Assessing the given–find–solution method in an undergraduate thermodynamics course

Durward K. Sobek, II (corresponding author), Vic A. Cundy and Vicki L. Briggeman
Mechanical and Industrial Engineering Department, Montana State University, Bozeman, MT 59717-3800, USA
E-mail: dsobek@ie.montana.edu

Abstract This paper describes the given–find–solution method as a general approach to solving engineering problems. An in-classroom study was conducted to provide evidence of its effectiveness in increasing, students' problem-solving proficiency. The target course was an undergraduate service thermodynamics course taught at Montana State University. Results indicate that the given–find–solution method seems to be effective for most students. The implications for assessment and methodological improvements are discussed.

Keywords problem-solving methods, assessment, classroom research

Introduction

Outcomes assessment is receiving prominent attention in engineering programs worldwide [1, 2]. The Accreditation Board for Engineering and Technology (ABET), for example, has made outcomes assessment the cornerstone of its new criteria for US engineering programs seeking accreditation [3]. Outcomes assessment simply means that an engineering program states its outcome objectives, then demonstrates that it is achieving them. It turns accreditation from an input model (we teach engineering mechanics) to more of an output model (our students can solve these classes of engineering mechanics problems, and here are the data to prove it).

ABET criterion 3(e) states that engineering graduates must demonstrate 'an ability to identify, formulate, and solve engineering problems' [3, p. 32]. Also, by criterion 3(k), they must have 'an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice' [3, p. 33]. This strongly suggests that an important component of engineering education is, as it has been for many years, explicit instruction in and practice of problem-solving methods and tools. But how do engineering educators know that the methods and tools they are teaching to students are effective?

Many in the mechanical engineering (ME) faculty at Montana State University teach the given–find–solution method in a variety of ME courses. From experience, those in the faculty have found the method very helpful to students, but have never verified their intuitions experimentally. In other words, they had no data to support their claim that the given–find–solution method is effective. So the authors designed an in-classroom experiment to assess with quantitative measures whether the method is effective in aiding student comprehension and problem-solving ability.

This paper describes the given–find–solution method and presents a theoretical explanation for its effectiveness. It then describes an experiment designed to validate the method and discusses the experimental results in light of the constraints imposed by the classroom. The paper concludes with an in-depth discussion of the implications of such classroom experiments for assessing problem-solving tools, and how the experimental design might be improved.

The given–find–solution approach

The given–find–solution (GFS) method is a semi-structured approach to solving broad classes of engineering problems. The method is used primarily for textbook engineering problems, but it is easily adapted to 'real-world' problems. The beauty of the method is that it provides enough structure to help guide students through problems, but without so much detail that students arrive at a solution following a 'rote' method. It enables true problem solving and learning, not specialized cookbook approaches. Fig. 1 summarizes the approach, explained further in the following paragraphs. Fig. 2 applies the method to a simple heat exchanger problem.

The first step is to list what is 'given' in the problem statement. We require students to create a graphical representation of the system, and label it with the given information. Students are free to draw any diagram they deem appropriate for the problem, but simply copying the problem statement is disallowed. Fig. 2 shows one example of such a diagram, with the heat exchanger depicted and given information labeled as appropriate (e.g., water shown flowing at a certain rate, entering at one temperature and leaving at another). Students must use standard nomenclature for all parameters and use proper units.

The next step is to list the 'find' – what is the problem asking us to do? The student should add this variable to the diagram if possible. In Fig. 2, the problem asks the student to determine the volumetric flow rate ($\dot{V}_1$) of the air.

Textbook problems are usually given in narrative form. More often than not, the problem statement does not include a complete figure.

Students then formulate the problem into:

Given

1. Diagram the system.
2. Label the diagram with given information.

Find

3. List the variable value(s) to be determined.
4. Label 'find' variable on diagram.

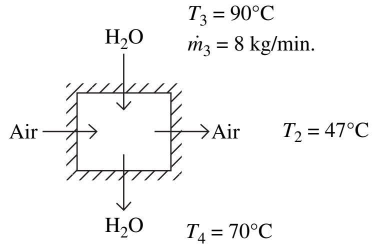
Solution

5. List assumptions (leave space).
6. Write governing equation(s) in general form.
7. Simplify, substitute, and solve.

Fig. 1 *The GFS approach.*

Problem 4-50: In a water-basded heating system, air is heated by passing it through the fins of a radiator. Hot water enters the radiator at 90°C at a rate oof 8 kg/min and leaves at 70°C. Air enrters at 100 kPa and 25°C and leaves at 47°C. Determine the volume flow rate of air at the inlet. (Cengel, 1997, p. 178)

GIVEN:



FIND: $\dot{V}$

SOLUTION:

Assume:
- steady state, steady flow
- air acts as ideal gas
- constant pressure specific heat capacity (c_p) for water

Equations: First law of thermodynamics

$$\dot{Q} + \sum \dot{m}_i\left(h_i + \frac{1}{2}V_i^2 + gz_i\right) = \sum \dot{m}_o\left(h_o + \frac{1}{2}V_o^2 + gz_o\right) + \dot{W}$$

$$\sum \dot{m}_i h_i = \sum \dot{m}_o h_o$$

$$\dot{m}_1 h_1 + \dot{m}_3 h_3 = \dot{m}_2 h_2 + \dot{m}_4 h_4$$

Mass conservation:
$\dot{m}_1 = \dot{m}_2,\ \dot{m}_3 = \dot{m}_4$

$$\dot{m}_1(h_1 - h_2) = \dot{m}_3(h_4 - h_3)$$

$$\dot{m}_1 c_p(T_1 - T_2) = \dot{m}_3(h_4 - h_3)$$

Ideal gas law:
$P\text{v} = RT,\ \Delta\text{h} = c_p\Delta T$

$$\dot{m}_1 = \frac{\dot{m}_3(h_4 - h_3)}{c_p(T_1 - T_2)}$$

$$\dot{m}_1 = \frac{8\,\frac{\text{kg}}{\text{min}}(292.98 - 376.92)\,\frac{\text{kJ}}{\text{Kg}}}{1.005\,\frac{\text{kJ}}{\text{kg·K}}(298 - 320)\text{K}} = 30.37\,\frac{\text{kg}}{\text{min}}$$

$$v_1 = \frac{R_1 T_1}{P_1} = \frac{\left(0.287\,\frac{\text{kPa·m}^3}{\text{Kg·K}}\right)(298\text{K})}{100\text{KPa}} = 0.855\,\frac{\text{m}^3}{\text{Kg}}$$

$$\dot{V}_1 = \dot{m}_1 v_1 = \left(30.37\,\frac{\text{kg}}{\text{min}}\right)\left(0.855\,\frac{\text{m}^3}{\text{kg}}\right) = \underline{\underline{26.0\,\frac{\text{m}^3}{\text{min}}}}$$

Fig. 2 *Example of GFS approach. (Adapted from actual student solution to a problem from Cengel [11, p. 178].)*

Stating the given and find as outlined above initiates the problem-solving thought process. In many engineering problems, what is given and/or what is desired may not be obvious. Often just determining the parameters takes some time, and explicitly identifying them often helps to better determine where difficulties may arise as the problem progresses. In addition, students often think ahead as they formulate the given and find, and consider what assumptions may be required to arrive at a problem solution.

The 'solution' step starts by listing assumptions and then writing the applicable governing equations in general form. The list of assumptions may grow as the problem unfolds, so students are encouraged to leave some space to add assumptions. The equations are then simplified using the assumptions and information from the problem statement. Factors are rearranged algebraically to express the unknown variable(s) (the 'find') in terms of the known variables (the 'givens'). Students are now ready to substitute known (or assumed) values into the equation, and use dimensional analysis to solve for the unknown variable(s).

Theoretical explanation

Why should this method work, aside from experiential evidence? First, models of metacognition – that is, the process of understanding what we are doing and why – indicate that new learners often suffer from not understanding the process by which to think about problems [4, 5]. The result is they do not know where to start formulating a problem strategy. The GFS method gives the student a default starting point, one that starts with problem understanding.

Second, several studies comparing expert and novice problem solvers have noted that experts tend to spend more time understanding the problem than do novices [6, 7]. Novices simply dive into the problem solution without clearly understanding the problem parameters or objectives. The GFS method provides a simple mechanism for synthesizing and organizing the problem information into a coherent, readily understandable framework. Forcing students to restate the problem in a different form gets them to think about it more deeply, understanding what the variables are and how they relate. It also clearly separates given information from what we are trying to solve, a critical step in defining the problem that novice problem solvers often fail to do.

Third, the first author has observed that many students struggle to solve engineering problems because they try to jump directly from problem statement to a final formulation without intermediate steps to study the problem and potential solutions. Prior work [8, 9] hypothesizes that making this jump is difficult because of the significant cognitive gap that exists, and that use of intermediate representations (often embodied in problem-solving approaches) helps one traverse the gap by taking it in smaller steps. The GFS method does precisely that, helping students tackle problems in smaller cognitive steps.

Thus, the GFS method has some theoretical grounding for its usefulness in addition to the experiential evidence. We now turn to the in-classroom experiment to validate the method's effectiveness.

Methods and procedures

The course selected for the study was ME 324: Engineering thermodynamics, taught during the spring 2000 semester. The course targets non-ME junior-level engineering majors taking the course to satisfy requirements of their degree programs. The section taught by the second author had 78 students enrolled at the beginning of the semester. This course and section were selected for their availability, course content (strong problem-solving focus), instructor interest and class size.

We chose a before-and-after format for this study, where student work was evaluated before learning the GFS method, then again after learning the method. Problem-solving proficiency can be measured in terms of quality and time to achieve that quality, so we created measures of both. Student could demonstrate improvement by achieving a higher-quality solution with similar levels of effort, or by achieving the same level of quality in less time.

Data collection and coding

Student homework assignments provided the data for the study. Homework problems were collected weekly throughout the semester. They represented a sample of the assigned homework for the previous week. The students were taught the GFS method approximately 6 weeks into the semester, after the first mid-term examination. For this study, two before-treatment assignments and two after-treatment assignments were collected. The homework sequence was designed to determine the 'before and after' effect of the problem-solving method.

The instructor collected homework assignments on dates specified in the course syllabus. Collected homework represented a subset of the homework problems assigned the previous week. A graduate teaching assistant (not involved with this study) graded the assignments. The students self-reported problem-solving time. They were asked to record the time it took them to solve each problem, and document it on the homework assignment.

Thus, the particular variables of interest for this study were:

(1) *Score*. A numeric value between 0 and 10 (with 10 being a perfect score) to evaluate the quality of the solution, regardless of solution approach.
(2) *Time*. Self-reported time to complete the assigned problem (in minutes) to evaluate the efficiency of problem solving.
(3) *Approach*. 'GFS' or 'alternative method' to evaluate the use or non-use, respectively, of the GFS method.

Solutions set up in the GFS fashion described above were classified as using the method. The specific terms 'given', 'find', and 'solution' did not have to be written down explicitly as long as they were obvious. All solutions that followed a different solution approach were classified as alternative (alt) method. (Note that problems in which the students simply copied the problem statement in the given step were considered alternative method.)

All homework assignments concerned new material covered in class the previous week. Like many engineering courses, the material builds on itself throughout the course. But the challenging parts of the homework assignments concerned the new material, assuming students achieved sufficient mastery of previous material. With this assumption, and the fact that successive homework problems covered new material, we can achieve reasonable comparability between homework problems.

Data analysis

Data analysis involved two steps. First, a group-level analysis used a two-tailed *t*-test (unequal variances) to determine the significance of overall differences in score and time. Outliers for both score and time were identified by visual inspection of histograms and removed from the analysis.

The homework assignments collected before the GFS method was introduced were considered the 'before' sample. The homework assignments after students were taught the method and were familiar with it constituted the 'after' sample. A number of students used the GFS method before it was formally taught in the classroom, having learned it in other classes that use the method or from other students. Thus, the following combinations of variables were used for the *t*-tests on scores and times:

(1) *Overall before versus after*. Comparison of all 'before' solutions with all 'after' solutions.
(2) *Before-alt-method versus after-alt method*. Comparison of problems in which students did not use the GFS method before it was taught with those who did not use it after it was taught.
(3) *Before-GFS versus after-GFS*. Comparison of problems in which students used the GFS method before it was taught with those in which it was used after they were taught it.
(4) *Before-alt-method versus after-GFS*. Comparison of problems in which students did not use the GFS method before it was taught with those in which it was used after they were taught it.

The second analysis step examined individual student results. This analysis compared 'before' and 'after' scores and times of individual students to determine the number of students who improved with the use of the GFS method. The mean scores and times in the 'before' set were compared with the mean scores and times in the 'after' set, and were coded as an increase, a decrease, or remained the same *for each individual student*. An increase or decrease in score and time were defined as a shift of at least ±10%. Then the students' performance was classified as better, same, worse, or indeterminate, as defined in Table 1.

Results

Initial data collection resulted in information (scores, times, and approaches) for four homework assignments: HW1, HW2, HW4, and HW5. HW3 was not used in this study. The first two assignments were completed before the professor taught the GFS method, while the other two assignments were completed afterwards. Thus, HW1 and HW2 were designated as the 'before' set, and HW4 and HW5 were designated

TABLE 1 *Classification of outcomes*

Classification relative performance after introduction of GFS method	
Better	Increase score, decrease time Increase score, same time Same score, decrease time
Same	Same score, same time
Worse	Same score, increase time Decrease score, same time Decrease score, increase time
Indeterminate	Increase score, increase time Decrease score, decrease time

TABLE 2 *Homework assignment summary*

			Score		Time (min)	
	Assignment	Maximum score	Mean	*N**	Mean	*n***
Before	HW1	10	8.43	58	28.88	56
	HW2	10	8.55	60	25.57	56
After	HW4	10	9.29	46	32.23	46
	HW5	10	8.85	53	16.30	53

**N* = number of scores obtained; ***n* = number of times reported.

as the ‘after’ set. Table 2 displays a summary of the homework assignments, scores, times, and sample sizes.

HW1 consisted of three problems graded together to give a maximum of 10 points total. HW2, HW4, and HW5 consisted of one problem each worth 10 points. Therefore, the time figures used for data analysis of HW1 is the average time for one problem (the total self-reported time for HW1 divided by three).

The instructor rated each homework problem for difficulty and complexity. HW2, HW4, and HW5 received comparable ratings (8–9 on a scale of 1 to 10, 10 being very difficult), while HW1 was rated considerably less difficult, but still of moderate difficulty. All the problems were of comparable complexity. No adjustments were made for difficulty or complexity.

Analysis results

Table 3 shows *t*-test results for the group analysis. A positive difference in scores indicates an improvement in solution quality, while an improvement in speed is marked by a negative time difference. Test 1 compares the overall scores and times before and after the method was taught, regardless of the solution method the

TABLE 3 *The t-test results for the group analysis*

		Test 1: Overall before versus after	Test 2: Before-alt-method versus after-alt-method	Test 3: Before-GFS versus after-GFS	Test 4: Before-alt-method versus after-GFS
Score	Mean before	8.68 ($N = 118$)	8.47 ($N = 74$)	9.02 ($N = 44$)	8.47 ($N = 74$)
	Mean after	9.24 ($N = 96$)	9.71 ($N = 7$)	9.20 ($N = 89$)	9.20 ($N = 89$)
	Difference	0.56**	1.24**	0.18	0.73**
Time (min)	Mean before	26.30 ($n = 112$)	25.08 ($n = 71$)	28.41 ($n = 41$)	25.08 ($n = 71$)
	Mean after	20.69 ($n = 96$)	17.86 ($n = 7$)	20.91 ($n = 89$)	20.91 ($n = 89$)
	Difference	–5.62**	–7.23*	–7.50*	–4.17*

*$p \leq 0.05$ **$p \leq 0.01$

students actually used. It shows an overall improvement in both scores and times at the $p \leq 0.01$ significance level. Mean scores increased from 8.68 to 9.24, and speed improved from 26.30 minutes on average to 20.69 minutes. Thus, assuming all else equal, student problem-solving efficacy and efficiency seemed to have improved over the study period.

Tests 2, 3, and 4 looked more closely at different segments within the sample. For test 2, before-alt-method versus after-alt-method, both the score and times improved significantly despite the very small after-alt-method sample size. Closer scrutiny of the after-alt-method solutions, however, strongly suggested that at least one of these students had access to a solutions manual, since the form of solution submitted was nearly identical to that given in the solutions manual. These data must be treated skeptically since the after-alt-method group was likely not to have solved the problem on their own (although we were not able to verify this).

In test 3, before-GFS versus after-GFS, the scores increased from 9.02 to 9.20, but this was not significant; however, the decrease in time was significant at the $p \leq 0.05$ level. The time dropped 7.5 minutes, from 28.41 minutes on average to 20.91 minutes, after the students had received formal instruction in the problem solving method.

Test 4 shows that both the scores and times improved significantly (at the $p \leq 0.01$ and $p \leq 0.05$ levels, respectively) between those students who did not use the GFS method before it was taught (as expected) and those who did use it after it was taught.

The *t*-test results, on the whole, suggest that students who used the more formal GFS method tended to enhance the quality of their solutions and simultaneously improved their problem-solving efficiency (see tests 1 and 4). The method helps students organize and understand the problems, leading to proper formulation and direct application of problem-solving techniques.

For individual student performance evaluation, the average score for HW1 and HW2 was compared with the average score for HW4 and HW5 for each student; the same procedure was followed for time comparisons. If a student completed only one assignment in either the 'before' or the 'after' set, the single measure was used as the mean. Only those students who completed and reported times for at least one homework assignment in the 'before' set and one homework assignment in the 'after' set were included.

Of 52 comparisons, 22 students (42%) performed better after the introduction of the GFS method (i.e., they improved score and/or time without detriment to the other). About a quarter of the sample was indeterminate (students either improved scores but took more time, or spent less time for a worse score). Of the remaining students, about 15% performed the same (scores and times did not change significantly), and just under 20% of the students did worse (i.e., score and/or time was worse without increasing the other). Fig. 3 displays these results.

The individual results seem to show that the GFS approach helped many, though not all, students improve their thermodynamics problem-solving skills. Qualitative analysis of the individual homework solutions indicates that the GFS approach is marginally useful to those students without a basic understanding of the foundational thermodynamics concepts. The indeterminate cases are difficult to assess. Some students may still have been on the learning curve with the new technique and will improve with practice. Others may have given the technique only superficial treatment and so do not constitute a legitimate attempt at the method. We were unable to distinguish between these cases, even after close scrutiny of the students' solutions.

Discussion

Overall, the group analysis results seem to support the hypothesis that the GFS method is helpful in enhancing students' problem-solving proficiency; the individ-

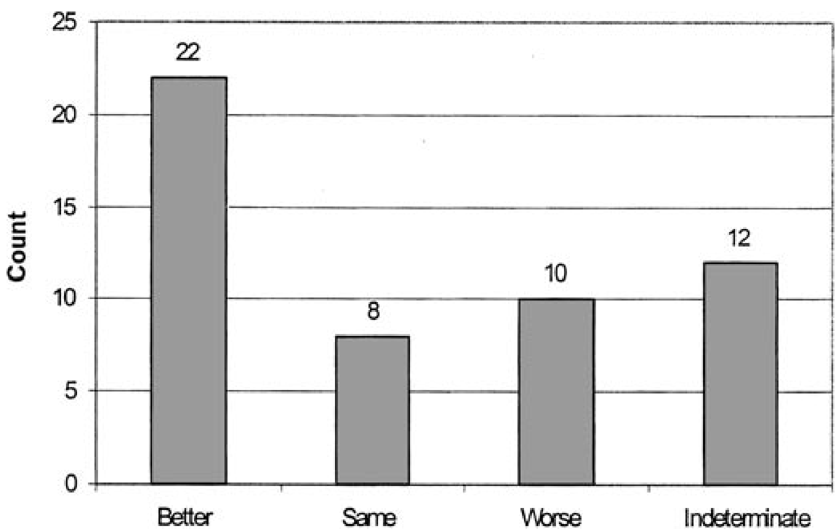


Fig. 3 *Individual analysis results.*

ual analysis results lend mild support. On average, the section performed better on homework assignments after being exposed to the method. Even though over half the class did not improve performance in the individual analysis (see Fig. 3), this could indicate that these students simply need more practice in the method to become proficient. At the very least, adding the 'extra' steps did not seem to add much to the overall solution time. The approximately 20% of students who performed worse with the GFS method may actually represent a population of students for which this structured method was more of a hindrance than a help – a very interesting outcome of the study. But they could also represent students who were not as grounded in prior material as expected.

The results, however, are not entirely conclusive, due to a number of methodological issues. First, and perhaps most prominent, is a comparability issue. Are the different homework problems comparable? We took pains to select comparable problems, considering the students' proficiency level in the course. The instructor rated the homework on difficulty before seeing any study results, and judged that the 'before' set on the whole should actually have been a bit easier for the students than the 'after' set. But these were subjective evaluations, and what the instructor predicts will be challenging for students may in fact be quite easy for them, and vice versa. The reason that students on average scored higher and spent less time on the problems could simply result from the later problems being easier for the students than the earlier problems.

Asking students for their assessment of difficulty is problematic as the GFS method is designed to make problem solving easier, so in fact students may perceive the problem to be easier than they might have otherwise. Or, some students may have found the problem more difficult because they were simultaneously learning the new method. We did not attempt to measure problem difficulty through student assessment.

An obvious solution to the comparability issue is to set up more of a controlled experiment, where a control group solves a problem using any method they want, and an experimental group solves the same problem using the GFS method. But in a classroom setting, this is nearly impossible to achieve (how do you keep the experimental group from 'contaminating' the control group?), and is ethically dubious. Having the two groups in different sections helps mildly, but does not completely alleviate the problem as students in different sections will talk and work together, and the ethics question remains. Separating the two groups by time, such as having the control group in one semester and experimental group in the next, is problematic because one would want to use the same homework problems, and students may share solutions. One way to completely alleviate the issue is to conduct an experiment outside the classroom. But this has its own set of issues, such as funding and time availability.

A second factor to consider is whether learning between homework assignments significantly influenced the scores and times. Students may have gained efficiency in solving problems in the thermodynamics domain, and a number of the weaker students may have dropped the course after the first mid-term examination, causing an artificial increase in overall class performance. However, since each homework assignment covered new material and consisted of different types of problem, the student knowledge level relative to that required for the problem was similar at each stage. Furthermore, the instructor judged the difficulty of the homework problems before seeing the results. HW1 was judged the easiest, while HW2, HW4, and HW5 were of comparable difficulty (though more difficult than HW1).

Third, self-reported times are often an approximation, with varying degrees of accuracy among the students. In this instance, the approximations are considered to be acceptable as the students were asked to record the times as they completed the homework assignments. Only in rare instances were students were asked to recall the amount of time they spent on a particular assignment. Another problem with self-reported times is the tendency for students to inflate the numbers to show the instructor that they are working hard. This is, in theory, not a problem as we are interested in relative differences, so as long as students were consistent in the inflation factor used, the factor should have negligible effect.

Additionally, the overall grading of homework tended to assign scores in the high end of the range. Therefore, the scores may not have been an accurate measure of quality of work for statistical analysis procedures. Also, the spread of scores was perhaps not as large, and the distributions not as symmetrical about the mean, as one would like for statistical analysis. In the future, we recommend stricter scoring criteria to create a more favorable distribution. However, the same grader evaluated all the homework problems in an effort to establish consistent and comparable scoring.

Finally, as with most studies of this nature, subjectivity of the grader and researchers may have introduced a bias in the evaluation and data collection. The grader could have, for example, unwittingly graded more leniently as the semester progressed. This problem was acknowledged and minimized through written criteria for scoring, data collection, and analysis procedures. A more rigorous approach would be to have more than one rater for homework scores. If inter-rater reliability were sufficiently high, we would have greater confidence in the results. Blind scoring by a panel of faculty members on a sample of the homework problems could also help validate study results. But these extensions, of course, would require more resources.

Despite the number of methodological issues, the before-and-after approach has some pedagogical merit. The instructor was able to present the analysis to students, to show them that the method seems to work for at least a good portion of the class, and that it may be even more helpful if they continue to work at it. Thus, the study itself can be used as a motivating tool to show the students the utility of using a rigorous and systematic method to solve problems, and that the diagram sketching and so on are valuable.

Finally, we also compared the examination results with those of another section of the course not participating in the study. This second section was required to use the GFS method throughout the semester, whereas the first section was not required to use it until after the first mid-term examination, approximately six weeks into the course. Both sections took the same examinations at the same time throughout the semester. Interestingly, on average the second section outscored the first section on the mid-term examinations by quite a margin. However, average scores on the following examinations were essentially equal. On its own, this finding may not have much significance (many factors could account for this), but in conjunction with the previous data provides more evidence in support of our hypothesis.

Implications

Studies of this nature are notoriously messy, due to a large number of confounding factors, many of which are impossible or impractical to control. Conducting the experiment in the classroom introduces significant constraints on the experimental design and precludes a highly controlled experiment. Even though the results are not as conclusive as we might like, the data do seem to indicate that the GFS method is a useful tool for many students. So, at the very least we can advocate the GFS method as one approach to solving engineering problems. Whether an instructor should require use of the method is open to debate. One model an instructor could use is to require that students use the method until they have gained sufficient mastery of it. Thereafter the instructor would indicate to students that the approach is likely to be useful to many of them but not all, and that other approaches are acceptable.

Furthermore, the method has broad applicability to many classes of engineering problems. Any engineering problem that gives the student a description of a system and asks for a quantitative solution would be a candidate for this method. The first author has successfully applied the method to production planning problems and to system control problems, while the second author routinely uses the method in heat transfer and heating, ventilation, and air conditioning (HVAC) courses in addition to thermodynamics. Use of the method on small problems can be useful preparation for solving larger, open-ended problems. It gets students in the habit of explicitly asking crucial first questions, like 'What exactly do we know?', 'What do we need to find out?' and 'What assumptions must we make?'

Perhaps a broader implication of this study concerns the utility of such classroom experiments in engineering education assessment. The before-and-after approach can give the instructor a way to gather quantifiable evidence that a tool or method has educational utility, without the ethical considerations associated with conducting a controlled experiment in the classroom (i.e., treating the students as guinea pigs and potentially giving one group an advantage over the other). This kind of experiment can also be done in the classroom easily with the part-time help of an undergraduate.

The results of any given experiment will often not be entirely conclusive, but they do give the instructor some indication as to the usefulness of the method. The methodological issues can be addressed, as noted in the previous section, but not without significant increase in effort, time, and/or cost. Even so, if a number of before-and-after experiments were conducted on different groups of students in different classes working on different problems, and all had results pointing in the same direction, the collective evidence would become quite compelling. A particularly telling variation would be to swap the 'before' and 'after' cases, that is, require the methodology for a time, then not require it. Undoubtedly some students would elect not to use it when not required, and some interesting comparisons could be made.

We are considering a follow-up study involving an in-class problem-solving exercise. Examinations, for example, represent a fairly controlled environment in which we could evaluate alternative approaches, and perhaps learn about innovative approaches that may be more or less effective for different learning styles.

We have also proposed in this study and an earlier one [8] that, when evaluating a problem-solving tool or method, one should consider not just quality of outcome (i.e., is the solution correct?) but also problem-solving efficiency. We have also demonstrated that an easy way to collect these data is for students to self-report their completion times, and that students (at least at MSU) are willing to track and supply this information. In future studies, one might validate the accuracy of the reported times. We could also extend this rubric by assessing other qualities of problem-solving approaches, such as robustness, depth of understanding, or ingenuity.

The pedagogical benefits of the study should not be overlooked. A common complaint among students when we have taught this and other methods is that it is just 'busy work', that is, the required sketching and explicitly listing of the 'find' variable and so on have no value and are done just to satisfy a quirky instructor. The before-and-after experiment then gives the instructor an opportunity to explain to the students why the method is useful, and give the students first-hand experience in its utility. As one last bit of anecdotal evidence, several students confronted the instructor a few weeks into the course wanting to know why he was not teaching (and requiring) the GFS method 'like he had done in other courses'. This event alone was poignant enough for the instructor to conclude that the method was useful. Of course, the instructor told the students not to worry, and that he would soon begin teaching them the method. The exciting thing for the instructor was that he now had a group of students poised and receptive to learning this problem-solving approach.

A final implication is that engineering programs should teach problem-solving methodologies as an explicit part of the curriculum. This paper describes one such method, along with an approach to determining its effectiveness.

Conclusions

This paper has addressed two issues. First, it describes a general problem-solving method and presents data that supports its effectiveness in enabling engineering students to become more able problem-solvers. The results suggest that the GFS method seems to help many (though not all) students improve their problem-solving effectiveness. According to the growing body of work on differing learning styles [e.g. 10], it is unlikely that any one method will benefit all students. So we emphasize that engineering students should be exposed to multiple solution approaches, and be encouraged to find those that work well for their individual learning style. We have described one method that seems to work well for a sizeable group of students.

Second, this paper describes an experiment to validate the efficacy of a problem-solving method, and provides a critique of the experimental method in light of outcomes assessment for engineering programs. We suggest a number of enhancements, augmentations, and follow-on studies to improve experimental validity.

Acknowledgements

This project was funded as a Faculty Fellow mini-grant under the Institutional Reform project at Montana State University, NSF grant #DUE-9850116. Many thanks to Elisabeth Swanson for her help and advice, and many more thanks to the students who willingly participated in the study.

References

[1] P. T. Ewell, 'National trends in assessing student learning', *Journal of Engineering Education*, **87**, (1998), 107–113.

[2] J. McGourty, 'Strategies for developing, implementing, and institutionalizing a comprehensive assessment process of engineering education', *Proceedings of the Frontiers in Education Conference*, (4–7 November 1998), 117–121.

[3] Accreditation Board of Engineering and Technology, Inc. (ABET), *Criteria for Accrediting Engineering Programs*, November 1999 (revised March 2000). see http://www.abet.org/eac/2000.htm.

[4] J. H. Flavell, 'Metacognitive thinking: the struggle for meaning', in L. Resnick (ed.), *The Nature of Intelligence*, (Lawrence Erlbaum, Hillsdale, NJ, 1976).

[5] V. Vadhan and P. Stander, 'Metacognitive ability and test performance among college students', *Journal of Psychology*, **128** (1994), 307–309.

[6] C. J. Atman, J. R. Chimka, K. M. Bursic and H. L. Nachtmann, 'A comparison of freshman and senior engineering design processes', *Design Studies* **20** (1999), 131–152.

[7] M. B. Waldron and K. J. Waldron, 'The influence of the designer's expertise on the design process', in M. Waldron and K. Waldron (eds), *Mechanical Design: Theory and Methodology*, (Springer-Verlag, New York, 1996), 5–20.

[8] D. K. Sobek, II, 'The role of intermediate representations in engineering problem solving', *Proceedings of the ASEE 62nd Annual Pacific Northwest Section Meeting*, (Bozeman, MT, 27–29 April 2000).

[9] D. K. Sobek, II, 'Understanding the importance of intermediate representations in engineering problem-solving', *Proceedings of the 2001 American Society of Engineering Education Annual Conference* (Alburquerque, NM, June 2001).

[10] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development* (Prentice-Hall, Englewood Cliffs, NJ, 1984).

[11] Y. A. Cengel, *Introduction to Thermodynamics and Heat Transfer*, (Irwin McGraw-Hill, Boston, MA, 1997).