total" score) was 51%, whereas in Fall 2019 the mean was 72%. The authors are not comfortable using these nascent results to determine the effectiveness of the method at this time, but look for further data before drawing any conclusions.

Table 3: Number of requests for partial credit by exam.

| Exam | Requests |
|---|---|
| Exam 1A | 127 |
| Exam 1B | 64 |
| Exam 2A | 123 |
| Exam 2B | 120 |
| Exam 3A | 64 |
| Exam 3B | 47 |
| Final A | 84 |

Table 4: Final exam performance in Fall 2018 versus Fall 2019. The "Total" score is taken as the mean of the three parts of the problem.

| | Fall 2018 | Fall 2019 |
|---|---|---|
| Part A (diagram) | 70% | 89% |
| Part B | 59% | 84% |
| Part C | 25% | 43% |
| Total | 51% | 72% |

## 4. Faculty Observation of Students

In addition to examination scores themselves, faculty perceived changes in student behavior throughout the semester. We offer a number of observations but this list is not comprehensive. Also, behaviors are coded either a "positive" or "negative" behavior for student learning based on faculty judgement.

*Positive Behaviors:*

- Students use practice problem sets as a study tool despite there being zero credit given for completion of these problems.
- Attendance of instructor/teaching assistant office hours increased noticeably.
- Students generally adapted quickly to method, after instructor messaging and coaching around the time of the first exam, three weeks into the term.
- Defined partial-credit scheme magnified problem areas not necessarily taught in thermodynamics, such as unit conversions and significant figures (very useful for instructors of the course)

- Anecdotal evidence for Level 1 questions suggest that students who add a sketch, or write an equation to visualize their thinking perform slightly better than students who 'just know' an answer.

*Negative Behaviors:*

- Student skepticism of the method and instructors, which required substantial effort by instructional staff to get students "on board" with methodology.
- Students "gamifying" the superscoring system in order to maximize grade; students either treated the first attempt as a "practice" exam or only attempted a handful of questions on each unique exam attempt (leaving a number of questions completely blank), which led to low and statistically insignificant mean exam scores for a unique exam attempt.
- Students with 'simple' errors will persist without fixing them if partial credit allows.
- Students, on average, did better on Level 3 questions than Level 1, confirming an instructor belief that a student's ability to solve a problem is not necessarily indicative of their understanding of the fundamental concepts driving that problem.

*Role of identity and equity*. One concern of the instructional staff when implementing the Conceptual Fluency Method was the possibility of underlying biases that could disproportionately disadvantage underrepresented minorities. Minority students' lessened sense of belonging [8] and self-efficacy [9]-[10] has shown that these students are more susceptible to losing interest and succumbing to failure [8]. Given that underrepresented minorities, particularly women, are not persisting at the same level as their peers in engineering [8] the Conceptual Fluency's inherent need for self-advocacy and questioning authority (requesting partial credit) and increased frequency of high-stakes events (examinations) may pose an increased threat to underrepresented students.

In an attempt to understand if requesting partial credit results in possible bias, one author reviewed all requests for partial credit to compute the proportion of women and underrepresented minorities submitting these requests. Of 138 students in the course in Fall 2019, 83% were male and 17% female. From a total of 629 unique requests submitted, 87% were by men (548) and 13% by women (81). The author also investigated what percentage of students submitted at least one request for partial credit at any point in the semester: of the 114 men, 95 (83%) submitted a request; of the 24 women in the course, 21 (88%) submitted a request. Therefore, while men submitted a higher number of individual requests across the entire semester (average of 4.8 requests per male student versus 3.4 per female student), a slightly higher percentage of women are requesting partial credit overall[1]. We computed proportions of partial credit requests by underrepresented racial minorities, and those results are shown in Table 5.

The above results are indicative of 'inputs' to our partial credit system. We also wish to also examine 'outputs' - i.e. what % of student requests are valid and result in awarding of credit, and also 'false negatives' - the proportion of students who could have submitted requests and received

[1] Note that due to a low population count, a change in response rate by 1 woman is a 5% change in proportional representation, vs < 1% variation for one more request by men. Meaning if 20/24 women submit requests the percent requesting by gender is an even 83% for both populations. Similar caution should be used when interpreting low population counts and drawing conclusion based on proportion by race. Larger data sets are required to draw strong conclusions.

points, but did not. These analyses were tabled in lieu of space constraint - but are avenues of future inquiry.

Table 5: Percentage of population and percentage of requests for partial credit submitted by underrepresented minorities

| | % of class | % of requests |
|---|---|---|
| Black or African American | 3.6% | 3.8% |
| Asian | 15.9% | 18.4% |
| Native Hawaiian or Other Pacific Islander | 0.7% | 0.0% |
| Hispanic | 5.8% | 4.8% |

## 5. Conclusions and Future Work

This work-in-progress paper summarizes the student performance on Exams (Table 1 and Figure 1), data regarding requests for partial credit (Table 2), and a point by point comparison of a final exam question (Table 4). These preliminary results show that:

- With superscoring, student exam averages exceed 70% for every exam, indicating students achieved conceptual understanding of a majority of concepts
- Partial Credit requests were a vehicle for 'flipped struggle' and there was a higher percentage of requests in the first half of the semester.

With regard to faculty observation of students, we perceived

- A 'Gamified' approach to exam taking due to super scoring
- Students will use practice problems even when no credit is offered.
- Students with 'simple' errors will persist without fixing them if partial credit allows
- Original concerns about this effect of the Conceptual Fluency Method on underrepresented minorities must be further analyzed. Students' self-advocacy with respect to requests for partial credit indicate mixed results. A greater number of requests per from men than women, while a larger % of women submit requests.

The method has continued to be applied in spring 2020, with two new and one persisting instructor. Multiple changes were made to the implementation based on the 2019 results.

- Exams are superscored in their entirety, rather than per problem.
- A small percentage of the final course grade (5%) is awarded for problem sets (homework).
- Fewer quizzes are given in response to student testing exhaustion.
- The number of questions per exam was changed from 9 to 7.
- Additional problem solving time and instructional support was added by hosting an optional Guided Study Sessions (GSS) two times a week, where a TA guides students through relevant problems that are unique from problem set or in-class problems.

The authors hope to continue improving this methodology and are searching for additional methods of assessing the effectiveness of this method with regards to student learning. Longitudinal studies of student success are needed to determine rigorous evidence of effectiveness.

[1] R. Averill, G. Recktenwald and S. Roccabianca, "Effect of Assessment Methods on Performance in Mechanics of Materials," in *Proceedings of the 2018 ASEE Annual Conference & Exposition, Salt Lake City, UT, June 24-27, 2018, Paper ID #23572,* American Society for Engineering Education, 2018.

[2] Turns, S. R. (2012, June), Best Paper PIC I: Applying Knowledge from Educational Psychology and Cognitive Science to a First Course in Thermodynamics Paper presented at 2012 ASEE Annual Conference & Exposition, San Antonio, Texas. https://peer.asee.org/23332

[3] Karimi, A., & Manteufel, R. D. (2012, June), *Assessment of Student Knowledge in an Introductory Thermodynamics Course* Paper presented at 2012 ASEE Annual Conference & Exposition, San Antonio, Texas. https://peer.asee.org/20988

[4] Karimi, A., & Manteufel, R. D. (2014, June), *Assessment of Fundamental Concept in Thermodynamics* Paper presented at 2014 ASEE Annual Conference & Exposition, Indianapolis, Indiana. https://peer.asee.org/20104

[5] Karimi, A., & Manteufel, R. D. (2017, June), *Use of Adaptive Questions and Electronic Pooling to Promote Mastery of Fundamental Thermal Science Concepts* Paper presented at 2017 ASEE Annual Conference & Exposition, Columbus, Ohio. https://peer.asee.org/29064

[6] Wright, K., & Milanovic, I., & Yavuzturk, C. C. (2019, June), *Testing Prerequisite Knowledge of Thermodynamics during a Thermodynamics II Course* Paper presented at 2019 ASEE Annual Conference & Exposition , Tampa, Florida. https://peer.asee.org/33366

[7] J. Clark, J. Thompson, D. Mountcastle. "Investigating Student Conceptual Difficulties in Thermodynamics Across Multiple Disciplines: the First Law and P-V Diagrams," in *Proceedings of the 2014 ASEE Annual Conference & Exposition, Indianapolis, IN, June 15-18, 2014,* American Society for Engineering Education, 2014. https://par.nsf.gov/servlets/purl/10018819

[8] C. Hill, C. Corbett, and A. St. Rose. "Why So Few? Women in Science, Technology, Engineering, and Mathematics," American Association of University Women (AAUW), Washington, DC, Tech. Report, 978-1-879922-40-2, February 2010.

[9] K. Jordan, S. Sorby, and S. Amato-Henderson, "Pilot Intervention to Improve "Sense of Belonging" of Minorities in Engineering," in *Proceedings of the 2012 ASEE Annual Conference & Exposition, San Antonio, TX, June 10-13, 2012,* American Society for Engineering Education, 2012. https://peer.asee.org/21800

[10] C. Vogt, L. Hagedorn, "A Social Cognitive Construct Validation: Determining Women's and Men's Success in Engineering Programs." *The Journal of Higher Education*, Vol. 78, No. 3, pp. 337-364, May – Jun. 2007.

[11] M.A. Pyc and K.A. Rawson, "Why testing improves memory: mediator effectiveness hypothesis." *Science,* 330(6002), 335.