Appendix Tables

Table A1. The conceptual learning goals for these activities and the way in which we scored them. The language of these objectives is modified from the Working Group on Teaching Evolution (1998). The criteria used for scoring is shown in this table.

	Scores		
Learning Goals	**0**	**1**	**2**
1. Scientific conclusions are based on data that have been observed, modeled, and/or derived from experiments.	Not depicted	Diagram hints at experimental design, but is missing either hypothesis-testing or empirical &/or modeled data.	Diagram includes aspects of experimental design, including hypothesis-testing and empirical &/or modeled data.
2. Current scientific research depends on previous scientific investigations and influences future ones.	Not depicted	Diagram incorporates the impact of other research incompletely. For example, the student may only recognize the impact of an investigation on future research, without recognizing the influence of previous research on the current investigation. Alternatively, the student may explain how the research relies on previous studies, but does not mention how it might inspire additional research.	Diagram indicates that a scientific investigation relies on previous scientific results *and* influences future, novel research questions.
3. Scientific discovery creates knowledge that is new and unpredictable.	Not depicted	Diagram acknowledges that a hypothesis can be accepted or rejected, but does not indicate that the conclusion may be unexpected or that it may change the direction of research in a discipline. Diagram may indicate that experiments only lead to publishable results when they are consistent with hypotheses.	Diagram acknowledges that a hypothesis can be accepted or rejected, and also indicates that the conclusion may be unexpected and change the direction of a discipline. Diagram may connect experimental results—regardless of outcome—to theory-building.

Table A2. Results for each prediction		
We predicted that after the module, more of the diagrams would…	**We found that…**	**Statistical result**
be complex.	post diagrams are more complex and less linear.	Complexity: (paired t test, $p << 0.01$, Cohen's d=0.75). Linearity: paired t test, $p << 0.01$, $d = 0.51$
reference Testing Ideas.	all diagrams included components of Testing Ideas in their pre- and post-diagrams.	NA
illustrate multiple ways of generating testable ideas, as explained in the Exploration and Discovery component of the *Flowchart*.	most diagrams contain reference to making observations and asking questions. Significantly more post-diagrams include finding inspiration and reading the literature.	Observations: χ^2 test, $p = 0.05$ Questions: χ^2 test, $p = 0.43$ Finding inspiration: χ^2 test, $p << 0.001$ Reading literature: χ^2 test, $p << 0.001$
reference the Benefits and Outcomes and Community Analysis and Feedback components of the *Flowchart.*	only 12-17% of the diagrams mention Benefits and Outcomes before and after intervention, but significantly more post-diagrams reference Community Analysis and Feedback.	Community Analysis and Feedback: χ^2 test, $p<<0.001$ Benefits and Outcomes: χ^2 test, $p = 0.41$
emphasize the concepts that[1] 1) scientific conclusions are based on data that have been observed, modeled, and/or derived from experiments. 2) current scientific research depends on previous scientific investigations and influences future ones. 3) scientific discovery creates knowledge that is new and unpredictable.	 1) all diagrams indicated the target level of understanding . 2) a significant increase in performance, but with majority of post-diagrams continue to represent incomplete understanding. 3) pre- and post-diagrams do not differ significant, and indicate poor-to-intermediate understanding of the importance of discovery.	 NA Mann-Whitney U test, $p << 0.001$ Mann-Whitney U test, $p = 0.23$

[1]The language of these objectives is modified from the Working Group on Teaching Evolution, 1998.