Supplementary information for

Uncovering students' incorrect ideas about foundational concepts for biochemistry

Sachel M. Villafañe[1], Jennifer Loertscher[2], Vicky Minderhout[2], and Jennifer E. Lewis*[1]

[1]Department of Chemistry, University of South Florida

[2]Department of Chemistry, Seattle University

*E-mail: jennifer@usf.edu

Published in *Chemistry Education Research and Practice* (2011), vol. **12** issue 2,

This supplementary material provides information about the content of the instrument, descriptive statistics for each group, and reliability and validity results for the scores.

Concepts in the instrument

The first column of Table S1 provides a brief description of the seven concepts in the instrument in the form of a correct statement about each concept. The second column of Table S1 lists the sets of three potential incorrect ideas that served as the basis for the construction of the common distractors across the multiple-choice items associated with a concept.

Distribution of scores:

The descriptive statistics for groups A and B are shown in Table S2. As can be seen from the table, the distribution of the scores is approximately normal, since most of the skewness and kurtosis values are within the range of ± 1. However, for group B, the kurtosis value obtained for the posttest was -1.6, which indicates that the distribution of the scores is slightly flatter than the normal distribution, in other words that the scores are more dispersed from the mean. When we look at the standard deviation in this case, it also indicates a greater variation of the scores on the posttest. Slight deviations from normality are common with small samples.

Validity and reliability of the instrument's scores

It is very important to gather evidence about the validity and reliability of an instrument's scores before drawing conclusions about data gathered with the instrument. Therefore, these aspects of measurement need to be evaluated both when developing and when using an instrument.

Validity is concerned with whether the instrument's scores reflect what the instrument is designed to measure. In this case, three different aspects of validity were evaluated. First, content validity evidence was gathered during our instrument's development process, through panels of experts who evaluated the concepts, the incorrect ideas, and the items to be included on the instrument. Second, cognitive interviews with students were used at the beginning stages of our instrument's development to gather face validity evidence for the items. The last validity aspect evaluated was construct validity. This type of validity was evaluated using Confirmatory Factor Analysis (CFA) for the pretest and the posttest. Different fit indices are used to determine how well the proposed model (seven concepts with three parallel items for each concept) fits the data. In this case, three indices are used including the chi square test of model fit (χ^2), Comparative Fit Index (CFI), and Weighted Root Mean Square Residual (WRMR). The chi square test of model fit (χ^2) compares the proposed model with the best possible model, with a p value less than 0.05 indicating a good fit (Brown, 2006). CFI compares the proposed model with a completely uncorrelated model, and obtaining values greater than 0.90 (Cheng & Chan, 2003) or 0.95 (Hu & Bentler, 1999) is desirable. The last index used is WRMR, which indicates how close the proposed model is to the data, with values less than 1.0 representing a good fit (Brown, 2006). CFA results from the pretest and posttest are shown below in table S3. Those results indicate a good model fit for our instrument according to the accepted cut-offs associated with each fit index.

Reliability is another important aspect of measurement to be evaluated, since it is concerned with the reproducibility of an instrument's scores. There are different approaches to determine reliability. In our case, we were concerned with the degree to which the items related to the same concept would yield scores that were consistent. Do students tend to respond similarly across the set of items associated with a concept? This type of reliability measures the internal consistency of the items and is usually assessed using Cronbach's alpha coefficient. Table S4 presents the Cronbach's alpha coefficients for each concept and for the complete instrument for both pretest and posttest. These results indicate weaker correlation among students' responses to items in concepts where alphas are low, for example in hydrogen bonding and protein function concepts. However, since Cronbach's alpha depends on the number of items, the values are also attenuated simply because each concept has only three items (Cortina, 1993; Murphy & Davidshofer, 2005).

Table S1 Describing the Seven Concepts in the Instrument

Declarative Statement (concepts)	Incorrect Ideas
Bond Energy: When a chemical bond forms, energy is released.	Bond formation requires energy.
	Bond formation sometimes requires energy and sometimes releases energy.
	The strength of the bond determines when energy is released or absorbed when bonds are formed.
Free Energy: The free energy change for a process (ΔG) indicates whether or not a process is spontaneous at a given temperature.	The free energy change for a process indicates whether or not the process releases heat.
	Heat is released in all spontaneous processes.
	A spontaneous reaction proceeds quickly.
London Dispersion Forces: London dispersion forces are the only type of non-covalent interaction that can occur between non-polar molecules.	London dispersion forces are only found in non-polar molecules.
	There are no attractions between non-polar molecules.
	A dipole is not involved in the interaction between non-polar molecules.
pH/pKa: Comparing the pH value of an aqueous solution of substance to the pKa values of an ionizable group gives information about the ionization state of that group.	At the pH=pKa, the group is totally protonated or totally deprotonated.
	When pH is below pKa species are deprotonated or when pH is above pKa, species are protonated.
	The ionizable groups are unaffected by pH.
Hydrogen Bonding: A hydrogen bond is a non-covalent interaction typically between N, O, or F and a hydrogen atom bonded to N, O, or F.	All hydrogens are capable of hydrogen bonding.
	A covalent bond with a hydrogen in it is a hydrogen bond.
	Any polar molecule can make a hydrogen bond.
Alpha Helix: The interior of an alpha helix contains atoms from the protein backbone in close contact.	The interior of an alpha helix contains the side chains (R-groups) of the amino acid residues.
	The interior of an alpha helix contains water molecules.
	The interior of an alpha helix is empty.
Protein Function: Changes in amino acid sequence of a polypeptide sometimes changes protein function.	Changes in amino acid sequence always change protein function.
	Changes in amino acid sequence never change protein function.
	Changes in amino acid sequence only decrease protein function.

Table S2 Descriptive statistics and distribution of the scores

Group	Test	N	Mean	SD	Skewness	Kurtosis	Min	Max
A	Pretest	125	9.1	3.1	.14	-.31	2	16
	Posttest	125	12.5	3.3	.05	-.07	5	20
B	Pretest	11	11.3	3.9	.87	.33	6	19
	Posttest	11	13.5	5.1	.07	-1.6	7	21

Table S3 Chi-square (χ^2) test of model fit and fit indices from CFA (N=136)

Model	df	χ^2	p-value	CFI	WRMR
Pretest	48	53.396	.2746	.993	.763
Posttest	53	66.488	.1009	.987	.814

Table S4 Cronbach's alpha coefficient values for the pretest and posttest (N=136)

Concept	Cronbach's alpha pretest	Cronbach's alpha posttest
21-item instrument	**.62**	**.66**
Bond Energy	.77	.70
Free Energy	.70	.44
London Dispersion	.57	.66
pH/pKa	.48	.68
Hydrogen Bonding	.27	.47
Alpha Helix	.86	.89
Protein Function	.10	.38