PROBLEM SOLVING: THE DIFFERENCE BETWEEN WHAT WE DO AND WHAT WE TELL STUDENTS TO DO

George M. Bodner
Arthur E. Kelly Professor of Chemistry and Education
Purdue University, West Lafayette, IN 47907
Email: gmbodner@purdue.edu

INTRODUCTION

Most of our insight into problem solving has come from research that uses qualitative methods, in which we interview people struggling to solve problems and ask them to talk about what they are doing or what they are thinking while they are involved in this process. Another useful source of information has been the analysis of answers to exam questions coupled with informal discussions with students about why they gave a particular answer when they took the exam. We’ve worked in a variety of courses, from general chemistry through the sophomore organic and inorganic courses, to physical chemistry, and even advanced organic chemistry courses taken by graduate students.

Samples of the kinds of questions we have asked participants in our interviews are shown below. The first question is from an early study of problem solving by science and engineering majors enrolled in a general chemistry course at Purdue. The second is from a study of undergraduates, graduate students, and faculty trying to predict the product of
an inorganic reaction. The third comes from a study of students enrolled in a graduate-level organic chemistry course. The fourth is from a study of graduate students' understanding of aspects of 2D FT NMR.

- Uranium reacts with fluorine to produce a compound which is a gas at 57°C. The density of this gas is 13.0 g/L at 57°C and 1 atm pressure. What is the molecular formula of this compound? (a) UF₂ (b) UF₃ (c) UF₄ (d) UF₅ (e) UF₆

- Predict the products of the following reactions:
  
  \[
  \begin{align*}
  \text{Na} + \text{H}_2\text{O} & \rightarrow \text{NaOH} + \text{Cl}_2 \\
  \text{MgO} + \text{H}_2\text{O} & \rightarrow \text{H}_2\text{S} + \text{Cl}_2 \\
  \text{Ba}_3\text{N}_2 + \text{H}_2\text{O} & \rightarrow \text{NaOH} + \text{SO}_2 \\
  \text{XeF}_4 + \text{D}_2\text{O} & \rightarrow \text{NO}_2 + \text{H}_2\text{O} \\
  \end{align*}
  \]

- Explain the following reaction.
  
  Insert Figure 1 here

- Starting from thermal equilibrium (with M° aligned along the Z axis) and assuming no delays between pulses, predict in which plane the magnetization vector M will lay after experiencing the following pulse sequence. Assume the RF transmitter is aligned along the +X axis. 90°, 90°, 180°, , 90°, , 270°, , 90°, , 90°

PROBLEMS VERSUS EXERCISES

To help the reader understand one of the insights we’ve developed from this work, I’d like you to look at a problem that we have given to perhaps as many as a thousand practicing chemists who were either working in industry or participating in a training program for teaching assistants.

Two trains are stopped on adjacent tracks. The engine of one train is 1000
yards ahead of the engine of the other. The end of the caboose of the first train is 400 yards ahead of the end of the caboose of the other. The first train is three times as long as the second. How long are the trains?

Let's assume that the two trains are headed in the same direction and remind ourselves of the definition of a caboose — the car that used to be placed at the end of a train, which was used by the crew on the train.

Industrial chemists and teaching assistants to whom this problem is given do essentially the same thing. They start with a drawing, in which they use some convention to identify the engine versus the caboose. They typically label the length of one train as “x” and the other as “3x.” They label the distance between the engine of one train and the engine of the other; between the caboose on one train and the caboose on the other.

Insert Figure 2 here

They then write an equation in one unknown, solve for “x”, and report the answer.

\[ 3x + 400 = x + 1000 \]

\[ 2x = 600 \]

\[ x = 300 \]

The only fundamental difference between the two groups is the tendency for those in industry to write “x = 300” and for those in academics to write “x = 300 yd.”

When I tell the industrial chemists that there is no partial credit in this course, and they therefore get a zero, they get mad. They get a zero for the obvious reason — they haven't answered the question! The graduate teaching assistants, when they're told that they get a zero, shrug this off. They're use to not getting the credit they feel they deserve on exams.
For now, let’s focus on two observations about this problem. First, when faced with a novel problem, practicing chemists almost always start with a drawing, of some kind. Second, practicing chemists stop their problem solving activities when they get to the point that they fully understand the problem; not when they get the “answer.”

Now let’s look at another question:

What is the molarity of an acetic acid solution, if 34.57 mL of this solution is needed to neutralize 25.19 mL of 0.1025 M sodium hydroxide.

\[
\text{CH}_3\text{CO}_2\text{H}(aq) + \text{NaOH}(aq) \rightarrow \text{Na}^+(aq) + \text{CH}_3\text{CO}_2^-(aq) + \text{H}_2\text{O}(l)
\]

What would you expect practicing chemists to do? Would they start with an equation or formula, such as: \( n = M \times V \)? Or with a drawing such as the following?

Insert Figure 3 here

The answer should be obvious — in the absence of explicit instruction to do so, no practicing chemist would draw a picture when solving this problem. They would all start by feeding numbers into an equation.

These results suggest that a given individual might exhibit fundamentally different behaviors on different problem-solving tasks. To help the reader understand the source of these differences, we need to define the terms problem and problem solving. We’ll start with John Hayes’ definition of the term problem [1].

Whenever there is a gap between where you are now and where you want to be, and you don't know how to find a way to cross that gap, you have a problem.

According to Hayes, the presence of a gap between where you are and where you want
to be isn’t sufficient to generate a problem. There also has to be an element of uncertainty, confusion, if not downright ignorance about how one is going to cross that gap.

Almost 20 years ago, one of my colleagues in mathematics education (Grayson Wheatley) proposed the only acceptable definition of problem solving I have found. He said, problem solving is “what you do, when you don’t know what to do” [2]. If you’ll accept the two definitions I’ve offered, then you have to accept their logical consequence: There is a fundamental difference between a routine exercise and a novel problem.

Some have said that the difference between an exercise and a problem is one of difficulty; others have said it is one of complexity. I’m going to argue that problems are neither inherently difficult or complex. The only difference between an exercise and a problem is the element of familiarity. Consider the following question from a general chemistry exam.

What weight of oxygen is required to burn 10.0 grams of magnesium?

\[ 2 \text{ Mg(s)} + \text{O}_2(g) \rightarrow 2 \text{ MgO(s)} \]

This is a routine exercise for a practicing chemist, but a novel problem for students who encounter chemistry for the first time. Or consider the following question from a sophomore organic chemistry course.

Robinson annulation reactions involve two steps: Michael addition and aldol condensation. Assume that Michael addition leads to the following intermediate. What would be produced when this intermediate undergoes aldol condensation?

Insert Figure 4 here

This is a problem for most chemists, but a routine exercise for those who teach organic
chemistry.

The distinction between the exercises or problems is important because it is a potential source of miscommunication between instructors and their students. We tend to put a content expert in the classroom, for whom tasks that arise in the course of the semester are routine exercises, and expect that individual to “teach” students for whom the same task is a novel problem.

The difference between the way exercises and problems are worked is particularly well demonstrated by the examples that appear in so many textbooks. These examples have several fundamental characteristics.

- They are logical sequences of steps. (It is rare for a textbook author to be deliberately illogical.)
- They string together in a linear fashion like links on a chain.
- They proceed from the initial information to the solution. (Except, of course, in organic synthesis, where we start from the solution and work back to the initial information.)

These textbook solutions, which are often mirrored by instructors in the classroom, are examples of a phenomenon that has been called “forward-chaining” or “forward-working.” As such, they are examples of how routine exercises are worked by individuals with years of experience with similar tasks [3]. As our work has shown, however, they have little, if any, similarity to the approach successful problem solvers use when they encounter novel problems.

MODELS OF PROBLEM SOLVING

The goal of our work is the development of a model of problem solving that has two
characteristics. First, it must fit our experimental data from interviews with successful problem solvers working on what is, for them, a novel problem. Second, it must be “teachable.” It must be a model that can be given to students that can improve their problem solving performance in chemistry.

Let’s therefore look at several models of problem solving that have been proposed. The first of which is Polya’s stage model [4].

- Understand the problem
- Devise a plan
- Carry out the plan
- Look back

This model makes sense. It seems logical that we would start by understanding the problem, then devising a plan, then carrying out the plan, and then looking back to check our work and consolidate our gains. It is so logical this model can be described as “exocharmic” [5]. It seems to exude charm, particularly for high-school chemistry and physics teachers, who argue that this is exactly what they do.

Unfortunately, Polya’s model is not consistent with our work. To try to convince the reader of this, consider the following problem which is based on the experimental data collected when one of the first xenon fluoride compounds was analyzed [6].

A sample of a compound of xenon and fluorine was confined in a bulb with a pressure of 24 torr. Hydrogen was added to the bulb until the pressure was 96 torr. Passage of an electric spark through the mixture produced Xe and HF. After the HF was removed by reaction with solid KOH, the final pressure of xenon and unreacted hydrogen in the bulb was 48 torr. What is the
empirical formula of the xenon fluoride in the original sample?

We’ve given this problem to many practicing chemists who do not have a history of teaching general chemistry. They inevitably get to the correct answer, but they almost never get there by following the four stages of Polya’s model. Indeed, a common comment heard when they finally get the answer is: “Oh, its an empirical formula problem!” In other words, this problem — like the two trains problem cited earlier — suggests that the process of problem solving is over when one gets to the point that they understand the problem.

Several other models of problem solving that are logical extensions of Polya’s model have been discussed elsewhere [3]. They all have the disadvantage of not being consistent with the patterns we’ve observed for successful problem solvers working on novel problems. Let’s therefore turn to a model proposed by Alex Johnstone and co-workers [7]. This model assumes that each learner has a working-memory capacity (X) and that each problem has a working-memory demand (Z), which is defined as the maximum number of steps activated by the least able individual.

The Johnstone-ElBanna model assumes that when the working-memory capacity of the individual is larger than the demand on working memory (X ≥ Z), we have a necessary, but not sufficient condition, for success. It isn’t sufficient because success depends on prior knowledge; whether the prior knowledge is easily accessible; on the student’s motivation (inclination, interest, etc.); and so on.

This model assumes that students won’t be successful when the demand on working memory exceeds the capacity of working memory (Z > X), unless the student can organize the demand on working memory so that it is smaller than his or her working-memory capacity. Johnstone and co-workers note that there is a sharp drop in performance when
the demand of the problem exceeds capacity. But, some students (\( \approx 10\% \)) seem to be able to solve problems for which the demand exceeds capacity \((Z > X)\) because of chunking devices that reduce the demand on working memory.

Let’s assume that the Johnstone-ElBanna model is correct when it is applied to situations that meet the six criteria proposed by Tsarpalis, et al. [8]. Furthermore, let’s assume that Niaz is correct when he concludes that: “Teachers can facilitate success by decreasing the amount of information required for processing, and thereby avoiding working memory overload” [9]. Now what? From the perspective of this model, there isn’t much we can do to improve student performance in our classes. We simply have to accept the limitations our students bring to the classroom, and conclude that the only way we can improve their performance is to lower the intellectual rigor of the tasks we give them.

I’m going to argue that we can do more than this. Based on research in mathematics education, Grayson Wheatley proposed an anarchistic model of problem solving that describes what successful problem solvers do when they work on novel problems [2]. As noted most recently by Calimsiz [10], this model is consistent with the results of our problem-solving interviews.

An Anarchistic Model of Problem Solving

- Read the problem
- Now read the problem again
- Write down what you hope is the relevant information
- Draw a picture, make a list, or write an equation or formula to help you begin to understand the problem
- Try something
• Try something else
• See where this gets you
• Read the problem again
• Try something else
• See where this gets you
• Test intermediate results to see whether you are making any progress toward an answer
• Read the problem again
• When appropriate, strike your forehead and say, “son of a ...”
• Write down “an” answer (not necessarily “the answer”)
• Test the answer to see if it makes sense
• Start over if you have to, celebrate if you don't

There are several stages in the model that deserve explicit attention. In the “two trains” problem, we saw the role that a drawing that is annotated with relevant information can play in solving a novel problem. We’ve also seen, in the molarity calculation, that drawings aren’t done when we encounter a routine exercise.

I’ve often described the steps “try something” and “try something else” as “playing with the problem.” Unfortunately, all too many of our students believe you can’t “play” with a problem. This is important to me because I have talked to far too many beginning students — particularly those who are struggling with the course — who believe that “trial and error” is not a legitimate strategy for problem solving. That scares me, because it is the most powerful strategy I own.

There is abundant evidence in our data that successful problem solvers routinely
encounter a cue during problem solving that causes them to ask: Am I getting anywhere? Many beginners forget to do this. They exhibit a “garden-path syndrome,” working the problem the way they might walk through a garden — smelling the roses along the way — but not noticing that they aren’t getting anywhere. Successful problem solvers tend to start over when they find that they aren’t making any progress toward the answer; beginners often fail to do this.

The penultimate step in this model is particularly important to me. In 1966, I took a PChem lab in which I was asked to measure the “heat of reaction” in an acid-base neutralization reaction. I did the experiment, worked out the value of $\Delta E$ for the reaction, and reported something on the order of 13,000 kcal/mole.

When the TA’s handed back the lab, they told me my answer was stupid. (They were right; it should have been 13 kcal/mole). They told me that I should have known better. I asked: How? At no point in my academic career had anyone begun to give me the information that would have allowed me to deduce what would have been a reasonable answer for a heat of reaction measured in a calorimeter.

We have found that beginners seldom test their answer to see if it makes sense for two reasons. First, they haven’t ever seen anyone do this when they’ve watched their instructors work out the solutions to tasks that are exercises for the instructors. Second, they have been given the information they would need to do this.

In summary, I would like to argue that Polya’s model is an ideal approach to working a routine exercise. One reads the question, understands the task, devises a plan, and so on. I would like to argue that one of the characteristic tests of whether a task is an exercise is to ask: How is the solution found? Exercises are worked in a linear, forward-
chaining, rational manner. Our model suggests that problem solving is cyclic, reflective, and can appear irrational. Experts who watch students struggle with a problem are tempted to intervene; to show the “correct” way of obtaining the answer. This makes the expert feel good, but it doesn’t necessarily help the individual struggling with the problem for the first time.

**REPRESENTATIONS AND REPRESENTATIONAL SYSTEMS**

Our first hint into the role that representations and representational systems play in problem solving in chemistry came from a study in which we looked at students’ answers to the following question [11].

Predict the product of the following reaction:

\[
\text{PhCOOH} + \text{SOCl}_2 \rightarrow
\]

A couple of typical incorrect answers are given below.

\[
\text{PhCOOH} + \text{SOCl}_2 \rightarrow \text{PhCl} + \text{SO}_2 + \text{HCl}
\]

\[
\text{PhCOOH} + \text{SOCl}_2 \rightarrow \text{PhCOOCl} + \text{SO}_2 + \text{HCl}
\]

When people who teach general chemistry look at these equations, they often comment: They aren’t balanced. That doesn’t bother me because organic chemists seldom worry about the mundane details of writing balanced equations. What bothers me is the fact that this equations are absurd. There is no way to go from the starting materials to the products of these equations by making and breaking of bonds.

When we did this experiment, we noted that some of the students who answered this question successfully wrote a symbolic representation in the margin of the exam paper. It seldom looked as regular as the symbolic representation one would fine in a journal:

Insert Figure 5 here
Sometimes the ring looked as if it had six “bumps” corresponding to the six carbons of a benzene ring, often it did not. The ring sometimes contained one double bond, sometimes two, sometimes three. Sometimes it didn’t contain any double bonds. But the -CO₂H portion was always clearly written.

Some would argue that the “PhCO₂H” with which the starting material was presented is a “symbolic” representation. I’d like to argue that it can be a symbolic representation, but it often is not. For many students, particularly those who struggle with organic chemistry, it is a verbal/linguistic representation that consists of letters and numbers that aren’t symbols because they don’t symbolize anything.

Interviews with students struggling with organic chemistry has lead us to conclude that there is a fundamental difference between what the instructor writes on the blackboard and what students write in their notebooks, in spite of the fact that one seems to be a direct copy of the other. Similar external representations are written by individuals with different internal representations. What students write in their notebooks seems to be a direct copy of what the instructor writes on the blackboard. In spite of this, there is a fundamental difference between what the instructor and some students write. The instructor writes symbols, which represent a physical reality. All too often, students write letters and numbers and lines, which aren’t symbols because they have no physical meaning to them. Interviews suggest that it is the students who are trapped in verbal/linguistic representation systems who are most likely to write the following equation for attack by a Grignard reagent on a ketone.

Insert Figure 6 here

It isn’t until the letters, lines and numbers in this equation become symbols that this answer
becomes wrong. One can draw as many lines to a “C” as one wants; it is only bonds to a carbon atom that are limited to four.

In an earlier paper in *University Chemistry Education* we presented the results obtained when we analyzed student answers to an examination question in which they were asked to predict the products of the photochemical bromination of methylcyclopentane [12]. We used this example as the basis for arguing that successful problem solvers construct significantly more representations while solving a problem than those who aren’t successful, even though neither group constructed very many representations while solving the problems.

This time, let’s look at another example of the power of a second representation at getting to the answer to a novel problem. Several years ago, one of my physical organic chemistry colleagues put the following question on a homework assignment for a graduate-level organic course.

*Explain why electrophilic aromatic substitution occurs at the indicated position.*

*Insert Figure 7 here*

As “everyone” knows, amines are notoriously good ortho/para directors. Thus, if one follows the arguments in virtually any introductory organic textbook, electrophilic attack should occur at C(4). Experiment, however, suggests that electrophilic aromatic substitution occurs at C(3).

I was introduced to this “problem” because one of the 15 students in the class was in my research group. He noted that he had no idea how to answer this question, and asked me if I knew the answer. I admitted that I didn’t. I was in good company, however,
because none of the students knew how to answer the question.

When I talked with the instructor, he was surprised. He said the answer is obvious. You just need to look at the molecule from a different perspective. What I would call using a second representation.

Insert Figure 8 here

When this is done, the answer becomes obvious. The steric effect of the methyl groups on C(2) and C(6) keeps the N, N-dimethyl amine substituent from lying in the same plane as the aromatic ring. This turns off the resonance effect of the amine, with the net effect that attack at C(3) becomes favorable. Regardless of what system I use to make my point, our work has repeatedly shown that people who use more than one representation to solve a problem are more likely to get the correct answer than those who struggle with a single representation.

I’m not arguing that it is impossible to get to the answer from the first representation one builds of the problem. There is abundant evidence that this can happen, particularly for those for whom the task is an exercise.

I’m not asking for many representations to be applied to a problem. Our work suggests that all one typically needs is a second representation.

I’m not arguing that a second representation will always get you to the answer to a problem because the first representation (when incorrect) can lead to a second incorrect representation, which leads to an erroneous answer. I’m just arguing that bringing two representations to a problem increases the probability of getting the right answer.

As a test of this hypothesis, consider the following question:

Which weighs more, dry air at 25°C and 1 atm, or air at this temperature
and pressure that is saturated with water vapor? (The average molecular weight of air is 29.0 g/mol.)

When they first see this question, some claim that it isn’t fair. They’re right; it isn’t. I don’t tell you the volume of either the sample of dry air or the sample of air saturated with water vapor. That’s deliberate because it is only in general chemistry (and perhaps organic chemistry) that one encounters tasks that have all the information needed to solve them, and only the information needed to solve them. In the real world, tasks come to us that are poorly defined. The only way to answer this question is to make a reasonable assumption: That we are talking about the same volume of the two gases. I.e., one liter of dry air versus one liter of air saturated with water vapor.

A lot of very bright chemists, when they encounter this question for the first time, conclude that the air saturated with water vapor has to be heavier than the dry air. They are in good company; the vast majority of beginning students reach the same conclusion. The source of the error in their thinking is simple: They build a verbal/linguistic representation of the question when they read it, and try to get to the answer from that representation. When shown the symbolic representation given below, they often change their mind.

Every representation one brings to a problem carries different information. This one reinforces what we should have known — that the number of particles in a given volume of a gas at constant temperature and pressure remains the same. Thus, when we replace an \( \text{O}_2 \) molecule with a mass of 32 amu (or an \( \text{N}_2 \) molecule with a mass of 28 amu) by an \( \text{H}_2\text{O} \) molecule with a mass of 18 amu, the gas doesn’t get heavier. It gets lighter.

**IMPLICATIONS FOR TEACHING**
For some time now, we’ve been recommending that instructors draw a picture for every single task they work in class from the beginning of the Fall semester until the end of the Spring semester. We find that when this is done, the number of “C’s” in the class goes down and the number of “B’s” and “A’s” goes up. An example of this is shown below.

Question: What is the pH of 100 mL of water to which one drop of 2 M HCl has been added?

This diagram contains all of the information extracted from the statement of the problem as well as information that is derived while one struggles to get to the answer. Part of its success comes from the fact that it is a symbolic representation. Part comes from the fact that it is a second representation, which often brings to our attention details that might not have been obvious from reading the problem. It is also important to recognize that this diagram is a chunking device, as called for in the Johnstone-ElBanna model. It brings together information, thereby reducing the demand on working memory.

REFERENCES


Figure 5

\[
\text{CO}_2\text{H}
\]

Figure 6

\[
\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3 + \text{CH}_3\text{MgBr} \rightarrow \text{Et}_2\text{O} \rightarrow \text{CH}_3\text{CH}_2\text{CH}_2\text{CCH}_3
\]

Figure 7

\[
\text{CH}_3 \text{NCH}_3
\]

Figure 8

\[
\text{CH}_3 \text{NCH}_3
\]

Figure 9

- Dry Air
- Air Saturated with Water
Figure 10

1 Drop (0.05mL)
2 M HCl (K_a=10^6)

100 mL H_2O

What is the pH?

HCl + H_2O → H_3O^+ + Cl^-
100%