# SUPPORTING INFORMATION

**Supplementary Table S1.** Demographics of teachers participating in APEP

| Category          | Conference-based (N = 100) | Distance Learning (N = 56) |
|-------------------|-----------------------------|----------------------------|
| **Subject**       |                             |                            |
| Biology           | 36                          | 28                         |
| Chemistry         | 40                          | 21                         |
| Biology & Chemistry | 21                        | 10                         |
| Other             | 3                           | 4                          |
| **School Type**   |                             |                            |
| Public            | 84                          | 53                         |
| Private           | 13                          | 2                          |
| Parochial         | 3                           | 1                          |
| **School Size**   |                             |                            |
| < 500             | 19                          | 19                         |
| 501-1000          | 28                          | 9                          |
| 1001-2000         | 36                          | 10                         |
| >2000             | 11                          | 18                         |
| **School Minority** |                         |                            |
| Population        |                             |                            |
| < 10%             | 30                          | 14                         |
| 10-39%            | 41                          | 25                         |
| 40-79%            | 16                          | 15                         |
| > 80%             | 13                          | 2                          |
| **School Locale** |                             |                            |
| Urban             | 30                          | 5                          |
| Suburban          | 46                          | 26                         |
| Rural             | 24                          | 25                         |
| **US Location**   |                             |                            |
| Northeast         | 9                           | 0                          |
| Southeast         | 39                          | 20                         |
| Midwest           | 49                          | 6                          |
| West              | 2                           | 30                         |
| Military          | 1                           | 0                          |
**Supplementary Table S2.** Number of Teachers Participating in the Distance Learning Workshops as a Function of Geographic Location

| Location      | Number of Teachers | (N=56) |
|---------------|--------------------|--------|
| **Southeast** |                    |        |
| Coral Glades, FL | 3                  |        |
| Lugoff, SC**   | 6                  |        |
| Memphis, TN    | 6                  |        |
| Abingdon, VA   | 5                  |        |
| **Midwest**    |                    |        |
| Thornton, CO*  | 2                  |        |
| Houston, TX    | 3                  |        |
| Fort Worth, TX† | 15                 |        |
| Johnston, IA*  | 4                  |        |
| Kansas City, MO* | 2                |        |
| **West**       |                    |        |
| Concord, CA    | 3                  |        |
| Hayward, CA    | 3                  |        |
| Pine Valley, CA| 1                  |        |
| Bellingham, WA | 3                  |        |

*Streaming participants; **included 2 broadcast sites; †included 5 broadcast sites
### Supplementary Table S3: Teacher Evaluation of APEP Workshops

| Evaluation Item                                                                 | Distance Learning Workshop | Conference-based Workshop |
|--------------------------------------------------------------------------------|----------------------------|---------------------------|
|                                                                                  | Mean Score (N=51) | % Answering 4 or 5 | Mean Score (N=99) | % Answering 4 or 5 |
| I learned something new in chemistry.                                           | 4.3             | 84.3            | 4.4             | 85.9            |
| I learned something new in biology.                                             | 4.6             | 94.1            | 4.7             | 94.9            |
| I learned something new in pharmacology.                                        | 4.8             | 98.0            | 4.8             | 97.0            |
| I learned something new about integrating biology & chemistry through pharmacology. | 4.7             | 98.0            | 4.8             | 100.0           |
| This workshop stimulated my thinking about new ways of teaching my subject.     | 4.7             | 100             | 4.9             | 100.0           |
| This workshop stimulated my thinking about new ways to integrate chemistry and biology. | 4.6             | 96.1            | 4.8             | 100.0           |
| I found the DL approach effective.                                              | 4.2             | 86.3            | N/A             | N/A             |
| I had no trouble following the lectures delivered via TV.                       | 4.0             | 80.4            | N/A             | N/A             |
| I prefer the DL approach to traveling overnight to a workshop.                  | 4.0             | 76.5            | N/A             | N/A             |
| I prefer traveling to a workshop versus a distance learning approach.           | N/A             | N/A             | 4.8             | 99.0            |
Example of the Story (the Problem)
Module 4 – Alcohol and the Breathalyzer Test

Spritney Beers Blows it Bigtime

The big news is that celebrity singer, Spritney Beers has been arrested for driving under the influence (DUI). Spritney blew a "0.08%" on the Breathalyzer™ when she was stopped by the police after driving over the center line. This value reflects the blood alcohol concentration (BAC) and, in California, it is the value that legally designates intoxication. Despite considerable evidence that the Breathalyzer™ test gives an accurate approximation of the BAC, there is still much discussion about the validity of this test. While in jail, Spritney used her one phone call to contact her attorney.

A court date has been set for three weeks from today. Paparazzi photographed her as she left the courthouse in a heated discussion with her attorney. Reporters at the scene quoted him saying "I am very suspect of the validity of the Breathalyzer™ test given to my client. I only have to establish reasonable doubt in this case."

Back at school, you are discussing the case with a classmate when your science teacher interrupts and decides that the topic is worthy of a class debate. To prepare for the class debate, you decide to learn how a Breathalyzer™ test is used to determine the BAC so that you can debate whether it is accurate or not.

Spritney Beers' level of alcohol intoxication was assessed by measuring the concentration of ethanol she expelled from her lungs. This technique works because a water-based molecule like ethanol that is absorbed from the gut into the bloodstream will reach the lungs, where it is exhaled as a vapor in the air.

1. Diagram the path that alcohol will take from the gut (stomach & small intestine) to the lungs via the circulatory system. Which membranes will ethanol have to cross? Although some alcohol is exhaled from the lungs, the rest stays in the blood. Where does the ethanol in the blood go once it leaves the lungs?

Ethanol is dissolved in the blood and is distributed to organs around the body. It is a volatile molecule and can be vaporized quite easily. In the lung, ethanol is converted from a liquid to a gas, so it can be exhaled in the air.

2. Identify and describe the chemical and physical properties of ethanol that contribute to its volatility. It will help to draw the chemical structure of ethanol.

3. Where in the lung is ethanol vaporized? What role does the lung play in the vaporization of ethanol?

The Breathalyzer™ can approximate a person's blood alcohol concentration (BAC) because the concentration of alcohol vapor in the lungs is directly related to its concentration in the blood. Alcohol
vaporizes in the air sacs (alveoli) of the lungs and achieves an equilibrium with the concentration of alcohol that is still in the blood.

4. What is meant by equilibrium? Are ethanol molecules still moving across the membrane between the capillary and the alveolus?

5. What would happen to the equilibrium if some of the ethanol leaves the alveolar sac by exhalation?

When a person exhales into a breath analyzer such as the Breathalyzer™ tube, the exhaled alcohol reacts with compounds in the Breathalyzer™ chamber to produce a change in color from orange to green. The chemical reaction indicated by the color change involves oxidation and reduction. Silver nitrate catalyzes the reaction.

6. What is oxidized in the Breathalyzer™? What is reduced? What role does the catalyst play in the chemical reaction?

The degree of the color change indicates how much alcohol is present in the expired air and the instrument calculates an actual concentration. However, for the Breathalyzer™ to calculate how much alcohol is present in the expired air sample it must take into account the volume of blood from which the alcohol originated. There is a standard way of describing this relationship; it is called the blood-to-breath ratio or partition ratio. The average blood-to-breath ratio is 2100:1 and this is the value used for legal purposes. The ratio assumes that an equilibrium exists between the blood and the alveolar air.

7. Explain what this ratio means in terms of the concentration of alcohol in the blood and the breath. Does the ratio change as one exhales?

Actually, the ratio can vary between 1500:1 and 3000:1 depending upon a number of factors including a person's age, gender, and genetic makeup. In Spritny Beers' case, the reported BAC was 0.08%, based on the 2100:1 ratio. In fact, she may be anywhere in this range from 1500:1 to 3000:1.

8. Calculate the underestimation and overestimation of Spritney's BAC assuming she had a blood-to-breath ratio of 1500:1 and 3000:1. Would Spritney still be considered legally intoxicated based on your answers?

As a point of discussion, consider the advantages and disadvantages of using the Breathalyzer™ test and decide for yourself whether you would prosecute based on the Breathalyzer™ evidence in this case. Is there other evidence to consider?
Sample Teacher Assessment Items

(Correct answers are underlined.)

Chemistry-related items:

When hydrogen is burned it combines with oxygen to produce water molecules. Which describes the state of each atom?

A. Hydrogen is oxidized, Oxygen is reduced
B. Oxygen is oxidized, Hydrogen is reduced
C. Both atoms are oxidized
D. Both atoms are reduced

Ethanol, an alcohol, is a molecule that has both a polar end (OH group) and a non-polar end containing a chain of 2 C atoms. Which of the following describes the solubility of alcohol molecules with longer C atom chains?

A. they become more soluble in water because they are more polar
B. they become more soluble in water because they are more non-polar
C. they become less soluble in water because they are more polar
D. they become less soluble in water because they are more non-polar

Biology-related items:

Why do neurons use both electrical and chemical signals to communicate with each other?

A. Neurons need chemical signals to stop the electrical signals
B. Neurons need fast and slow ways to pass information between them
C. Electrical signals can’t jump across the synaptic space
D. Electrical signals can degrade over the distance of an axon

A molecule is in the capillaries. What path does it take to move through the circulation?

A. veins, right side of the heart, lungs, left side of the heart, arteries
B. veins, left side of the heart, lungs, right side of the heart, arteries
C. arteries, right side of the heart, lungs, left side of the heart, veins
D. arteries, left side of the heart, lungs, right side of the heart, veins
Sample Student Assessment Items

(Correct answers are underlined.)

Basic knowledge assessment of chemistry

What feature of the hydroxyl group (–OH) makes it possible to form hydrogen bonds with other molecules?

A. electronegative oxygen contains unshared pairs of electrons
B. electronegative oxygen contains an extra electron
C. electropositive oxygen contains unshared pairs of electrons
D. electropositive oxygen contains an extra electron
E. don’t know

Methanol is more volatile than water because:

A. it has a higher boiling point
B. it has a lower boiling point
C. it can form more covalent bonds
D. it can form more hydrogen bonds
E. don’t know

A chain of carbon atoms single-bonded together and saturated with hydrogen atoms (-CH2-) produces a non-polar molecule because:

A. the electron pairs between the carbons are shared unequally
B. the electron pairs between the carbons are shared equally
C. there is an unpaired electron between the carbons
D. there is an equal number of electrons at each end of the molecule
E. don’t know

Basic knowledge assessment of biology

Protein synthesis proceeds in the following order:

A. DNA is translated to mRNA, mRNA carries out transcription to a protein
B. DNA is transcribed to mRNA, mRNA carries out translation to a protein
C. RNA is transcribed to DNA, DNA carries out translation to a protein
D. RNA is translated to DNA, DNA carries out transcription to a protein
E. don’t know

The ability of an electrical impulse to flow along neurons is based on the distribution of ions inside and outside the cell. What is the distribution of ions when the neuron is at rest?

A. K+ high on the outside, Na+ high on the inside
B. K+ high on the outside, Ca++ high on the inside
C. Na+ high on the outside, K+ high on the inside
D. Na+ high on the outside, Ca++ high on the inside
E. don’t know
The brain depends on having enough oxygen to work properly. Trace the path that the oxygen takes to get up to the brain once you breathe it in:

- A. lungs, right side of the heart, arteries, brain
- B. lungs, left side of the heart, arteries, brain
- C. lungs, right side of the heart, veins, brain
- D. lungs, left side of the heart, veins, brain
- E. don’t know

**Advanced knowledge assessment of chemistry**

Why would a police officer ask a person stopped for driving while impaired to take 3 deep breaths before blowing into the Breathalyzer?

- A. to increase the alcohol vapor in the mouth enough to get a reading on the Breathalyzer
- B. to get rid of alcohol vapor in the stomach so it won’t confound the measurement
- C. to accurately reflect the equilibrium concentration of alcohol in the mouth with respect to that in the blood
- D. to accurately reflect the equilibrium concentration of alcohol in the lungs with respect to that in the blood
- E. don’t know

**Advanced knowledge assessment of biology**

Alcohol causes many effects such as poor judgment, drowsiness, and incoordination. Which brain areas are affected to produce these problems (in order)?

- A. cerebellum, limbic system, hypothalamus
- B. frontal cortex, brainstem, cerebellum
- C. limbic system, cerebellum, frontal cortex
- D. hypothalamus, frontal cortex, limbic system
- E. don’t know
Statistical Methods

We use Bayesian methods to estimate the models via the WinBUGS software package (Spiegelhalter et al., 2003). Bayesian methods start with statements of prior beliefs about the model parameters. The prior beliefs are combined with the observed data to update knowledge about the model parameters, resulting in a posterior distribution for the parameters. We assume vague prior beliefs so that the posterior distribution is dominated by the observed data. We fit the models with alternate prior assumptions and found that the results are insensitive to reasonable specifications of prior beliefs.

We estimate the effects of the modules using binomial logistic regressions with random effects for teachers. These models enable us to control for demographic variables that could influence test performance. They also account for the fact that students taught by the same teacher are more likely to have similar test scores than two students taught by different teachers; that is, the models control for teacher effects. We use separate models for the basic knowledge and advanced knowledge tests.

Formally, we model the number of correct answers $Y_{ij}$ of student $i$ in the class of teacher $j$ as $Y_{ij} \sim \text{binomial}(n, p_{ij})$, where

$$\log \left( \frac{p_{ij}}{1-p_{ij}} \right) = X_{ij} \beta + m_{ij1}b_1 + m_{ij2}b_2 + m_{ij3}b_3 + m_{ij4}b_4 + \lambda_{ij}.$$  

Here, $n$ refers to the number of questions in the test under consideration. $X_{ij}$ is a vector of student attributes including race, gender, current course, and two indicator variables: the first indicator equals one for students in the experimental condition and whose teacher attended the distance learning workshop (and the indicator equals zero otherwise), and the second indicator equals one for students in the experimental condition and whose teacher attended the conference-based workshop (and the indicator equals zero otherwise). $\beta$ is a vector of regression coefficients for the students’ attributes. The $m_{ij}$ terms are indicator variables for the number of modules used in the student’s class. For example, if two modules are used, we have $m_{ij2} = 1$ and $m_{ij1} = m_{ij3} = m_{ij4} = 0$; the corresponding $b$ coefficients indicate the effect associated with using that number of modules. We use four indicator variables instead of one continuous variable to allow the response to vary flexibly with module use.

The $\lambda_{ij}$ random effects are interpreted as student-specific ability; for example, $\lambda_{ij} > 0$ means that the student performs better than the average student with the same array of covariates. We assume that these random effects follow the normal distribution.
\( \lambda_{ij} \sim \text{normal}(\alpha_j, \sigma_j^2) \). Finally, we assume that the teacher-level random effects follow the normal distribution \( \alpha_j \sim \text{normal}(0, \tau^2) \).

We restrict the sample to include students whose teachers were in both the experimental and control conditions. Because teachers are present in both conditions, this eliminates differences in teacher characteristics across the conditions and hence reduces the potential for confounding. The conclusions do not change substantially when we expand the analysis to include teachers with only control (2877 students) or only experimental scores (147 students).

Finally, we assessed the fit of our models in two ways. First, we fit a more complex model that relaxes the assumption that the questions are of equal difficulty. This did not improve the fit according to the DIC, which is a model selection criterion for hierarchical models (Spiegelhalter et al., 2002). Second, for the models used to construct the tables, we performed posterior predictive checks (Gelman et al., 2004). This involves simulating new outcomes for the students in the dataset based on the fitted model and then comparing the simulated outcomes to the true \( Y_{ij} \). Discrepancies in the simulated and actual outcomes indicate inadequate fit. These predictive checks do not reveal problems with the model specification.
Supplementary Table S4: Basic knowledge regression.

| Predictor          | Mean    | Standard Error | 95% interval        | Mean Probability of Correct Answer (95% interval) |
|--------------------|---------|----------------|---------------------|--------------------------------------------------|
| Intercept          | -1.426  | 0.032          | (-1.491, -1.365)    | 0.19 (0.18 , 0.20)                                |
| Male               | 0.098   | 0.012          | (0.074, 0.121)      | 0.21 (0.20, 0.22)                                |
| Asian              | 0.100   | 0.025          | (0.052, 0.148)      | 0.21 (0.20, 0.22)                                |
| Black              | -0.080  | 0.022          | (-0.123, -0.037)    | 0.18 (0.17, 0.19)                                |
| Hispanic           | -0.032  | 0.024          | (-0.078, 0.014)     | 0.19 (0.18, 0.20)                                |
| Native American    | 0.069   | 0.042          | (-0.014, 0.149)     | 0.20 (0.19, 0.22)                                |
| Chemistry 1        | 0.178   | 0.030          | (0.119, 0.236)      | 0.22 (0.21, 0.23)                                |
| Chemistry 2        | 0.611   | 0.043          | (0.527, 0.693)      | 0.31 (0.29, 0.33)                                |
| Biology 2          | 0.460   | 0.026          | (0.409, 0.511)      | 0.28 (0.26, 0.29)                                |
| Conference workshop| 0.090   | 0.023          | (0.044, 0.136)      | 0.21 (0.20, 0.22)                                |
| Distance workshop  | 0.034   | 0.028          | (-0.022, 0.088)     | 0.20 (0.19, 0.21)                                |
| 1 Module           | 0.084   | 0.031          | (0.023, 0.146)      | 0.22 (0.21, 0.24)                                |
| 2 Modules          | 0.118   | 0.032          | (0.054, 0.180)      | 0.23 (0.21, 0.24)                                |
| 3 Modules          | 0.165   | 0.036          | (0.093, 0.235)      | 0.24 (0.22, 0.25)                                |
| 4 Modules          | 0.197   | 0.030          | (0.137, 0.256)      | 0.24 (0.23, 0.26)                                |

Supplementary Table S5: Advanced knowledge regression.

| Predictor          | Mean    | Standard Error | 95% interval        | Mean Probability of Correct Answer (95% interval) |
|--------------------|---------|----------------|---------------------|--------------------------------------------------|
| Intercept          | -1.138  | 0.038          | (-1.210, -1.064)    | 0.24 (0.23, 0.26)                                |
| Male               | -0.054  | 0.017          | (-0.087, -0.020)    | 0.23 (0.22, 0.25)                                |
| Asian              | 0.024   | 0.035          | (-0.044, 0.094)     | 0.25 (0.23, 0.26)                                |
| Black              | -0.186  | 0.031          | (-0.246, -0.127)    | 0.21 (0.20, 0.23)                                |
| Hispanic           | -0.079  | 0.035          | (-0.147, -0.012)    | 0.23 (0.21, 0.25)                                |
| Native American    | -0.010  | 0.061          | (-0.124, 0.109)     | 0.24 (0.22, 0.27)                                |
| Chemistry 1        | 0.070   | 0.040          | (-0.006, 0.152)     | 0.26 (0.24, 0.27)                                |
| Chemistry 2        | 0.419   | 0.060          | (0.309, 0.537)      | 0.33 (0.30, 0.35)                                |
| Biology 2          | 0.404   | 0.038          | (0.331, 0.477)      | 0.32 (0.31, 0.34)                                |
| Conference workshop| 0.042   | 0.034          | (-0.023, 0.109)     | 0.25 (0.23, 0.27)                                |
| Distance workshop  | -0.121  | 0.041          | (-0.201, -0.039)    | 0.22 (0.20, 0.24)                                |
| 1 Module           | 0.186   | 0.046          | (0.096, 0.277)      | 0.29 (0.27, 0.31)                                |
| 2 Modules          | 0.200   | 0.046          | (0.110, 0.289)      | 0.29 (0.27, 0.31)                                |
| 3 Modules          | 0.348   | 0.052          | (0.249, 0.448)      | 0.32 (0.30, 0.35)                                |
| 4 Modules          | 0.420   | 0.043          | (0.335, 0.504)      | 0.34 (0.32, 0.36)                                |

For both regressions, the base condition (i.e., the intercept) is a white female in a Biology 1 class whose teacher did not yet attend the workshop and did not use any modules. Probabilities in other rows change one of those characteristics, except for the module use rows, which also assume that the teacher attended the conference-based workshop. Coefficients of variables with 95% intervals that do not contain zero are significantly different from zero (bolded). Interpretations of the coefficients as odds ratios are obtained by exponentiation of the
point estimates and confidence interval limits. For example, a Biology 2 student has $e^{.404} = 1.50$ times the odds of answering an advanced knowledge question correctly compared to a Biology 1 student. The 95% interval for the odds ratio is $e^{.331} = 1.39$ to $e^{.447} = 1.56$.

Gelman A, Carlin JB, Stern HS, and Rubin DB (2004) *Bayesian Data Analysis*, 2nd ed, Chapman and Hall: Boca Raton, FL.

Spiegelhalter DJ, Best NG, Carlin BP, and van der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*. (64)4:583-639.

Spiegelhalter DJ, Thomas A, Best NG, Gilks WR, and Lunn D (2003) BUGS: Bayesian inference using Gibbs sampling. MRC Biostatistics Unit, Cambridge, England. [www.mrc-bsu.cam.ac.uk/bugs](http://www.mrc-bsu.cam.ac.uk/bugs)