

Article

Development of the Neuron Assessment for Measuring Biology Students' Use of Experimental Design Concepts and Representations

Annwesa P. Dasgupta,[†] Trevor R. Anderson,[‡] and Nancy J. Pelaez^{†*}

Departments of [†]Biological Sciences and [‡]Chemistry, Purdue University, West Lafayette, IN 47907

Submitted April 2, 2015; Revised January 22, 2016; Accepted February 4, 2016
Monitoring Editor: Ross Nehm

Researchers, instructors, and funding bodies in biology education are unanimous about the importance of developing students' competence in experimental design. Despite this, only limited measures are available for assessing such competence development, especially in the areas of molecular and cellular biology. Also, existing assessments do not measure how well students use standard symbolism to visualize biological experiments. We propose an assessment-design process that 1) provides background knowledge and questions for developers of new "experimentation assessments," 2) elicits practices of representing experiments with conventional symbol systems, 3) determines how well the assessment reveals expert knowledge, and 4) determines how well the instrument exposes student knowledge and difficulties. To illustrate this process, we developed the Neuron Assessment and coded responses from a scientist and four undergraduate students using the Rubric for Experimental Design and the Concept-Reasoning Mode of representation (CRM) model. Some students demonstrated sound knowledge of concepts and representations. Other students demonstrated difficulty with depicting treatment and control group data or variability in experimental outcomes. Our process, which incorporates an authentic research situation that discriminates levels of visualization and experimentation abilities, shows potential for informing assessment design in other disciplines.

INTRODUCTION

Given the rapid pace of new discoveries in the biological sciences and the need to extend what students learn about biology beyond the knowledge presented in textbooks, there is a growing trend and body of literature on research experi-

ences for biology students. In support of this, the *Vision and Change* report advocates the development of formal practices like observation, experimentation, and hypothesis testing among core competencies for biology disciplinary practice (American Association for the Advancement of Science [AAAS], 2011). One challenge facing instructors and educational researchers is how to know when students are actually learning these practices from their research experiences or other types of instruction. In this regard, the literature contains several recent reports on the development of assessments of competence in biological experimentation, particularly of experimental design (Hiebert, 2007; Sirum and Humburg, 2011; Brownell *et al.*, 2014; Dasgupta *et al.*, 2014; Deane *et al.*, 2014). However, none of these studies has situated such assessments in a scenario in which the concepts, procedures, and research tools are relevant to all biological contexts or educational environments. Assessments that ask students about drug design, salinity and shrimp growth (Dasgupta *et al.*, 2014), pesticide effects on fish or plants,

CBE Life Sci Educ June 1, 2016 15:ar10
DOI:10.1187/cbe.15-03-0077

*Address correspondence to: Nancy J. Pelaez (npelaez@purdue.edu).

© 2016 A. P. Dasgupta *et al.* CBE—Life Sciences Education © 2016 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Non-commercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

growth of different plant species (Deane *et al.*, 2014), effect of insecticides on frog respiration (Hiebert, 2007), effectiveness of an herbal supplement (Sirum and Humburg, 2011), and poppy growth experiments (Brownell *et al.*, 2014) may be relevant to some biology programs but not others.

Research has shown that scientists describe evidence from experiments, graph data, and draw models to locate interacting cellular and subcellular components when they investigate molecular and cellular mechanisms (Trujillo *et al.*, 2015). However, none of the above-mentioned assessments of competence in experimentation relate to a molecular or cellular context. Furthermore, one might expect appropriate symbolic conventions and representations used by scientists in a particular subdomain of biology to be learned by students. But there is a dearth of research on, let alone appropriate assessments for, the visualization of experimental design, such as abilities to interpret or generate representations or to analyze features like the shape of a graph in predicting how variables for an experiment might be related. Such representations spontaneously generated by scientists or their research tools have been documented in the literature (Trujillo *et al.*, 2015), and the spontaneous use of symbols and visualizations from research tools has been well defined for the field of chemistry (Kozma *et al.*, 2000). However, we do not yet know how to assess student understanding and related conceptual, reasoning, and visual difficulties with the symbol systems used for designing experimental investigations in biology. Thus, given that many modern biology research tools are quite different from those used 25 yr ago, we recognize the value of designing an assessment instrument that provides students with visual representations and prompts them to generate their own diagrams to explore the role of visual representations in the process of experimental design.

In our view, the Rubric for Experimental Design (RED) published by Dasgupta *et al.* (2014), although exclusively

focusing on design and not the rest of the experimental process, comes closest to informing how to develop assessments that would also acknowledge experimentation scenarios in other subdomains of biology. Thus, in the present study, we used the RED in multiple ways in the following three-step process to address the above gaps in our knowledge of what and how to assess competence in biological experimentation. Our three-step assessment development and evaluation process is summarized below and in the flowchart presented in Figure 1:

Step 1A. Draft and fine-tune an “experimentation assessment” in the context of a real-life situation for a particular subdomain of biology with appropriate background knowledge and questions informed by the RED.

Step 1B. Use the novel experimentation assessment to probe for the use of conventional symbol systems and visualizations, including graphs and tables, and to elicit current practices (within a particular subdomain of biology) of representing different experimentation concepts.

Step 2. Determine how well the experimentation assessment reveals the nature of expert knowledge about the biological context and the experiments used to elucidate new knowledge in that context. This means we make use of the RED to determine how the response from a biologist to the experimentation assessment compares with his or her explanation of an experiment without the instrument.

Step 3. Determine how well the experimentation assessment exposes student knowledge about and related difficulties with experiments that investigate new knowledge in the biological research context. This means we make use of the RED to determine how the responses from the target student subjects to the experimentation assessment compare with their explanations of an experiment without the instrument.

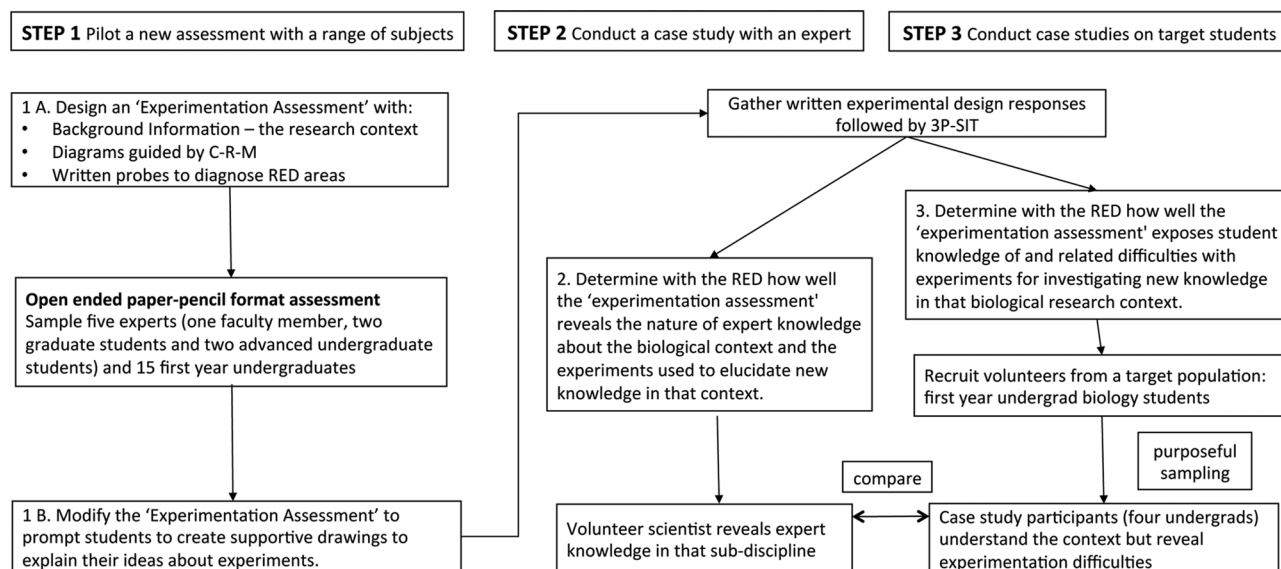


Figure 1. Three steps were used in developing the Neuron Assessment with an authentic biological research situation as the context to measure how well students visualize experiments compared with an expert.

Support for our three-step assessment development process is evident in that it aligns well with all three elements of the assessment triangle framework (cognition, observation, and interpretation) proposed by the National Research Council (NRC, 2001). Steps 1 and 2 involve considering what scientists know and how they demonstrate competence in designing experiments (cognition) in a particular subject domain of biology. By giving the students a background story as the context for an assessment task in a real-life biological research situation, the assessment makes it possible to collect evidence about student abilities to design experiments (observation). Step 2 further compiles the information needed for drawing inferences from the observations by establishing how an expert responds to the assessment task (cognition and observation). In step 3, we propose criteria for evaluating student responses by interpreting how students are performing in comparison with responses and visualizations in the expert response (interpretation).

The Neuron Assessment was designed as an experimentation assessment to understand how scientists and students approach reasoning about experiments. The assessment task asks students to design experiments when given a background story and diagrams to illustrate a biological research context based on neurons involved in a disease. The Neuron Assessment prompts students to investigate the cause for a disorder with impaired mitochondrial movement within neurons by designing experiments premised on the function of biological molecules in a neuron. A limitation is that the Neuron Assessment is situated in a specialized subdomain of biology and therefore may not be transferable as a general assessment tool. However, we provide here a useful methodology for others to guide the design of new ways to probe student understanding of experimental design in other subdisciplines of biology.

BACKGROUND

The Neuron Assessment was designed to provide a context for designing biological experiments. Of relevance to this, previous work reveals that undergraduate students face challenges with aspects of experimental design like knowledge about the experimental subject (Salangam, 2007), manipulating variables (Picone *et al.*, 2007; Shi *et al.*, 2011), identifying measurable experimental outcomes (Hiebert, 2007; Harker, 2009), recognizing sources of variation (Kanari and Millar, 2004; Kuhn and Dean, 2005), and drawing causal inferences (Klahr *et al.*, 1993; Schauble, 1996). On the basis of research using published assessment tasks, Dasgupta *et al.* (2014) developed a Rubric for Experimental Design (RED) organized around five broad areas of students' experimental design difficulties: 1) variable property of an experimental subject, 2) manipulation of variables, 3) measurement of experimental outcome, 4) accounting for variability, and 5) scope of inference of findings. Difficulties in these areas were detected in student responses to the Shrimp Assessment, which presents a context wherein students manipulate various growth-enhancing nutrients and salt levels to design an experiment to track growth of tiger shrimp, and the Drug Assessment, which examines abilities to design an experiment to test a blood pressure drug.

Experimental design can be classified as scientific thinking. Schönborn and Anderson's (2009) CRM model proposes that engagement in any kind of scientific thinking (requires interactions among three factors: conceptual knowledge (C), reasoning skills (R), and mode of representations or visualizations (M). Factor CM, or concepts and the mode of representing them, involve conventional symbols used by scientists when they visualize an experiment. Various skills are involved in recognizing and creating visual representations (Schönborn and Anderson, 2009), such as decoding the symbolic language and interpreting and using the representations when creating your own graphs. Visualization skills are required for scientists to interpret and design experiments. Thus, our rationale was to evaluate whether the skills that experts apply are also applied by students. Similarly, describing the design of a hypothetical experiment requires application of knowledge of the concepts relevant to the subject matter and also experimental design concepts. Therefore, in this study, we examine and compare knowledge of concepts that experts and students present as they propose an experiment using the subject matter of the Neuron Assessment as context. A glossary of vocabulary terms (Dasgupta *et al.*, 2014) was used as a guide to identify experimental design concepts presented by experts and students in their explanations.

In addition to conceptual and visual reasoning abilities, in step 1 (Figure 1), we are also interested in exploring conventional symbol systems and visualizations of mitochondrial transport in the context of neuronal functions. For instance, mitochondrial transport could be depicted using globular or spherical mitochondria moving along elongated rod-like axons, as shown by experts and textbook images. Similarly, conventional symbol systems representing an experiment would be graphical representations of data with the dependent variable on the y -axis to display experimental findings. Reasoning about graphical representations involves organizing the treatment and outcome variables appropriately on the x - and y -axes, whereas reasoning about the concepts related to an experimental design involves, for example, reasoning about treatment and outcome variables to show presence or absence of a causal association in an experiment.

In previous work with the RED, student difficulties with experimental design were only characterized for reasoning about the concepts, because the assessments used to develop RED did not include any diagrams, and students were not prompted to create any visual representations of experiments. Thus, visual reasoning abilities such as construction of graphical representations or reasoning about experimental variables using a graph were not examined. The current study builds on previous work by exploring how students use visualizations when they design experiments. Thus, we define the cognitive element for the current study as including visualization abilities.

For step 1 in the current study, the CRM model was used to guide the design of an original assessment in the context of a cutting-edge research problem. The research problem posed by the Neuron Assessment asks for a method to investigate the source of a disorder associated with mitochondrial movement along axons in neurons. Steps 2 and 3 of this study also apply CRM, along with the RED, in an exploratory qualitative study that examines the usefulness of the Neuron Assessment as a way to observe and compare

expert and undergraduate student knowledge about experimental design. We were interested in applying the three-step assessment-design process (Figure 1) to develop an experimentation assessment to observe expert ways of designing an experiment and to test whether the question is useful to discriminate novice student answers that show difficulties previously reported in the literature from more expert responses (interpretation element). To deeply explore differences in how students and experts think about experiments, we used a case study method. Case studies allow exploration of situations in which the intervention has no preconceived set of outcomes but rather involves examining expert and student knowledge and visual representations of experimental evidence (cognition element) without any relevant behaviors being manipulated. It also covers contextual conditions and allows understanding of the underlying participant experiences and how they influence outcomes from the study (Yin, 1984). For the Neuron Assessment to be a useful measure for discriminating different levels of understanding of experimental design, we would expect it to provide an opportunity for experts and students to present their knowledge and visual depictions related to experiments (observation element) regardless of their prior knowledge of neurons and the movement of mitochondria in neurons.

METHODS

To understand the usefulness of the Neuron Assessment as a probe for revealing expert (Figure 1, step 2) as well as students' knowledge (Figure 1, step 3) about experimental design, we initially designed and piloted several draft versions of the Neuron Assessment as represented in step 1 (Figure 1) and described below. The assessment format was modified to provide clear background information and to minimize any confusion. A neuroscientist was recruited as an expert research participant for the case study oral interview to examine the potential of the Neuron Assessment to reveal the nature of expert knowledge about experimental design concepts and visualizations (Figure 1, step 2). Then, student interviews were conducted and analyzed for the presence of difficulties with experimental concepts and visuals, using expert responses for comparison (Figure 1, step 2) and RED as a tool to characterize expected difficulties (Dasgupta *et al.*, 2014). Each of these steps is detailed in the following sections.

Step 1A: Design of the Neuron Assessment

The Neuron Assessment prompts design of an experiment to investigate a disorder related to organellar movement in neurons (See Box 1 for the question; answers appear in Supplement B in the Supplemental Material). Each part of Box 1, a–c, was logically organized to represent complementary perspectives of organellar movement along neurons based on visual design principles as recommended by Mayer and Moreno (2003). Background information and the diagrams were provided to control for some of the differences in students' prior knowledge of molecular and cell biology so that the Neuron Assessment would more specifically assess their knowledge of experiments. Visual representations have been shown to alleviate misinterpreta-

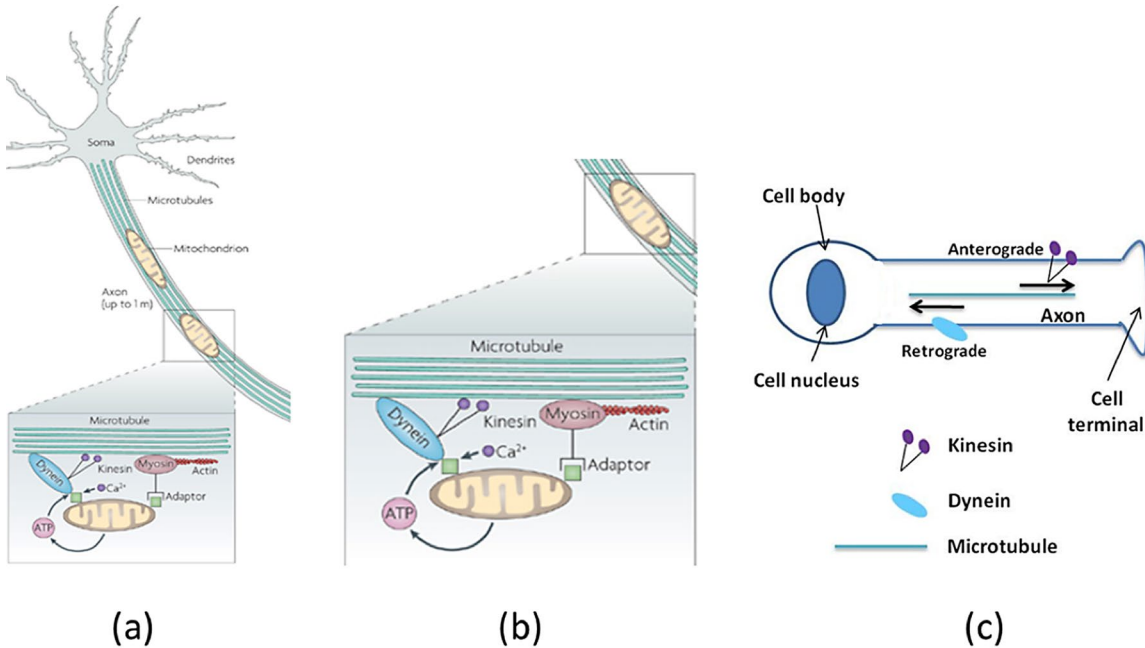
tion by translating across multiple modalities (Stenning and Oberlander, 1995; Moreno and Mayer, 1999). The Neuron Assessment was designed with written probes to diagnose understanding in each of the five RED areas (Dasgupta *et al.*, 2014) (Figure 1, step 1). To probe understanding of experimental subjects, the assessment (Box 1) asks, "How will you assign subjects to groups for your experimental study? Explain." To probe for knowledge of treatment/control conditions, the prompt asks, "Which factors will you vary and which will you keep the same in your study? Why?" The questions "How would you present the results of your experiment?" and "Do you think you can establish a cause-and-effect relationship between the treatment and a response variable in this experiment? Justify your answer" probe for knowledge of measurable outcomes; the assessment probes abilities for dealing with variation and interpreting and representing experimental ideas by asking "How will you improve the validity of your experiment?" and "What results do you expect to get and what would those mean? Using complete sentences, explain what criteria will be used to indicate the success or failure of your experiment." Once designed, the assessment was piloted with a small sample, including the intended study population.

Step 1B: Piloting the Neuron Assessment

Before the study, research procedures were reviewed and approved by the institutional review board (IRB protocol 1008009581). Two sessions were conducted in Fall 2010 and Spring 2012 to pilot the Neuron Assessment, as outlined in Figure 1, step 1. In 2010, 18 first-year undergraduate students and three advanced students (two graduate students and one advanced undergraduate student) participated as volunteers. The assessment was administered as a two-tier multiple-choice test in paper-and-pencil format. Analysis of responses showed that the two-tier format provided only limited information on the nature of students' problems with designing experiments (observation element). Therefore, a second pilot was conducted with a modified open-ended version of the assessment, which was also administered as a paper-and-pencil test. Five experts (one faculty member, two graduate students, and two advanced undergraduate students) and 15 first-year undergraduates participated as volunteers. The pilot study was followed by interviews with the participants, who clarified how the Neuron Assessment could be modified to probe for the five RED areas. This second pilot study also revealed that some students used drawings to explain their ideas about experiments, so the probes were modified to explicitly prompt for drawings to illustrate the role of visualization in designing experiments (Figure 1, step 1).

Steps 2–3: Research Participants

On finalization of the assessment, a scientist who studies neurobiology was recruited as the "expert" volunteer. This expert's research area of focus was related to but did not directly involve the same context as the story of mitochondrial movement for the Neuron Assessment (Figure 1, step 2). Student volunteers were recruited from a first-year undergraduate introductory course for science majors (Biology II: Development, Structure, and Function of Organisms). This course was appropriate, because a key learning objective was to gain biology knowledge through evidence from

Box 1. The Neuron Assessment includes background information and supporting figures

Background: mitochondria are one of the several organelles that get transported along the axon of a nerve (refer to figure above). They are transported in both directions along the length of the axon. The movement of mitochondria from the cell body to the cell terminal is termed as anterograde transport, while the movement from the cell terminal to the cell body, in the opposite direction, is termed as retrograde transport. Movement of mitochondria takes place on the microtubules present along the length of the axons. This complex movement is facilitated by the interaction of motor proteins, kinesin and dynein, present in the axons.

Directions: medical researchers at Seattle Grace Hospital are trying to diagnose the cause for a disorder associated with impaired mitochondrial movement within neurons in human subjects. Cell culture studies have been performed to observe the movement of mitochondria within neurons. The researchers think that kinesin or dynein activity might play a role in the cause of this disorder. Pretend that you work for a company called *MedResearch* that has been assigned to design an experiment to test how kinesin or dynein can affect mitochondrial movement. In your lab you have the following chemicals:

Compound K: inhibits kinesin

Compound D: inhibits dynein

Image software: measures mitochondrial movement in neurons.

- Describe what you see in the three diagrams above. Please summarize in detail what you think about it.
- What could be a potential hypothesis for your experiment? Create an illustration to represent your hypothesis.
- Which factors will you vary and which will you keep the same in your study? Why? Use a visual representation to illustrate the factors you will vary or keep the same.
- How will you assign subjects to groups for your experimental study? Explain. Create a representation to support your answer.
- Do you think you can establish a cause-and-effect relationship between the treatment and a response variable in this experiment? Justify your answer. Create a visual representation to illustrate a cause-and-effect relationship.
- How would you present the results of your experiment?
- What results do you expect to get and what would those mean? Using complete sentences, explain what criteria will be used to indicate the success or failure of your experiment.
- How will you improve the validity of your experiment? What visual representation would show how the validity will be improved?
- What do you think this diagram is not showing? Explain your answer.
- Is there anything about this question that you don't understand or find confusing? Explain.
- Consider yourself a diagram designer. If you could change the diagrams, what would you change or how would you improve them?

research and experimental design and also to practice drawing graphs to represent findings. In 2013, at the beginning of the semester, before any material dealing with experimental design was covered, and as a normal part of their class, students completed a survey using Qualtrics online survey software. The survey offered a sign-up opportunity to all enrolled students to participate in the experimental design activity. Thirteen students agreed to participate.

Using a purposeful sampling strategy, we selected four students for this study (Figure 1, step 3). The selection was based on the following criteria: each student was at the first-year undergraduate level, each student was interested in a different science major, and subjects were selected for broad representation of gender and ethnicities. Prior knowledge or ability was not a factor known to the instructor when these students were recruited at the beginning of the semester, but these students were identified by the instructor as verbally expressive and capable of sharing their own ideas with clarity. For confidentiality, the student participants were given the pseudonyms Juan, Daniel, Eve, and Li Na. The scientist is referred to as an “expert.” Juan is a male Hispanic who is a chemistry major. Daniel is a Caucasian male and engineering major. Eve is a Caucasian female and microbiology major. Li Na is an Asian female who majors in cell and molecular biology. The expert is a Caucasian male who is a neurobiology research scientist.

Case Study Procedure

The written experimental design activity was completed within 1 h by each participant. This was followed immediately by an oral interview that lasted on average 2 h. Oral interviews were recorded digitally and transcribed. On average, each interview involved 6 h of transcription. Data files were stored on a secured computer, and files were transferred using a secure, password-protected file-transfer system as per IRB protocol 1008009581.

The Three-Phase Single-Interview Technique. The three-phase single interview technique (3P-SIT) from Schönborn and Anderson (2009) includes an initial phase (phase 1) with questions to understand each participant’s knowledge of concepts related to mitochondrial transport in neurons and experimental design before exposing him or her to the Neuron Assessment. For example, questions asked were “*What comes to mind when I say ‘neurons’?*” or “*What comes to mind when I say ‘organelle movement’? Please draw to help me understand what you mean.*” In the next phase (phase 2), participants were provided with the Neuron Assessment to permit us to study the impact of the visuals and background information and further examine how participants presented their knowledge of experimental design presented when faced with a current research problem. To understand whether the story with diagrams about transport of mitochondria was intelligible and to find out whether the Neuron Assessment was clear enough to expose participants’ thinking about experimental design, we asked a third set of questions (phase 3) to gather reflections on phases 1 and 2. (See Supplement D in the Supplemental Material for the detailed interview protocol.)

CRM Coding of Interview Responses. The CRM model of Schönborn and Anderson (2009) was used to inform the

coding of the data, as described in Anderson *et al.* (2013). This involved inductively examining the data to code information into three categories: CM, RM, and RC. First, we identified CM or expert conceptual knowledge depicted by the mode of representations deployed by the expert. The expert drawings were examined to identify parts that depict conventional modes of representing both experimental and biological concepts related to neurons and organellar movement using visuals and associated symbolism. CM abilities were added to the glossary (Supplement A in the Supplemental Material) and to the RED (Supplement C in the Supplemental Material). To identify knowledge presented, we modified the RED to include Propositional Statements (meaning correct ideas about experimental design), and the corresponding visual representations for RED components. Further, our original glossary list of vocabulary terms associated with each of five RED areas (Dasgupta *et al.*, 2014) was modified to include how experimental concepts are visualized (Supplement A in the Supplemental Material). The second category, RM, or reasoning with mode of representations, involved inductively identifying the data that indicate reasoning with specific representations. The third category, or RC, indicates retrieving or reasoning with conceptual knowledge of biology subject matter and experimental design concepts in the design of an original experiment. The expert responses were examined to look for parts of an experiment depicted in the form of visuals. This information was added to the glossary, and the glossary list of vocabulary terms was modified and used as a guide to examine visual modes for depicting parts of an experiment drawn by the students.

For step 2, to determine how well the Neuron Assessment reveals the nature of expert knowledge about the biological context and the experiments used to elucidate new knowledge in that context, the expert 3P-SIT interview responses were transcribed (Supplement E in the Supplemental Material) and analyzed using the CRM framework (Figure 1, step 2). The transcript and associated drawings were examined for the conventional symbols (CM) used to describe mitochondrial transport, and these conventions are listed in Supplement F Table 1 in the Supplemental Material. The various visual abilities demonstrated by the expert as he reasoned with diagrams (RM abilities) to represent mitochondrial movement in neurons and experimental design, both before and with the Neuron Assessment, were analyzed. These findings are presented in Supplement F Table 2 in the Supplemental Material. Finally, to compare how the expert reasoned about concepts before and with the Neuron Assessment, the expert interview was coded for knowledge of concepts (RC) relevant to mitochondrial movement and for each component of the RED. Expert abilities were examined according to three categories, that is, evidence of correct ideas (green cells), of difficulties (red cells), and for lack of evidence (yellow cells) when information was missing for that subject matter or a particular RED component. The glossary (Supplement A in the Supplemental Material) was used as a reference to indicate correct knowledge of the experimental concepts presented by the expert. The RC abilities were organized into Supplement F Table 3 in the Supplemental Material to show specific underlying concepts the expert used related to each of the RED components. For example, Supplement F Table 3 in the Supplemental Material compares how the expert reasoned with an underlying concept

related to the RED component, variable property of the experimental subject, before and with the Neuron Assessment.

For step 3 (Figure 1), 3P-SIT interviews were used to probe how well the Neuron Assessment exposed the students' (Juan, Li Na, Eve, and Daniel) knowledge and related difficulties with experiments that investigate new knowledge in the biological research context. The interviews were transcribed (Supplement E in the Supplemental Material), and the transcriptions were analyzed with color coding, as had been done for the expert transcript. In Supplement F, Tables 4–7 (Supplemental Material) were generated to compare diagrams student created before and with the Neuron Assessment, and Tables 8–11 were generated to compare how well each student performed on concepts related to mitochondrial movement in neurons and each component of RED as he or she reasoned about his or her design of a hypothetical experiment.

RESULTS

Expert Abilities Revealed by the Neuron Assessment

Findings highlight the nature of expert knowledge revealed before and with the Neuron Assessment using the guiding CRM framework (Figure 1, step 2). In general, expert CM or conventional symbols associated with the Neuron Assessment include neurons, organelles, motor proteins, microtubules, arrows to point out features and show movement, an experimental design table with treatment groups, and graphs (Table 1). Expert RM abilities displayed in Table 2 show examples of reasoning with diagrams and experimental design visualizations both before and with the Neuron Assessment. Finally, Table 3 shows several examples to compare how the expert designed an experiment using knowledge of specific experimentation concepts (RC) both before and with the Neuron Assessment. Expert RM and RC abilities characterized according to evidence of correct ideas (green cells) and difficulties (red cells) and for lack of evidence (yellow cells) show that reasoning abilities that were missing for a certain

RED component without the Neuron Assessment were clearly expressed by the expert in response to the instrument (cognition and observation elements).

Expert CM Abilities. Table 1 summarizes the conventional modes of representing concepts illustrated in Box 2 when the expert depicted neuronal components or parts of experimental design. The expert used diagrams to illustrate several different conventional ways of presenting mitochondrial movement along axons (Box 2, A–C), including to show how information is organized for the design of experiments (Box 2, D–F). For example by convention, neurons are presented with a circular cell body and elongated axons (Table 1, top row), whereas experimental findings are represented using tables and graphs with various parts labeled (Table 1, last row).

Expert RM Abilities. Table 2 compares how the expert reasoned during the interview with modes of representing information (RM) before and with the Neuron Assessment. The expert both created visuals and used those provided when he reasoned about neuronal functions and experimental design (RM). Box 2, A–F, shows that the expert showcases visual representations that, together with the quotes from the interview (Supplement E in the Supplemental Material), provide evidence for the examples of abilities listed in Table 2 (a more complete list of reasoning abilities is provided in Supplement F Table 2 in the Supplemental Material).

Before seeing the Neuron Assessment, the expert produced diagrams of a neuron (Box 2A) and mitochondrial movement (Box 2B) and depicted tracking of labeled mitochondria (Box 2, C and D) but illustrated no experimental groups. However, with the Neuron Assessment, the expert provided figures and demonstrated RM abilities with experimental tables and graphs (Box 2, E and F) relevant to all five RED components (Variable property of an experimental subject, Manipulation of variables, Measurement of outcome, Accounting for variability, and Scope of inference). Thus, the expert visualized components of an experiment better with the assessment than before being prompted by the Neuron Assessment questions.

Based on the representational modes presented by the expert, the original glossary by Dasgupta *et al.* (2014) was revised (Supplement A in the Supplemental Material) to incorporate visual modes for representing parts of an experiment. Definitions for visual representation of a control, cause-and-effect relationship, factors, outcome variable, sample size, subject, treatment variable, and variability are provided. Consequently, the RED was also modified to incorporate visual evidence associated with each RED area (Supplement C in the Supplemental Material) as detailed in the next paragraph.

The expert depicted control and treatment variables in the experimental table (Supplement C, RED area 1, in the Supplemental Material) and as curves on the x -axis of his graph (Box 1F). Experimental factors were identified in the graph figure legend (Supplement C, RED area 2, in the Supplemental Material). Outcome variables and causal relationships could be interpreted from the x - and y -axes labels on a graph (Supplement C, RED areas 3 and 5, in the Supplemental Material). The expert showed variation with tracking of position of a mitochondrion and thus ways to represent variability in a graph were added (Supplement C, RED area 4, in the Supplemental Material).

Table 1. Knowledge presented by the expert with figures (CM)

CM	Modes of representation with symbols and conventions
Neurons	Circular cell body, elongated axons, small dendritic processes (Box 2A)
Organelles	Globular (Box 2, B and C)
Motor proteins (kinesin and dynein)	Stick figure (Box 2B)
Microtubules	Long strands (Box 2, B and C)
Arrows to identify components	Points at features, movement in anterograde and retrograde directions (Box 2, A, B, D, and F)
Arrows to show movement	Points at features (Box 2B)
Experimental design table	Control and treatment group variables organized into separate columns (Box 2E)
Graph	Independent variable on x -axis, dependent variable on y -axis, key to symbols on the graph show measures for treatment and control groups depicted as separate points or separate bars (Box 2F)

Table 2. Experts' reasoning with visualizations (RM) before and with the Neuron Assessment

Concept	RM ^a	Before	With
Neuron subject matter			
a. Neuron knowledge	Spatially manipulate a representation (Box 2A) ^c	Spatially manipulate a representation (figure of a neuron) to interpret and explain a concept (neuronal anatomy)	Decode the symbolic language composing a representation (Box 1, a–c) Translate horizontally across multiple ERs of organellar movement in neurons (Box 1, a–c)
	Decode a representation		
b. Organelle movement in neurons	Translate horizontally across representations		Temporal resolution of mitochondrial movement along neurons—position of organelle along axon over time (Box 2F)
	Interpret temporal resolution	Temporal resolution of steps in cargo transportation along microtubules during cellular processes of vesicular/organelle transport across neurons (Box 2B)	Translate horizontally across multiple representations of neurons (Box 2E)
	Translate horizontally across representations	Translate horizontally across multiple ERs of a concept (multiple figures representing various aspects of organellar movement) (Box 2B)	
RED area			
c. Experimental design representations Control group ^{2,b} Treatment group ¹⁸	Interpret and use a representation	Provided neuron figures were interpreted to demonstrate design of an observational experiment involving GFP-labeled tracking of mitochondria (Box 2C)	Provided neuron visuals were interpreted to design experimental groups and solve a problem of investigation of organellar movement in neurons (Box 2F)
	Construct a representation	The representation suggests an observational experiment (GFP-labeled tracking of mitochondrial movement along neurons) but no experimental groups were identified	The representation represents manipulation of control and treatment variables organized as separate groups in a table (Box 2F) ^{2,18}

■ Correct ideas; ■ difficulties.

^aRM refers to Reasoning about the Modes of representation with symbols and conventions listed in Table 1.

^bSuperscripts refer to the concepts listed in column 1 and defined in Supplement A in the Supplemental Material.

^cSuperscripts refer to the concepts listed in column 2 and defined in Supplement A in the Supplemental Material.

The expert figures highlight modes of representation that the scientist drew when designing an investigation of mitochondrial movement. The expert decoded neuron knowledge presented as symbols (Table 2, row a). He used the provided figures and constructed ones of his own to design an experiment (Table 2, row c). He used alternative representations to present knowledge of the organellar movement and thus showed horizontal translation (Table 2, rows a and b). Neuron structure was illustrated from organellar to cellular levels, thereby demonstrating his vertical translation abilities, and neuronal anatomy was spatially manipulated to explain a time course for various parts of an experiment (Supplement F Table 2 in the Supplemental Material).

Expert RC Abilities. Table 3 shows that the expert used concepts related to the neuron subject matter as well as experimental design concepts when explaining experimental evidence both before and when exposed to the Neuron Assessment. A superscript number for each concept corresponds to the glossary in Supplement A in the Supplemental Material. RC abilities in adjacent columns show what the expert did or how the concept was used at each stage of the interview. Evidence was identified when the participant

either used the specific term or provided an explanation that indicated knowledge of the concept as defined in the glossary (Supplement A in the Supplemental Material). For example, evidence of knowledge about “variability” using replication was marked as present when the participant stated “replicate the treatments to consider variability among outcomes” or “repeat the treatments to obtain a range of values for the same outcomes.”

Before the Neuron Assessment (phase 1), the expert demonstrated knowledge of neuron concepts but did not propose an experiment with a control group for comparison to test organellar movement in neurons. When the expert said, “Using live-cell imaging and a fluorescent tag to tag some mitochondrial specific protein and track fluorescence as it moves down the axon,” this revealed an observation with no experimental treatment variables. However, with the Neuron Assessment (phase 2), the expert said “To each of these kinesin-impaired and dynein-impaired cell lines I will add compound K, compound D respectively as treatments.” This demonstrates an experimental intervention with treatment variables. During phase 3, the expert said, “I think this is a fairly clear question. You can set up the experiment in a way that will give you some form of answer so it does lead you to derive a certain answer if you have the right ideas about designing an experiment.” These

Table 3. Experts' reasoning with experimental design concepts (RC) before and with Neuron Assessment

	Concept	Before (phase 1)	RC	With (phase 2)	RC
Neuron subject matter					
Neuron concept	b. Organellar movement	i. "[In organellar movement] the cargo is sorted to microtubules and kinesin. So we have microtubules bundles going down the axon and then the kinesin heavy chain help in transporting the cargo (could be organelles) across an axon in a neuron."	Apply knowledge of concepts (molecules like kinesin, microtubules, kinesin heavy chain) to explain organellar movement	ii. "In this study they are trying to test the mechanism for a particular set of neurons with impaired mitochondrial movement, to figure out how to correct the impairment and to repair neurons in patients with the disorder. They are already down to the idea that a defect with either kinesin or dynein is causing the disorder."	Apply knowledge of concepts (neurons, molecules like kinesin, microtubules, dynein) to explain investigation goal of diagnosing impaired mitochondrial movement
RED areas					
Manipulation of variables	Control variable Control ¹ Control group ²	i. "I am guessing since we are only tracking movement in the neurons, a control ^{1,2} won't be necessary at this point."	Lack of evidence	ii. "We will have a control (normal neurons ¹). When nothing is added, we get baseline for anterograde/retrograde speed. To a group of normal neurons we will add compound K and D respectively." ²	Transfer and apply knowledge of the concept of control ^{1,2} for comparison purposes
	c. Controlling outside variables Confounding variables ⁸ Control group ² Treatment group ¹⁸ Variation ²¹	i. "The axons in the study obviously should be picked from the same kind of neurons ²¹ to avoid confounding factors ⁸ that might contaminate our findings."	Apply knowledge of ways to reduce variation ²¹ by controlling confounding variables ⁸	ii. "The factors [across treatment ¹⁸ and control group] ² kept the same would be the imaging set up, conditions of the medium, the cell culture age, time window used to measure, effective concentrations of the inhibitors etc ⁸ . This ensures that any external sources of variation ²¹ are removed in the experiment."	Apply knowledge of matching treatment ¹⁸ and control group ² variables to propose ways to deal with variation ²¹ from confounding variables ⁸
■ Correct ideas; ■ lack of evidence.					

findings indicate that the Neuron Assessment carried sufficient information for design of an experiment to experimentally investigate organellar movement in neurons.

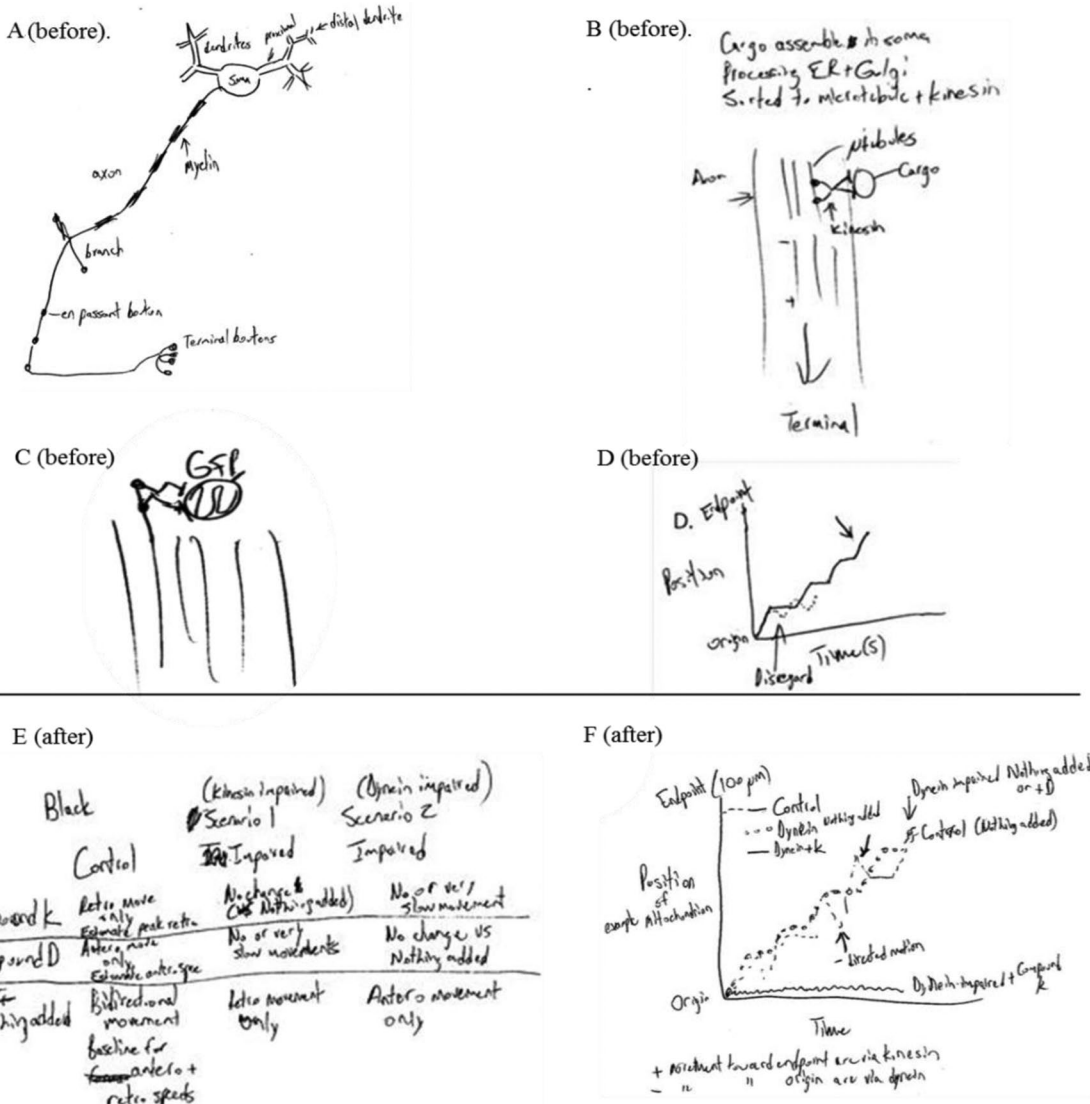
In summary, analysis of the expert response to the Neuron Assessment demonstrated that the assessment provided useful information about neurons and organellar movement in neurons and that the item was effective at revealing the experimental design components identified in the RED. Because the Neuron Assessment successfully revealed expert knowledge of experimental design concepts and ability to use that knowledge with visualizations, we decided to examine students' responses to the Neuron Assessment under the same conditions.

Students' Abilities Observed with the Neuron Assessment

The Neuron Assessment made it possible to observe and measure ideas from four student participants, Juan, Eve, Li Na and Daniel, who also created diagrams to illustrate their ideas about experimental design (Figure 1, step 3). When their responses were examined with the RED, Juan and Eve showed consistent difficulties reasoning with modes of representation both before and with the Neuron Assessment. In contrast, Li Na and Daniel, like the expert, corrected their difficulties when prompted with the Neuron Assessment. Examples of reasoning abilities demonstrated by Eve are

presented in Tables 4 and 5, with more detail provided for both Juan and Eve in Supplement F in the Supplemental Material to illustrate difficulties exposed with the Neuron Assessment. Because responses from Li Na and Daniel were more like the expert responses, for brevity, detailed analysis of their work is only presented in Supplement F. The findings differ for these two groups, so results including drawings by Juan and Eve are presented first, followed by Li Na and Daniel. As was done previously for the expert responses, participant abilities for RM (Table 4 and Supplement F Tables 4–7 in the Supplemental Material) and RC (Table 5 and Supplement F Tables 8–11 in the Supplemental Material) were characterized with RED according to evidence of correct ideas (green cells) and difficulties (red cells) and for lack of evidence (yellow cells) when information was missing for a subject such as a certain RED component. All of the students provided clear evidence of at least some difficulties, and the prevalence of difficulties varied across these four students as indicated by the frequency of red cells in each table. For facilitation of comparisons, the color pattern of RM abilities and difficulties identified from the student responses before and with the assessment from Supplement F Tables 4–7 in the Supplemental Material (Student RM) are summarized in Table 6 next to a list of the various ways the expert reasoned with modes of representation. Similarly, Table 7 summarizes findings presented in Supplement F Table 3 (Expert RC) and Supplement F Tables 8–11 (Student RC)

Box 2. Figures from an expert scientists' response to the Neuron Assessment



Before Neuron Assessment

- (A) Neuron concepts: spatially manipulate provided Neuron Assessment figures to interpret and explain a concept of a neuron. Visualize levels of organization, relative size, shape, and scale of cell body and axon.
- (B) Organellar movement in neurons: use representations to interpret temporal resolution of steps in cargo transportation along microtubules during vesicular/organelle transport across neurons. Translate horizontally across multiple representations of various aspects of mitochondrial movement.
- (C) Interpret and use a representation (provided neuron figures) to demonstrate design of an observational experiment (GFP-labeled tracking of mitochondria). Construct a representation to suggest an observational experiment (GFP-labeled tracking of mitochondrial movement along neurons). Note that experimental treatment groups were not indicated.
- (D) Interpret and use provided neuron figures to demonstrate design of an observational experiment to track GFP-labeled mitochondria. Construct a graph to represent findings from GFP-labeled tracking of mitochondria with independent variables and dependent variables on *x*- and *y*-axes, respectively. Specific treatments are represented as curves. Dotted line presents outliers as a result of variation. Translate horizontally across multiple figures of mitochondrial movement. Interpret the temporal resolution of mitochondrial movement along neurons—position of organelle along axon over time.

Continues

Continued

With Neuron Assessment

- (E) Neuron concepts: decode the symbolic language composing neurons in Neuron Assessment supporting figures (Box 1). Translate horizontally across multiple representations of neurons. Organellar movement: no additional figures were drawn to show organellar movement. RED areas: interpret provided neuron visuals to design experimental treatment groups. Construct experimental groups to represent manipulation of control and treatment variables. Interpret and use a representation (neuron figures) to solve a problem (investigation of organellar movement in neurons).
- (F) Construct a graph to represent curves corresponding to control and treatment outcomes. Construct a graphical representation with independent variables and dependent variables on *x*-axis and *y*-axis, respectively. Different treatments are represented as separate lines. Translate horizontally across experimental table and experimental graph with each treatment as a separate curve. Interpret the temporal resolution of mitochondrial movement along neurons—position of organelle along axon over time.

in the Supplemental Material for concepts pertaining to neurons, mitochondrial movement, and each RED concept (steps 2 and 3 in Figure 1).

Students' RM Abilities: Reasoning with Visualizations of Experimental Design. Students' knowledge and difficulties with modes of representation (RM) were coded using concepts

from the new glossary (Supplement A in the Supplemental Material; underlined sections represent modifications).

RM Abilities for Juan and Eve. Before the Neuron Assessment, when asked about neurons and organellar movement, both Juan and Eve showed spatial manipulation in their own neuron diagrams and visualized orders of relative scales for

Table 4. Examples for Eve's reasoning with visualizations (RM) before and with Neuron Assessment

Concept	RM ^a	Before	With
Neuron subject matter			
a. Neuron knowledge	Spatially manipulate a representation	Manipulated figures of a neuron to explain knowledge of neuronal anatomy (Box 4A) ^c	Decoded the symbolic language composing provided Neuron Assessment figures (Box 1, a–c)
	Decode a representation		Translated across provided representations of neuron and created own visuals of a neuron (Box 4C)
	Translate horizontally across representations		
b. Organellar movement in neurons	Spatially manipulate a representation	Spatial manipulation is flawed, as mitochondrion is depicted in cell body but shows no movement (for example by using arrows; Box 4A)	Lack of evidence, as no organellar movement represented in neuron figures (Box 4C)
RED Areas			
c. Experimental design table/figure	Visualize levels of organization	Relative size and scale of neurons depicted at the organ and cellular level (Box 4B)	
	Interpret a representation		Provided Neuron Assessment figures are used to design experimental groups (Box 4D)
	Control group ^{2,b}		Experimental table represents control and treatment group variables ² (Box 4D)
	Treatment group ¹⁸		
d. Graphs	Construct a representation	Lack of evidence, as no graph was drawn (Box 4B)	Graph drawn with independent variable on <i>x</i> -axis and dependent variable on <i>y</i> -axis ^{2,3} . Different treatments are represented as separate bars (Box 4E)
	RED areas: Manipulation of variables ¹⁷		Experimental table translated graphically with treatments shown as separate bars on the graph appropriately ⁵ (Box 4E)
	Measurement of outcome ⁷		
	Accounting for variability ²²		
	Scope of inference ¹⁵		

■ Correct ideas; ■ difficulties; ■ lack of evidence.

^aRM refers to Reasoning about the Modes of representation with symbols and conventions listed in Table 1.

^bSuperscripts refer to the concepts listed in column 1 and defined in Supplement A in the Supplemental Material.

^cSuperscripts refer to the concepts listed in column 2 and defined in Supplement A in the Supplemental Material.

Table 5. Examples for Eve’s abilities reasoning with concepts (RC) before and with Neuron Assessment

Concept		Before (phase 1)	RC	With (phase 2)	RC
Neuron subject matter					
Neuron concepts	a. Neuron knowledge	“A neuron is connected to other axons to distribute information.”	Memorize knowledge of neurons and axons	“In psychology I have seen similar types of neurons and axons in the brain.”	Apply knowledge of neurons to interpret the experimental context.
	b. Organellar movement	“What’s going on in the mitochondria determines how [organelle] transport occurs.” (RED, Area of Difficulty 1-b)	Reason locally (mitochondrial process) and globally (processes inside mitochondria regulate organellar movement)	“People with the disorder are unable to perform transport and scientists believe that it has to do with motor proteins—kinesin and dynein—not working and their effect on movement of mitochondria.”	Apply knowledge of concepts like transport, kinesin, dynein, mitochondria to explain the investigation goal.
RED Areas					
Manipulation of variables	Control variable Control ^{1a} Control group ²	“[Scientists] are going to need the control, ^{1,2} which would be people that don’t have the disorder so healthy neurons and experiment would be people that carry the unhealthy neurons.” (RED, area of difficulty 2-j)	Reason globally about control ^{1,2} (Experimental subjects carrying obvious differences are assigned to experimental vs. control group.)	“Neurons without any proteins ² [kinesin or dynein].” (RED, Area of Difficulty 2-h)	Transfer and apply knowledge of control (Control group ² does not provide natural behavior conditions, because absence of the manipulated variable in treatment group, results in conditions unsuitable for the experimental subject.)
	c. Controlling outside variables Confounding variables ³ Control group ² Treatment group ¹⁸		Lack of evidence	“Neurons in control ² and experimental group ¹⁸ will both carry same organelles. ⁸ ”	Apply knowledge of controlling confounding variables ³ to have uniform experimental subjects in control ² and treatment ¹⁸ groups.

■ Correct ideas; ■ difficulties; ■ lack of evidence.

^aSuperscripts refer to the concepts listed in column 2 and defined in Supplement A in the Supplemental Material.

various anatomical parts (Boxes 3A and 4A). However, they struggled to represent organellar movement; Juan showed no diagrams of an organelle before being given the Neuron Assessment, while Eve did not show any spatial manipulation, as her diagrams represent mitochondria but fail to show movement (Box 4A). Juan showed no evidence in his diagrams of reasoning about RED areas without the assess-

ment. Eve depicted neurons in an MRI scan at the organ level (Box 4A) and then zoomed in to a microscopic image (Box 4B). Hence, Eve represented these visualizations across orders of magnitude.

Once he was given the Neuron Assessment (Box 1), Juan demonstrated a range of visual abilities as he decoded the provided diagrams and spatially manipulated his own

Table 6. Expert and student reasoning with visualizations (RM) of experimental design

RM ^a	Expert		Juan		Eve		Li Na		Daniel	
	Before	With	Before	With	Before	With	Before	With	Before	With
1. Decode symbolic language		■		■		■		■		■
2. Interpret and use a representation	■	■		■		■		■		■
3. Construct a representation	■ (diff)	■	■	■ (diff)		■		■	■	■
4. Translate horizontally among alternative representations of the same phenomenon	■	■		■ (diff)		■		■		■
5. Visualize levels of organization	■		■	■	■		■		■	
6. Interpret the temporal resolution	■	■					■			
7. Spatially manipulate a representation	■		■	■	■ (diff)		■		■	

■ Before and with Neuron Assessment; ■ before and with Neuron Assessment only; ■ with Neuron Assessment only.

^aRM refers to reasoning about the Modes of representation with symbols and conventions listed in Table 1, and “x” means present, “x (diff)” means present with clear evidence of difficulty, and blank corresponds to lack of evidence.

Table 7. Expert and student reasoning with concepts (RC) of experimental design^a

RC concept	Expert		Juan		Eve		Li Na		Daniel		
	Before	With	Before	With	Before	With	Before	With	Before	With	
1. Neuron	x	x	x	x	x	x	x	x	x	x	
2. Organellar movement	x	x		x	x(diff)	x	x(diff)	x	x(diff)	x	
3. Experimental subject	x	x	x	x(diff)	x	x(diff)	x	x	x	x	
4. Variable	x	x	x	x(diff)	x	x(diff)	x	x	x	x	
5. Treatment variable		x	x	x(diff)	x	x(diff)	x	x	x	x	
6. Treatment group		x	x	x(diff)	x	x(diff)	x	x	x	x	
7. Control variable		x			x(diff)	x(diff)		x(diff)	x	x	
8. Control	x	x		x(diff)	x(diff)	x(diff)		x(diff)	x(diff)	x	
9. Control group		x		x(diff)	x(diff)	x(diff)		x(diff)	x(diff)	x	
10. Controlling outside variables	x	x				x		x	x	x	
11. Confounding variables	x	x				x		x	x	x	
12. Variation	x	x			x(diff)			x	x	x	
13. Outcome variable	x	x	x(diff)	x(diff)	x(diff)	x(diff)	x(diff)	x	x(diff)	x	
14. Replication	x	x			x(diff)			x	x	x	
15. Variability	x	x			x(diff)			x	x	x	
16. Randomization	x	x						x		x	
17. Representative sample	x	x				x(diff)		x		x	
18. Scope of Inference		x	x(diff)			x(diff)				x(diff)	x
19. Cause and effect		x		x(diff)	x(diff)	x(diff)				x(diff)	x
20. Correlations		x		x(diff)	x(diff)	x(diff)		x(diff)	x(diff)	x	

■ Before and with Neuron Assessment; ■ before and with Neuron Assessment only; ■ with Neuron Assessment only.

^a "x" means present, "x (diff)" means present with clear evidence of difficulty, and blank corresponds to lack of evidence.

images of neurons and organellar movement using appropriate orders of relative size and scale (Box 3B). However, he did not depict any organellar movement after being given the Neuron Assessment. Similarly, Eve decoded the provided neuron diagrams (Table 4, row a). With the Neuron Assessment, she spatially manipulated her diagrams to represent anatomical parts and motor proteins kinesin and dynein with a neuron (Box 4C) but still did not represent any movement of organelles in neurons (Table 4, row b). For RED areas, Juan was able to construct an experimental table (Box 3C) but showed difficulties with horizontal translation from table to graph, as there was a mismatch for experimental groups between the table and graph (Box 3D). In contrast, Eve demonstrated correct RM abilities, as she was able to construct an experimental table as well as design the corresponding graph (Box 4, D and E, and Table 4, rows c and d).

RM Abilities for Li Na and Daniel. Before the Neuron Assessment, both Li Na and Daniel were able to demonstrate a range of RM abilities, as they drew diagrams of a typical neuron with relative sizes for various anatomical parts but failed to depict any organellar movement (Boxes 5A and 6A). Regarding RED areas, Li Na did not provide any visualization, but Daniel constructed a representation of experimental groups (Box 6B) by drawing impaired and healthy patients. With the Neuron Assessment, both were able to decode neuron and organellar movement diagrams and translate between neuron images provided. They also represented corresponding experimental findings using graphs (Boxes 5C and 6D).

To summarize, before students were exposed to the Neuron Assessment, all four showed no evidence of depicting any movement of mitochondria along neurons and also no graphical representations of experimental results.

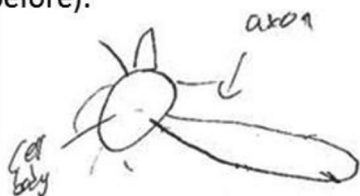
However, with the Neuron Assessment, Eve, Li Na, and Daniel were able to interpret the supportive diagrams and create their own experimental design tables and graphs, but Juan exhibited difficulties, with his Neuron Assessment response revealing no evidence of mitochondrial movement and clear evidence of difficulty with constructing a graph.

Students' CM Abilities: Reasoning with Concepts of Experimental Design. The students presented knowledge of the subject matter and experiments as they explained investigations designed to study a disorder with mitochondrial movement in neurons. Supplement F Tables 8 and 9 (Supplemental Material) show knowledge and difficulties with subject matter and experimental design (RC) before and with the Neuron Assessment. We characterized correct ideas (green boxes) and difficulties (red boxes) with concepts relevant to mitochondrial movement and each component of the RED. For example, Eve's consideration of a control group ("Neurons without any proteins [kinesin or dynein]") showed evidence of difficulty with the concept, as the variable being manipulated (*without any proteins*) results in conditions unsuitable for the experimental subject. A superscript number for each concept corresponds to the definition in the glossary (Supplement A in the Supplemental Material). RC abilities in adjacent columns show what students did or how the concept was used at each stage of 3P-SIT. Evidence was identified either when the students correctly used the specific term or provided an explanation that indicated knowledge of the concept as defined in the glossary.

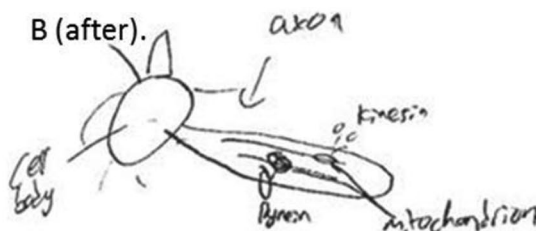
The RC analysis revealed difficulties or lack of evidence with concepts related to both mitochondrial movement in neurons and components of the RED. In brief, for Juan and Eve, RC abilities before and with the Neuron Assessment indicated

Box 3. Juan's Neuron Assessment figures

A (before).



B (after).



C (after).

Control	Group 1	Group 2
Contains both kinesin and dynein	Lacks kinesin (component K)	Lacks dynein (component D)

D (after).



Before Neuron Assessment

(A) Neuron concepts: spatially manipulate a representation to interpret and explain the concept of neuron knowledge with neuronal anatomy. Visualize levels of organization, relative size, shape, and scale of cell body and axon. Organellar movement: lack of evidence (no mitochondria or organellar movement is represented). Figures depict no experimental design skills.

With Neuron Assessment

- (B) Neuron concepts: decode the symbolic language composing provided Neuron Assessment figures. Spatially manipulate figure of a neuron to explain knowledge of kinesin, dynein, and a mitochondrion. Visualize levels of organization, relative size, shape, and scale of cell body, axon, motor proteins, and mitochondrion. Organellar movement in neurons: lack of evidence (no organellar movement is represented).
- (C) RED areas: interpret provided visuals to design experimental groups. Construct an ER to represent manipulation of control and treatment variables organized as separate groups.
- (D) Construct a graph (graph is flawed as inappropriate independent variables are represented on x-axis). Translate horizontally across experimental table to experimental graph (the groups represented in the experimental table do not correspond to the bars on the graph).

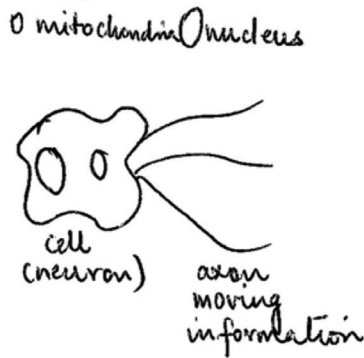
that, while there were some positive modifications to their knowledge, most of their difficulties before the assessment were consistent even when given the Neuron Assessment. In contrast, Li Na and Daniel showed many more correct ideas when given the Neuron Assessment. Concepts that showed “lack of evidence” were developed into knowledge when they were probed with the Neuron Assessment. Below is a detailed account of the interview findings from the students’ raw transcripts (Supplement E in the Supplemental Material).

RC Abilities for Juan and Eve. Findings related both to the neuron subject matter and the five RED areas are shown in Supplement F in the Supplemental Material. Without the Neuron Assessment, both Juan and Eve correctly depicted knowledge of a neuron but showed flawed or lack of knowledge

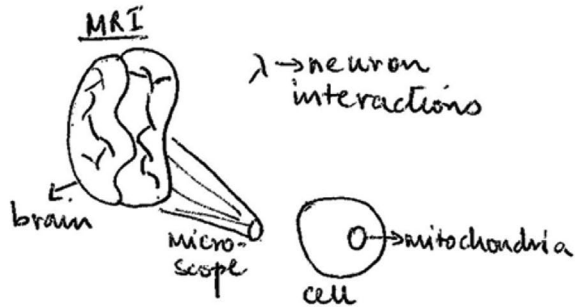
about organellar movement in neurons. When probed to think about how scientists discovered this information, both chose to describe experiments researchers may have carried out, which demonstrates ability to reason with concepts of experimental design. Their descriptions provided evidence of their existing knowledge for RED areas. Both integrated knowledge of subject matter concepts to propose the variable property of the experimental subject. For manipulation of variables, they presented mixed responses. Both appropriately applied knowledge of the treatment variable, but Eve had difficulties with reasoning about control groups (Table 5), while Juan showed lack of evidence for controls. Both participants also provided no information to control confounding variables in the studies they proposed, and both displayed difficulties applying knowledge of an outcome variable to

Box 4. Eve's Neuron Assessment figures

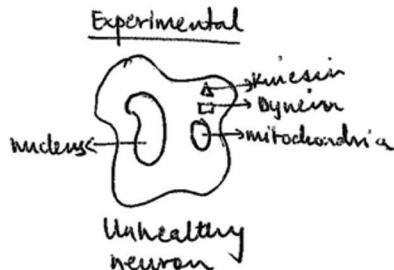
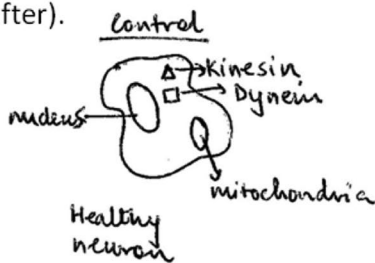
A (before).



B (before).



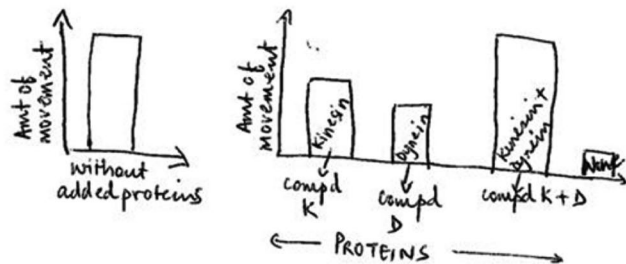
C (after).



D (after).

Control	Experimental			
	kinesin	dynein	kinesin + dynein	No proteins
Inject compound K	Inject compound D	Inject both K + D		

E (after).



Before Neuron Assessment

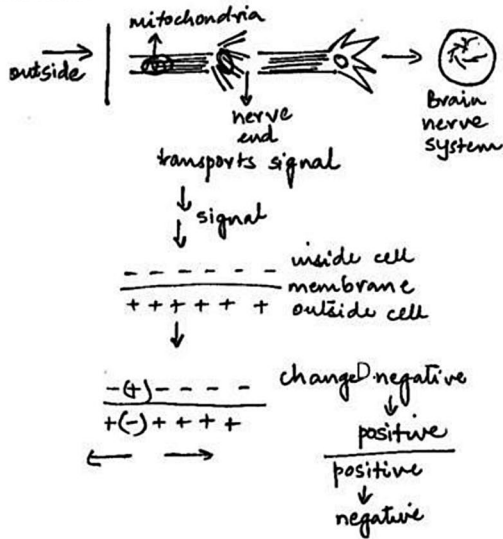
- (A) Neuron concepts: neuron knowledge: spatially manipulate a representation to interpret and explain the concept of neuron knowledge with neuron body and axons. Visualize levels of organization, relative size, and scale (relative size and shapes of cell body and axon). Organellar movement in neurons: spatially manipulate an ER to interpret and explain a concept. Mitochondria represented in the cell body, but its movement (for example by using arrows) is not represented.
- (B) RED areas: visualize levels of organization and relative size and position of neurons relative to the organ- and cellular-level diagrams.

With Neuron Assessment

- (C) Neuron knowledge: decode the symbolic language composing provided Neuron Assessment figures. Translate horizontally across provided representations of neuron and create own visuals of a neuron. Organellar movement in neurons: lack of evidence (no organellar movement represented in visual representation of neurons).
- (D) RED areas: interpret an ER (provided visuals) to design experimental groups. Construct an ER (experimental table constructed to represent control and treatment variables organized as separate groups).
- (E) Construct an ER (graphical representation) with independent variable on *x*-axis and dependent variable on *y*-axis. Different treatments are represented as separate bars. Translate horizontally across experimental table and graph representing each treatment in the table as separate bars.

Box 5. Li Na's Neuron Assessment figures

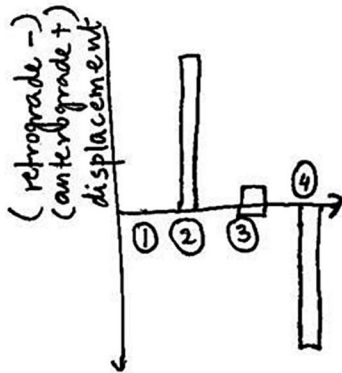
A (before).



B (after).

- Control group
 Experimental group
- ① Inhibiting kinesin + Inhibiting dynein
 - ② Activating kinesin + Inhibiting dynein
 - ③ Activating " + Activating "
 - ④ Inhibiting " + Inhibiting "

C (after).



Before Neuron Assessment

(A) Neuron concepts: spatially manipulate a representation to interpret and explain the concept of neuron knowledge with neuronal anatomy. Visualize levels of organization, relative size and scale of nervous system, cell body, axon, and mitochondria. Interpret the temporal resolution (shows signal transmission across cell as mode of neuronal communication).

Organellar movement in neurons: lack of evidence (mitochondria are represented, but movement of signals rather than of mitochondria is depicted). Lack of evidence (figure shows no evidence for experimental design skills).

With Neuron Assessment

Neuron concepts: no additional diagrams drawn. Organellar movement in neurons: no additional diagrams drawn.

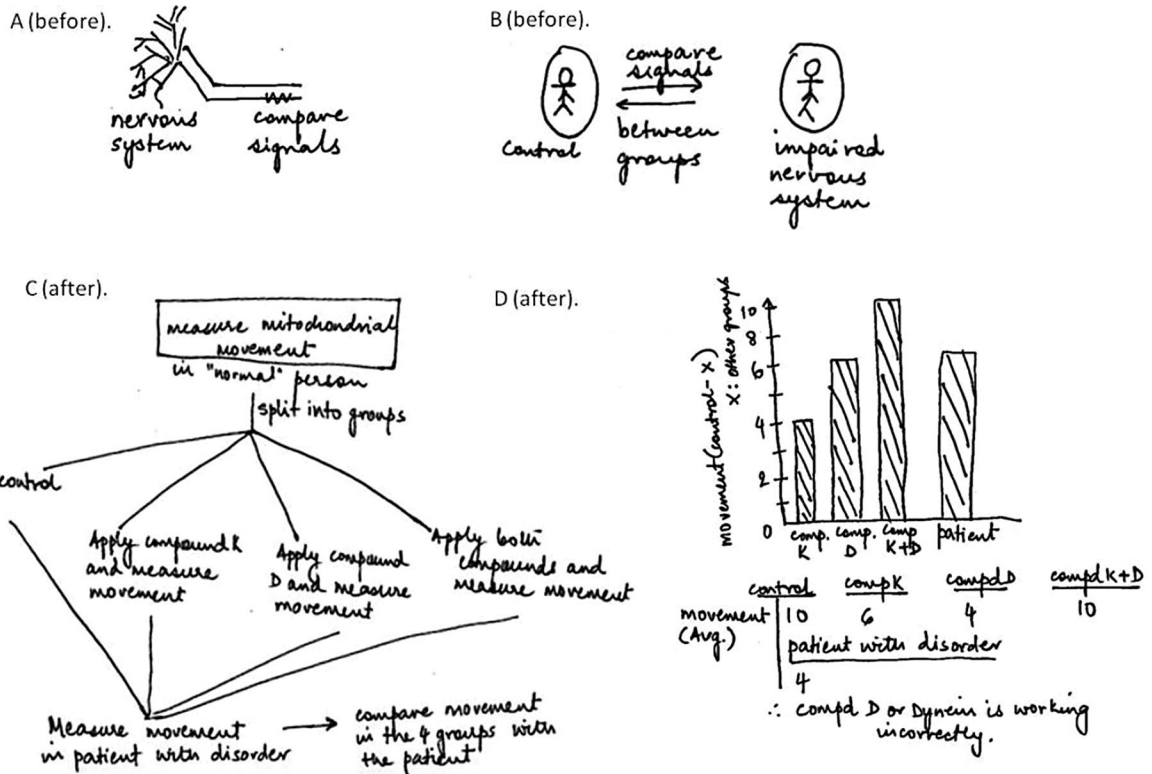
(B) RED areas: interpret provided neuron visuals to design experimental groups. Construct a graph to represent manipulation of control and treatment variables organized as separate groups. Note that treatments 1 and 4 are identical. Treatment 4 was meant to be inhibiting kinesin and activating dynein.

(C) Interpret provided visuals to design experimental groups. Construct a graph to represent control and treatment variables organized as separate groups; independent variables and dependent variables are represented on x-axis and y-axis, respectively. Translate horizontally across experimental table and experimental graph representing each treatment in the table as separate bars on the graph appropriately.

propose suitable measures. They shared no knowledge about ways to account for variability such as replication, randomization, and using a representative sample. Eve presented a difficulty with failure to show replication. For Juan and Eve,

flaws with knowledge of manipulation of variables and accounting for variability resulted in missing or deficient scope of inference and causal claims that did not align with the goal for the investigation.

Box 6. Daniel's Neuron Assessment figures



Before Neuron Assessment

- (A) Neuron concepts: spatially manipulate a representation to interpret and explain the concept of neuron knowledge with neuronal anatomy. Visualize levels of organization, relative size, and scale of axon and dendrites. Organellar movement in neurons: lack of evidence (no representation was created to depict organellar movement).
- (B) RED areas: construct a representation to explain experimental groups considered and measurement of outcome.

With Neuron Assessment

Neuron concepts: no new diagrams drawn. Organellar movement in neurons: no new diagrams drawn.

- (C) RED areas: interpret a representation (provided neuron visuals) to design experimental groups. Construct a representation to represent manipulation of control and treatment variables organized as separate groups.
- (D) Construct a graph representation with independent variables and dependent variables on x - and y -axes, respectively. Different treatment groups are represented as separate bars. Translate horizontally across experimental table and experimental graph representing each treatment in the table as separate bars on the graph appropriately.

With the Neuron Assessment, both Juan and Eve correctly interpreted the assessment context and supporting figures. When asked about how scientists would find the cause of the disorder, they suggested designing an experiment. When probed to elaborate ideas about how one would specifically plan that experiment, Juan had difficulty with knowledge of neuron concepts. He described experimental procedures that revealed problems in all five RED areas with reasoning about treatment variables and knowledge of control variables. Juan did not apply knowledge of outcome variables to propose a suitable measure. Also, no evidence was provided to show how variability measures would be handled. No causal conclusions would be possible from Juan's experimental design owing to missing variability measures

and inappropriate treatment suggestions. Even though Eve demonstrated correct knowledge of neurons and organellar movement along neurons (Table 5), when she designed an experiment, difficulties with concepts belonging to four RED areas became apparent. But she showed correct ideas for controlling outside variables (Table 5). Correct knowledge was shown for the variable property of the experimental subject. She also showed lack of evidence for replication and randomization.

In summary, before the Neuron Assessment, Juan's difficulties with RC abilities in all five RED components were consistent with difficulties revealed with the Neuron Assessment. Without the assessment, Eve was able to reason about the experimental subject but showed difficulties with manipulation

of variables, measurement of outcome, accounting for variability, and scope of inference. With the Neuron Assessment, she was able to reason with knowledge of the experimental subject overall and about controlling outside variables as part of accounting for variability. But Eve still revealed difficulties with at least one or more concepts under four RED areas: manipulation of variables, measurement of outcome, accounting for variability, and scope of inference.

When both Juan and Eve were asked to critically evaluate their experiments with the Neuron Assessment (phase 3 of 3P-SIT), both found the Neuron Assessment background easy to decipher (*“the background does sum up the basics”*). However, they asserted that designing an experiment was rather difficult when they did not know an expected outcome. For example, Eve said, *“It is very difficult to come up with an experiment if you don’t understand what you are supposed to find out eventually.”*

RC Abilities for Li Na and Daniel. In general, Li Na and Daniel performed better than Juan and Eve, both before and with the Neuron Assessment. Before the Neuron Assessment, both Li Na and Daniel accurately presented knowledge of neurons but showed difficulty applying knowledge of organellar movement in neurons (Supplement F Tables 10 and 11 in the Supplemental Material). Both were able to reason about experiments with concepts relevant to the variable property of experimental subject, but they presented mixed abilities with knowledge of manipulation of variables. Li Na did not show any knowledge about treatment variables or control of variables, in contrast to Daniel, who only exhibited difficulty applying his knowledge and reasoning to control of variables. Li Na also showed lack of knowledge about confounding variables, but Daniel presented correct knowledge of this concept. Li Na presented knowledge of outcome variables with flawed outcome measures by suggesting *“displacement of mitochondria”* as a measure, and Daniel also had difficulty measuring dependent variables, suggesting signal strength or pathway as a measure. Li Na did not address how to deal with or measure variability. In contrast, Daniel showed that he knew there was a need to replicate measures. Li Na did not provide evidence for reasoning about causal claims owing to lack of evidence for reporting variability in measures. Daniel showed difficulty with reasoning about inferences and causal claims from his experimental findings because he did not identify appropriate measurable outcomes or propose ways to measure variability as part of experimental findings.

With the Neuron Assessment (phase 2), Li Na and Daniel accurately presented their knowledge of neurons. Li Na also appropriately applied knowledge of RED components, variable property of experimental subject, and measurement of outcome and variability. She showed difficulty with the concept manipulation of variables, and she struggled to reason about controls and causal explanations. In contrast, Daniel sufficiently applied his knowledge of concepts from all five RED areas. He also reasoned locally and globally about concepts like variability measures and causal claims to draw appropriate inferences from findings of his experiment after he was given the Neuron Assessment.

In summary, without the assessment, Li Na showed knowledge of the RED components variable property of

experimental subject, measurement of outcome, and accounting for variability. This is also consistent with her response when given the assessment, but the assessment elicited a difficulty with *“control,”* whereas there was a lack of evidence before she was given the Neuron Assessment. Without the Neuron Assessment, Daniel displayed difficulties with manipulation of variables, measurement of outcome, and scope of inference. He corrected these difficulties when he reasoned about concepts of experimental design in response to the probing questions and background information of the Neuron Assessment.

As feedback (phase 3), Li Na and Daniel both found the experimental design activity to be quite enjoyable (*“I can come up with a lot of ideas, so I am comfortable with activities like this”*). They also considered the background information quite useful for designing an experiment (*“The diagrams definitely helped me think about the process more clearly, since I did not know about this process too much before this study. I think it helped me see how things like the mitochondria, kinesin, and dynein are placed within a neuron”*). Nevertheless, they expressed discomfort about being uncertain whether they had correctly given the expected answer for the experiment (*“I don’t know the right answer to this experiment so whether the question is good depends on the answer”*).

DISCUSSION

In this section, patterns for expert and student reasoning with modes of representations (RM) are presented (Table 6) followed by patterns for reasoning with experimental design concepts (RC) (Table 7). As discussed in this section, evidence suggests that the Neuron Assessment is especially useful as a probe for some specific details of the RED areas.

In terms of our step 2 (Figure 1), in which we determined how well the Neuron Assessment reveals the nature of expert knowledge about the biological context and the experiments used to elucidate new knowledge in that context, we found that the instrument provided evidence for us to observe expert reasoning about experiments. Furthermore, application of step 3 (Figure 1), in which we determined how well the Neuron Assessment exposes student knowledge about and related difficulties with experiments, revealed that the instrument provided students with adequate information to demonstrate how they either soundly or unsoundly reasoned with visual representations (RM) and experimental design concepts (RC) to support their ideas about investigating a current research problem. In general, the Neuron Assessment was far better than 3P-SIT (phase 1) at revealing evidence and details of student knowledge and difficulties. This was clearly apparent when the data from the *“before”* and *“after”* Neuron Assessment were compared for each of the four students and the expert. The instrument proved very effective in revealing that Juan and Eve displayed nearly all the difficulties documented by the RED. Furthermore, in the case of Li Na, Daniel, and the scientist, who demonstrated good understanding of experimental design, the instrument was very effective in exposing details of such sound knowledge. In the following sections, we elaborate on these findings with respect to how well the Neuron Assessment instrument revealed details of expert and student reasoning with visualizations (RM) and concepts (RC).

Expert and Student Reasoning with Visualizations (RM) of Experimental Design

Findings from the expert's knowledge (step 2), revealed by the Neuron Assessment, indicate that spatial manipulation across representations (Table 6, row 7) for experimental design could be observed. Trujillo *et al.* (2015) in the MACH model study showed that a neurobiologist and a cancer biologist infer a mechanism from experimental/temporal data, whereas a structural biologist infers a mechanism from spatial research findings. In reality, all mechanisms involve both spatial and temporal changes. Yet, the current findings with the Neuron Assessment indicate that experimental design by the expert scientist was often interpreted without referring back to the spatial (in most cases) or temporal (in some cases) features of the neuron.

Application of the CRM model to drawings made by the expert provided information about how the RED can be modified to capture an expert's experimental knowledge and use of visualizations during the design of an experiment. The expert created supportive illustrations each time he explained a concept of relevance to experimental design.

The diagrams provided as part of the Neuron Assessment (Box 1) proved suitable for the expert and all students to decode the information presented (Table 6, row 1). All participants used the information provided to construct their own diagrams relevant to investigations they designed for the Neuron Assessment (Table 6, row 3). The assessment was very effective in stimulating participants to provide evidence for interpretation and use of representations to solve a problem, as well as horizontal translation across representations (Table 6, rows 2 and 4). This was apparent from the fact that three out of four students who did not initially show these abilities were able to do so when responding to the Neuron Assessment (Figure 1, step 3).

In contrast, our limited data suggest that the Neuron Assessment may not provide reliable evidence to show visualization of the levels of organization (Table 6, row 5) to do with neuronal anatomy and mechanisms with neurons, like organellar movement and signal transduction. This was apparent from the fact that the two students and the expert visualized more of the neuronal features relevant to the experiments they were describing before taking the Neuron Assessment than when they took the actual assessment. However, after taking the assessment, they did continue to refer back to the ideas that they had already explained, suggesting that taking the assessment had been a stimulating experience for them. The expert, but not the students, was stimulated by the assessment to interpret temporal resolution. In fact, when taking the assessment, all students chose to represent comparison groups rather than time course graphs.

In summary, the Neuron Assessment provides useful evidence for RM abilities, as the more proficient students, Li Na and Daniel, demonstrated visual abilities like the expert before and with the assessment, while Juan and Eve, who demonstrated more limited visual abilities before taking the assessment, improved considerably in their visualization of neurons but not in their visualization of experiments once they were exposed to the assessment.

Expert and Student Reasoning with Concepts (RC) of Experimental Design

The context of the Neuron Assessment was sufficient for providing evidence related to concepts in the glossary (Supplement A in the Supplemental Material). The expert used all the experimental design concepts listed in Table 7 in his response to the Neuron Assessment. This included knowledge of treatment variables, which he had failed to mention before exposure to the assessment.

The Neuron Assessment provided evidence for knowledge of several experimental design concepts for students (Figure 1, step 3). The assessment revealed that Daniel showed knowledge of all concepts (Table 7). This was not true for Juan, Eve, and Li Na. In the case of Juan and Eve, the assessment showed no evidence that they understood or had difficulties with concept 12 (variation) or concepts 14–16 (replication, variability, and randomization). In addition, the instrument gave no evidence that Juan understood or had difficulties with concepts 7 (control variable), 10 and 11 (controlling outside variables, confounding variables), 17 (representative sample), and 18 (scope of inference). Furthermore, for Li Na, the instrument provided no evidence related to concepts 18 and 19 (scope of inference and cause and effect). Finally, across all four students, good evidence was found that the instrument revealed knowledge or difficulties with concepts 1–6 (neuron, organellar movement, experimental subject, variable, treatment variable, and treatment group), 8 and 9 (control and control group), and 20 (correlations).

To summarize findings from the case of Eve as a specific example, the Neuron Assessment prompted a display of correct reasoning with visualizations across six of the seven RM abilities in Table 6. However, even though the assessment prompted Eve to demonstrate good knowledge of organellar movement and neurons (Table 7, rows 1 and 2), her response to the Neuron Assessment revealed clear evidence of difficulties in 12 of 18 areas of experimental design (Table 7, rows 3–20). Based on additional data analyses to support the findings (tables in Supplement F in the Supplemental Material), the Neuron Assessment revealed knowledge and difficulties related to all five RED Areas of Difficulty (Supplement C in the Supplemental Material).

CONCLUSIONS AND IMPLICATIONS

This study has illustrated steps that can be taken to develop an assessment for experimental design that makes it possible to observe, measure, and interpret student knowledge and difficulties related to areas of concern reported by others. These include difficulties with the concepts related to "control" (Picone *et al.*, 2007; Shi *et al.*, 2011), "variability measures" (Kanari and Millar, 2004; Kuhn and Dean, 2005), and "causal outcomes" (Klahr *et al.*, 1993; Schauble, 1996). In addition, the assessment yielded information about major experimental design areas outlined by our own and other previous research (Sirum and Humburg, 2011; Brownell *et al.*, 2014; Dasgupta *et al.*, 2014; Deane *et al.*, 2014) and also revealed for the first time visual modes of presenting this knowledge. The latter knowledge contributed to modifications that improved the Dasgupta *et al.* (2014) glossary and the RED.

We found that all students in our sample, regardless of their level of experimental design abilities, were uncomfortable with not knowing the right answer to the Neuron Assessment. This suggests that we should give students more practice investigating uncertain ideas—a crucial aspect of scientific research. Because some experiments designed by students might be better capable of revealing new knowledge than others, assessments like this should be useful for classroom discussions.

We present this study as an exploratory study. The case study approach used here is not an appropriate method for validation of the Neuron Assessment, because case studies do not allow for the generalization of inferences. This study has demonstrated, though, that a case study approach that incorporates the three steps presented in detail in this paper can be an effective and useful method for the successful development of an instrument like the Neuron Assessment. A limitation is that the Neuron Assessment is focused on experimental design within the context of a specialized biological situation, so we do not claim it would be transferable as a general assessment tool. In fact, there is a need to more fully explore the degree to which reasoning about experimental design is grounded in situated contexts. In other words, we do not yet know if people who understand experimental design in the subdomain of biology in which they have developed expertise can apply or integrate good experimental design into an unfamiliar area of biology. The possibility that complexity in different subdomains may interfere with reasoning about experiments should be the focus of future research. The methodology described here may be useful for others who design assessments to observe and measure difficulty with reasoning about experiments in complex situations for other biology subdomains. As a future validation step, responses to the assessment could be gathered from other experts and students to establish whether the assessment can result in any false interpretations. Additional work is also needed to make sure the written responses are easy to score and to determine whether experts from other subdomains of biology would agree on the right answers.

In spite of the above limitations, we believe that our three-step process used in this study to develop the Neuron Assessment can be usefully applied to the development of similar assessments in other subdomains of biology. Support for this belief is found in the fact that our development process incorporates all three elements of the assessment triangle framework (cognition, observation, and interpretation) proposed by the NRC (2001). More specifically, regarding the observation and cognition elements, we demonstrate the usefulness of providing a real-life biological research context to probe (observation) thinking and visualizations of experiments (cognition). This study also allows us to begin exploring how well students use representations (interpretation) compared with a scientist in a particular subdomain of biology. This is similar to the work done in the field of chemistry (Kozma *et al.*, 2000). In our study, in response to the Neuron Assessment, the expert scientist graphed data using results from microscopy. Another study in the field of genetics might develop an assessment to observe and measure students' abilities to visualize experimental data as gene sequences. In contrast, an assessment task designed for physiology students might have a

sodium-potassium pump as the context and involve visualizing small interfering RNAs to knock down the expression of the gene sequence to disrupt normal functioning of the pump. Thus, even though the Neuron Assessment is not directly applicable to all areas of biology, we conclude that our findings confirm that our three-step process for assessment development might be usefully applied to assessment development in other subdomains of biology to gather and interpret observations about how well students are designing and visualizing experiments.

ACKNOWLEDGMENTS

We thank Drs. David Eichinger, Jeff Karpicke, and Dennis Minchella for their insightful advice, and our appreciation also goes to Dr. Peter Hollenbeck for sharing knowledge that led to the research problem for the Neuron Assessment. We also thank several anonymous reviewers for useful feedback that helped us greatly improve the manuscript. Members of the Purdue International Biology Education Research Group and the Visualization in Biochemistry Education Research Group are thanked for many fruitful discussions that contributed to the progress of this study. Thanks go to *Nature Reviews Neuroscience* as the original source for Box1, a and b, of the Neuron Assessment (Knott *et al.*, 2008). This figure is used with permission. Other material presented here is based on research supported by the National Science Foundation under grants 0837229 and 1346567. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- American Association for the Advancement of Science (2011). Vision and Change in Undergraduate Biology Education: A Call to Action Washington, DC. <http://visionandchange.org/finalreport> (accessed 15 March 2015).
- Anderson TR, Schönborn KJ, du Plessis L, Gupthar AS, Hull TL (2013). Identifying and developing students' ability to reason with concepts and representations in biology. In: Multiple Representations in Biological Education, Models and Modeling in Science Education Series, vol. 7, ed. DF Treagust and C-Y Tsui, Dordrecht, the Netherlands: Springer, 19–38.
- Brownell SE, Wenderoth MP, Theobald R, Okoroafor N, Koval M, Freeman S, Crowe AJ (2014). How students think about experimental design: novel conceptions revealed by in-class activities. *BioScience* 64, 125–137.
- Dasgupta AP, Anderson T, Pelaez N (2014). Development and validation of a rubric for diagnosing students' experimental design knowledge and difficulties. *CBE Life Sci Educ* 13, 265–284.
- Deane T, Nomme K, Jeffery E, Pollack C, Birol G (2014). Development of the Biological Experimental Design Concept Inventory (BEDCI). *CBE Life Sci Educ* 13, 540–551.
- Harker AR (2009). Full application of the scientific method in an undergraduate teaching laboratory: a reality-based approach to experiential student-directed instruction. *J Coll Sci Teach* 29, 97–100.
- Hiebert SM (2007). Teaching simple experimental design to undergraduates: do your students understand the basics? *Adv Physiol Educ* 31, 82–92.
- Kanari Z, Millar R (2004). Reasoning from data: how students collect and interpret data in science investigations. *J Res Sci Teach* 41, 748–769.
- Klahr D, Fay AL, Dunbar K (1993). Heuristics for scientific experimentation: a developmental study. *Cogn Psychol* 25, 111–146.

- Knott AB, Perkins G, Schwarzenbacher R, Bossy-Wetzel E (2008). Mitochondrial fragmentation in neurodegeneration. *Nat Rev Neurosci* 9, 505–518.
- Kozma R, Chin E, Russell J, Marx N (2000). The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning. *J Learn Sci* 9, 105–143.
- Kuhn D, Dean D (2005). Is developing scientific thinking all about learning to control variables? *Psychol Sci* 16, 866–870.
- Mayer RE, Moreno R (2003). Nine ways to reduce cognitive load in multimedia learning. *Educ Psychol* 38, 43–52.
- Moreno R, Mayer RE (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *Br J Educ Psychol* 91, 358–368.
- National Research Council (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*, Washington, DC: National Academies Press.
- Picone C, Rhode J, Hyatt L, Parshall T (2007). Assessing gains in undergraduate students' abilities to analyze graphical data. *Teach Issues Exp Ecol* 5. www.esa.org/tiee/vol/v5/research/picone/article.html (accessed 20 December 2013).
- Salangam J (2007). The impact of a prelaboratory discussion on non-biology majors abilities to plan scientific inquiry. Master's Thesis, California State University, Fullerton.
- Schauble L (1996). The development of scientific reasoning in knowledge-rich contexts. *Dev Psychol* 32, 102–119.
- Schönborn KJ, Anderson TR (2009). A model of factors determining students' ability to interpret external representations in biochemistry. *Int J Sci Educ* 31, 193–232.
- Shi J, Power JM, Klymkowsky MW (2011). Revealing student thinking about experimental design and the roles of control experiments. *Int J Sch Teach Learn* 5, 1–16.
- Sirum K, Humburg J (2011). The Experimental Design Ability Test (EDAT). *Bioscene* 37, 8–16.
- Stenning K, Oberlander J (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cogn Sci* 19, 97–140.
- Trujillo CM, Anderson TR, Pelaez NJ (2015). A model of how different biology experts explain molecular and cellular mechanisms. *CBE Life Sci Educ* 14, ar20.
- Yin RK (1984). *Case Study Research: Design and Methods*, Newbury Park, CA: Sage.