The Redundancy Principle of Multimedia Learning in a Next Generation Science Classroom: Measuring Learning Outcomes

Robert C. Wallon[1]

University of Illinois, Urbana-Champaign

Presented as a paper at NARST 2015, Chicago, IL

April 13, 2015

[1] Correspondence regarding this paper may be directed to the author via email at rwallon2@illinois.edu

Abstract

Previous research has identified circumstances when eliminating redundant information in instructional multimedia improved learning outcomes in laboratory and workplace settings. The goal of this study is to further clarify the boundaries of the redundancy principle by researching the extent to which it applies in a secondary science classroom context. This study used a pretest-posttest design during the enactment of a curriculum unit in three periods of a high school biology class. Fifty students were tested before and after watching either the redundant or nonredundant version of a video clip and at the conclusion of the curriculum unit. Comparison of student scores showed a redundancy effect on measures of retention but no redundancy effect on measures of transfer. This paper discusses implications of measuring student learning outcomes in authentic classroom settings with instruments modeled after those used in laboratory studies. Future research should explore the applicability of the redundancy principle using more authentic measures of transfer that take into account the social context of the classroom.

Keywords: redundancy principle, multimedia learning, science

The Redundancy Principle of Multimedia Learning in a Next Generation Science Classroom: Measuring Learning Outcomes

Research on multimedia learning over the past several decades has resulted in an extensive body of knowledge about how people learn from words and pictures. Mayer (2009) summarized his work in this area into statements called multimedia learning principles. The redundancy principle is one such principle, and it states that "people learn better from graphics and narration than from graphics, narration, and printed text"(Mayer, 2009, p. 118). The redundancy principle has been supported by research conducted in laboratory settings (e.g., Mayer, Heiser, & Lonn, 2001) and workplace settings (e.g., Kalyuga, Chandler, & Sweller, 1999) that compared learning outcomes from people who used redundant multimedia with learning outcomes from people who used nonredundant multimedia. However, no research has investigated the redundancy principle in classroom settings.

The lack of research on the redundancy principle in classroom settings is a cause for concern for two main reasons. First, the redundancy principle contradicts the common pedagogical practice of presenting the same information simultaneously in multiple formats. Sweller (2005) explains, "It is easy to assume that presenting the same information in multiple forms or presenting additional explanatory information could be advantageous and at worst, will be neutral. Such an assumption ignores what we now know of human cognitive architecture" (p. 166). Second, there is widespread use of and creation of multimedia by teachers. In a recent national survey, 46% of teachers reported that they use the Internet to find videos[2] to include in their curriculum materials (Project Tomorrow, 2014). It has also become relatively common for teachers to create their own educational multimedia. In the aforementioned national survey, 16%

[2] While the term multimedia includes many possible combinations of graphics and text, this study focuses specifically on video multimedia.

of teachers reported that they regularly create educational videos for their students (Project Tomorrow, 2014). If the redundancy principle applies to classroom settings, then teachers should consider it while selecting or creating multimedia materials for their students.

Mayer (2010) described a model for learning from multimedia that is called the *cognitive theory of multimedia learning* (*CTML*). The CTML accounts for the type of learning that enables people to form a mental model of a concept that they can manipulate such as a causal system. The CTML is based on three assumptions. First, the dual channels assumption states that humans have two distinct channels for processing auditory information and visual information. Second, the limited capacity assumption uses cognitive load theory to suggest that each channel is limited in the amount of information that it can process in a given time. Third, the active processing assumption states that "humans engage in active learning by attending to relevant incoming information, organizing selected information into coherent mental representations, and integrating mental representations with other knowledge" (Mayer, 2009, p. 63).

The CTML provides an explanation for how redundancy can lower learning outcomes. A learner watching a redundant video dedicates cognitive resources to processing graphics, narration, and printed text. Therefore relatively fewer cognitive resources are available in the visual channel when compared to a learner watching a nonredundant video who dedicates cognitive resources only to processing graphics and narration. Redundant multimedia can overload working memory and consequently harm integration to long-term memory.

Many of the previous research studies have used experimental designs in order to make causal claims about the effects of the redundancy principle. However, the conditions needed to make causal claims have also limited the environments in which the redundancy principle has been studied, favoring lab settings over classroom settings. Harskamp, Mayer, and Suhre (2007) studied a different multimedia learning principle in a classroom setting and succinctly described the implications of few studies in these contexts, explaining, "if design principles can be demonstrated in controlled lab environments but cannot be demonstrated in authentic school environments with students, their practical value for education and their theoretical value for multimedia learning are limited" (p. 446). To address the concern about few studies on multimedia learning principles in school settings, the research question investigated in this study is: To what extent does the redundancy principle of multimedia learning apply in a high school biology classroom?

By studying the extent to which the redundancy principle applies to a high school biology classroom environment, this study can further clarify boundary conditions of contexts in which the redundancy principle can be usefully applied.

Procedure

This study addresses the research question with a pretest-posttest design. Consistent with quantitative approaches, this design was appropriate because the objective of the study was to test existing theory (i.e., CTML).

The participating teacher was recruited from those who had received formal training on the *What can I learn from worms? Regeneration, stem cells, and models* curriculum unit (Project NEURON, 2013). The curriculum unit is described as follows from the Project NEURON web site:

> This unit is grounded in a cost-effective and student-driven investigation that teachers love! Intrigued by the fascinating behavior of regeneration, students examine the process of cellular division and visualize the process of planarian flatworm regeneration with fluorescent images from the University of Illinois. While students collect and analyze

> their own experimental data, students use computer models to simulate how DNA and protein affect behavior and explore applications of what they've learned to disease and stem cell research. (Project NEURON, 2013)

The curriculum unit was enacted over sixteen days of instruction. A timeline of the enactment can be found in *Table 1*, which has the aspects of the unit relevant to this study in bold.

Table 1. Timeline of instructional days in curriculum unit.

Instructional Day	Summary of Lesson
1	Introduction to regeneration discussion and jigsaw readings
2	Mini-lecture on planarian anatomy and planarian observations
3	Planarian cutting
4	Planarian observations, Day 1
5	Planarian observations, Day 2, and Neoblast Division packet
6	Planarian observations, Day 5, and Neoblast Division packet
7	**Pre-test for video,** Planarian observations, Day 6**,** and Cell cycle modeling activity
8	**BrdU Video, Post-test for video,** Planarian observations, Day 7**,** and BrdU packet
9	Planarian observations, Day 8, and BrdU packet
10	Planarian observations, last day, BrdU packet, RNAi reading
11	Lecture on constructing scientific explanations, Notes on RNAi, NetLogo RNAi modeling activity
12	NetLogo RNAi modeling activity
13	NetLogo RNAi modeling activity, Poster project
14	**Delayed post-test for video**, Letter to a family member explaining future of regenerative medicine
15	Planarian posters, Follow up interviews were conducted with 1 group from each class
16	Unit Test

The curriculum unit was selected because it included a video that explains a molecular biology technique that enables the visualization of stem cell regeneration in planarians, a type of flatworm commonly used in biology research. The teacher who agreed to participate in this study taught three periods of an elective second-level high school biology course. The study took place at a mid-sized high school in a town located near a small Midwestern city. The majority of the school's student population was white, with no federal race and ethnicity subcategory larger than 5%. The school low-income level was around 25%. The study took place in the context of an entire problem-based curriculum unit taught by the same teacher. Given that the class was an elective, students in the classes ranged from grade ten to grade twelve.

A quasi-experimental design was used in order to investigate the research question in the regular classroom environment. One class period was assigned to the nonredundant condition (n=20, 11 males and 9 females), and two class periods were assigned to the redundant condition (n=30, 15 males and 15 females). From consulting with the teacher there was no reason to believe that students in different classes differed from each other for systematic reasons (e.g., no students were ELLs or had IEPs).

Students took a pretest on instructional day seven in order to measure their prior knowledge of relevant biology terms. Students were allowed ten minutes to take the pretest, and all students finished within the allocated time. The pre-test (*Appendix A*) consisted of one page where students rated their understanding of six terms and an additional page where students explained the six terms.

During the next class period, instructional day eight, students watched the video clip (available online at https://neuron.illinois.edu/videos/brdu) that explained the molecular biology technique for fluorescently labeling new cells. The video clip was presented in either a redundant or nonredundant version. Both versions were identical except the redundant version displayed captioned text at the bottom of the screen that was redundant with the spoken audio (*Figure 1*).

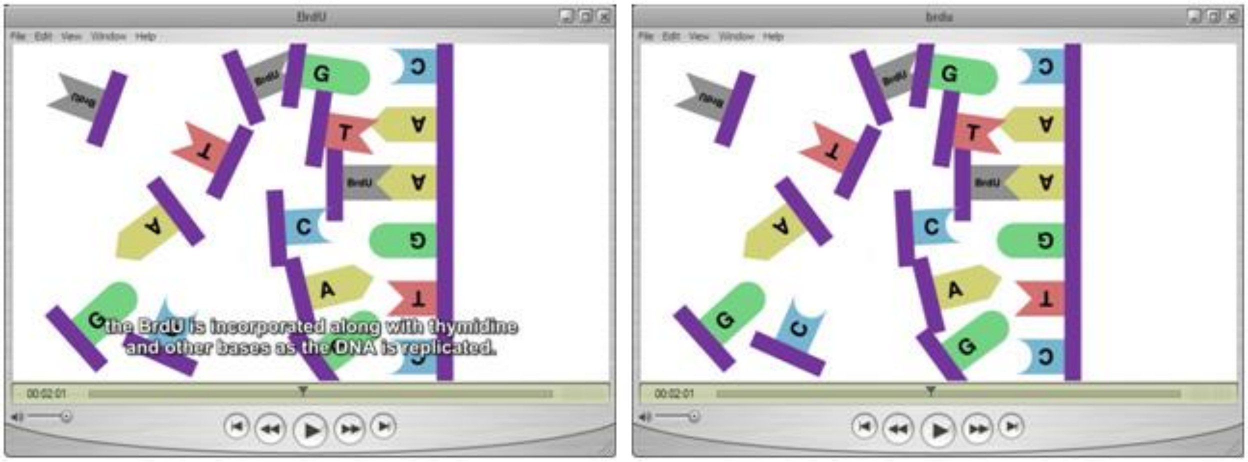

Figure 1. Screenshots of redundant video (left) and nonredundant video (right).

The video clip lasted approximately five minutes. Immediately after watching the video clip students completed the posttest (*Appendix B*). Students were allowed fifteen minutes to take the posttest, and all students finished within the allocated time. Mayer (2009) articulated that the two goals for multimedia learning are remembering and understanding. Remembering is the "ability to reproduce or recognize presented material" while understanding is the "ability to use presented materials in novel situations" (p. 20). Remembering can be measured with retention tests, and understanding can be measured with transfer tests (Mayer, 2009). The first two pages of the posttest were identical to the pretest. The second page was administered to measure students’ retention from the video. An additional third page included five open-ended questions that were administered to measure students’ transfer from the video. Question number one was excluded from analysis after it became apparent based on student responses that it was being answered in a way that measured retention more than transfer.

Students also completed the delayed posttest near the end of the curriculum unit on instructional day fourteen. Students were allowed fifteen minutes to take the delayed posttest, and all students finished within the allocated time. The delayed posttest was identical to the posttest.

Field notes and audio recordings were taken during classroom observations, but the test scores served as the primary data source for answering the research question addressed in this study. Student responses to 67% of retention items on pretests, posttests, and delayed posttests were independently scored by the author and a colleague using scoring guides (*Appendix C*) that were iteratively developed for each item. Scores were compared and all discrepancies were resolved by discussion. The author scored remaining items. Each of the six retention items was scored up to a maximum of five points, with a maximum total score of thirty points. Transfer items on post-tests and delayed post-tests were scored by adding up the number of acceptable responses, up to a maximum of five points per item for a maximum total score of twenty points.

Analyses and Findings

Analysis of variance (ANOVA) was run using SPSS to determine if there were statistically significant differences between scores for students in the redundant and nonredundant conditions on the pretests. Analysis of covariance (ANCOVA) was run using SPSS to determine if there were statistically significant differences between scores for students in different conditions on posttests and delayed posttests. Results are summarized in *Figure 2*. Effect sizes are reported as eta squared, η^2, for the ANOVA and partial eta squared, η_p^2, for the ANCOVAS. These effect size measures should be interpreted as the proportion of variance in the dependent variable (test scores) accounted for by the independent variable (redundant or nonredundant condition), which is calculated as a ratio of the sum of squares between groups to the total sum of squares[3] (Gravetter & Wallnau, 2011). Generally speaking, an η^2 of .02 should be interpreted as a small effect size, an η^2 of .13 should be interpreted as a medium effect size, and an η^2 of .26 should be interpreted as a large effect size (Cohen 1992). All tests were run at a confidence level of .05.

An ANOVA indicated that there was no statistically significant difference between the two conditions on the pretest, $F(1,48) = 0.491$, $p = .487$, $\eta^2 = 0.01$. An ANCOVA indicated that there was a statistically significant difference between the two conditions on posttest retention items favoring the nonredundant condition, $F(1,47) = 6.353$, $p = .015$, $\eta_p^2 = 0.119$. An ANCOVA indicated that there was no statistically significant difference between the two conditions on posttest transfer items, $F(1,47) = 1.765$, $p = .190$, $\eta_p^2 = 0.036$. An ANCOVA indicated that there was a statistically significant difference between the two conditions on delayed posttest retention items favoring the nonredundant condition, $F(1,47) = 5.089$, $p = .029$, $\eta_p^2 = 0.098$. An ANCOVA indicated that there was no statistically significant difference between the two treatments on delayed posttest transfer items, $F(1,47) = 3.558$, $p = .065$, $\eta_p^2 = 0.070$.

[3] Partial eta squared, η_p^2, also includes an error term based on additional predictors.

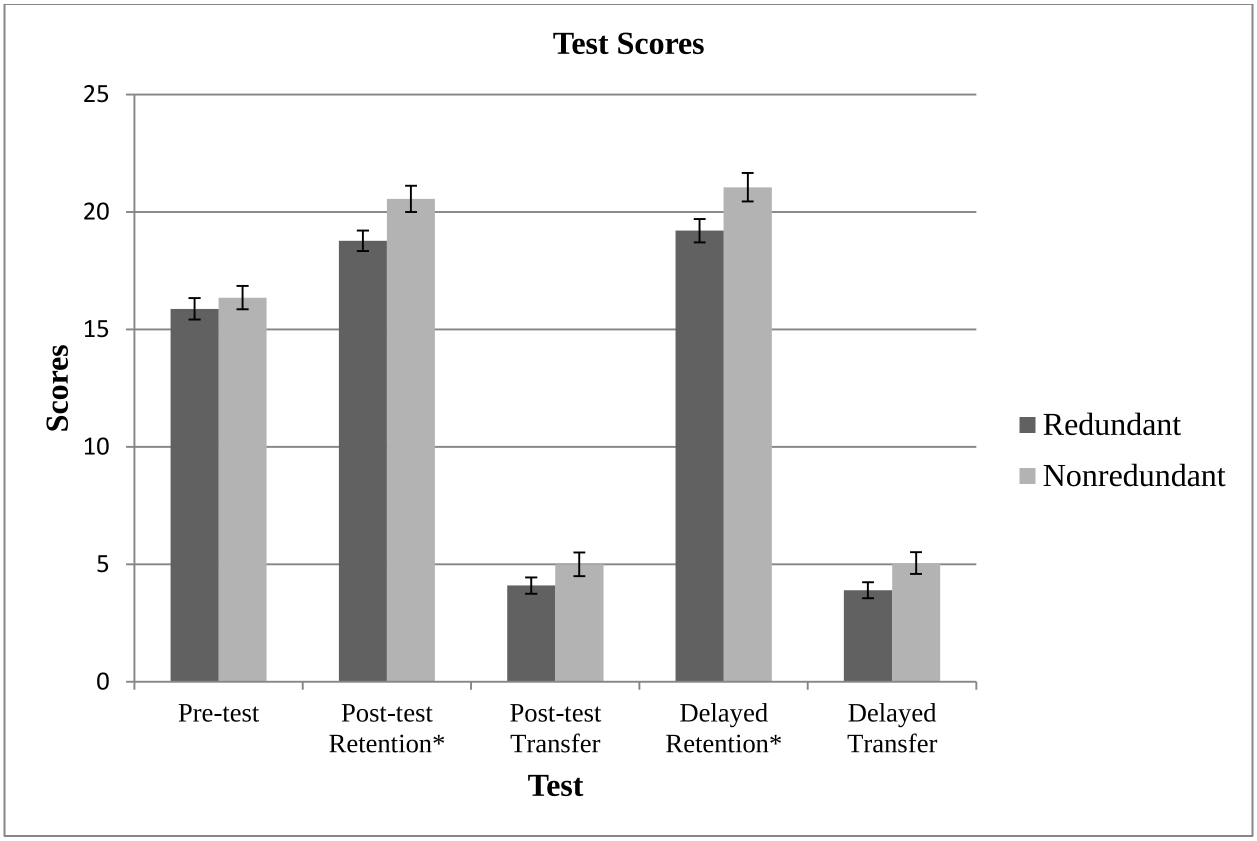

Figure 2. Mean test scores for redundant group (N=30) and nonredundant group (N=20) with standard error bars. * indicates statistically significant difference with $p < .05$.

Discussion

The test score results showed that the redundancy principle applied to a limited extent in this high school biology classroom. The redundancy principle applied to retention test items but not to transfer test items. Comparison of scores on the delayed posttest shows that the redundancy principle for retention items persisted over time and after additional instruction.

The results of this study are different from many studies of the redundancy principle, which have shown a redundancy effect either for both retention tests and transfer tests (e.g., Jamet & Le Bohec, 2007; Mayer, Heiser, & Lonn, 2001) or a redundancy effect for transfer tests but not for retention tests (e.g., Craig, Gholson, & Driscoll, 2002). The results for this study may have differed because of high student background knowledge, a lack of attention to redundant text, or student motivation while taking the tests. Each of these possibilities is discussed in more detail below.

The study took place in a second-level biology course so higher levels of student background knowledge may have affected the results. A previous study found that whether or not text is considered redundant may depend on the prior knowledge and experience of learners. The study found that the advantage of applying the redundancy principle disappeared for more expert learners (Kalyuga, Chandler, & Sweller, 2000). The next possibility explains why that may be the case.

Another possibility is that students in the redundant group may not have attended to the text displayed on the bottom of the screen. Instead they may have relied only on the audio narration, which was the case in another study with students in an immersive virtual reality environment (Moreno & Mayer, 2002). During classroom observations the author could see that all students were looking at the screen at the front of the room when the video was projected, but there was no way to discern whether or not students were attending to the text on the screen. Future research with eye tracking may be used to investigate this assumption, but it may be difficult to implement unobtrusively in classroom settings. Less obtrusive measures such as asking students if they noticed the text may be used, but students may not be consciously aware of their attention to different parts of the video. A connection between this possibility and the previous possibility is that students with higher levels of background knowledge may focus on the graphical elements and not the textual elements of the video. That is, higher expertise is associated with attending to relevant features.

An additional possibility is concerned with student motivation while taking the tests. While all students provided responses and there were no overt signs of students not applying themselves during testing, no students used the back of the test papers to provide additional responses as was encouraged in the directions. This is a concern because transfer test items were scored by counting a total number of acceptable responses, consistent with previous research on the redundancy principle. If students did not consider and list every possibility they could think of, then a redundancy effect may not be measured even if there was a difference between conditions.

Caution should be taken when interpreting the results reported in this study to avoid generalizing beyond the circumstances described in this paper. The context for this study resulted in a tradeoff of limitations. It was not practical to assign students to groups randomly, and students were nested within classes. Therefore statistical analyses factored in pretest scores as a covariate to attempt to account for any differences in prior knowledge. A classroom setting also afforded multiple exposures to content and the opportunity for students to ask questions of the teacher and of one another to help improve their understanding. These variables would likely be controlled for in a lab study, but they would be common features in most classroom environments. Therefore the implications of this study may have more relevance for teachers because it took place in a classroom. This study showed that the redundancy principle applied to a limited extent in a high school classroom. Therefore it may be useful for teachers to apply the redundancy principle when selecting and designing instructional multimedia.

Future research should seek to further clarify the conditions when the redundancy principle affects learning outcomes in classroom settings. Additionally, this study has addressed the redundancy principle from an individual cognitive perspective, consistent with previous studies. While this has allowed for incremental progress is defining boundaries for the redundancy principle, it has neglected to examine the social factors that also influence learning in classroom contexts. Future studies of the redundancy principle should expand to also include analysis of social learning processes. Methods of discourse analysis such as those found in systemic functional linguistics can potentially provide a window into differences in *how* students conceptualize ideas when they are learned in redundant or nonredundant conditions. These types of studies may be useful for evaluating benefits of applying the redundancy principle in social classroom settings, and they would afford the study of students engaged in more authentic problem solving than answering test questions.

References

Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*, 428-434.

Gravetter, F., & Wallnau, L. (2011). *Essentials of statistics for the behavioral sciences: Seventh edition*. Wadsworth Cengage Learning.

Harskamp, E. G., Mayer, R. E., & Suhre, C. (2007). Does the modality principle for multimedia learning apply to science classrooms? *Learning and Instruction, 17*, 465-477. doi: 10.1016/j.learninstruc.2007.09.010

Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology, 32*, 588-598.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351-371.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, *92*, 126-136. doi: 10.1037//0022-0663.92.1.126

Mayer, R. E. (2009). *Multimedia learning: Second edition*. New York, NY: Cambridge University Press.

Mayer, R. E. (2010). Fostering scientific reasoning with multimedia instruction. In H. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 160-175). New York, NY: Guilford Press.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*, 187-198. doi: 10.1037/0022-0663.93.1.187

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, *94*, 598-610. doi: 10.1037//0022-0663.94.3.598

Project NEURON. (2013). What can I learn from worms? Regeneration, stem cells, and models. Retrieved from http://neuron.illinois.edu/units/what-can-i-learn-from-worms

Project Tomorrow. (2014). Speak up 2013 national research project findings: A second year review of flipped learning. Retrieved from http://www.tomorrow.org/speakup/pdfs/SU13SurveyResultsFlippedLearning.pdf

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 159-168). New York, NY: Cambridge University Press.

APPENDIX A: Pre-test

Pre Test for “How do scientists visualize the regeneration of cells?” Video

Directions: Rate how familiar you are with each of the listed terms by placing a check in each row of the table below. The rating scale ranges from 1 to 5. Checking 1 means you are not at all familiar (have never heard the word), while checking a 5 means you are very familiar (understand the word and can use the concept in your thinking).

	Not at all familiar		*Somewhat familiar*		*Very familiar*
Term	**1**	**2**	**3**	**4**	**5**
Stem cells					
Regeneration					
Fluorescence					
BrdU					
DNA					
Antibodies					

Directions: In the space below, define each of the following terms in your own words to the best of your ability.

Stem cells:

Regeneration:

Fluorescence:

BrdU:

DNA:

Antibodies:

APPENDIX B: Post-test and Delayed Post-test

Post Test for "How do scientists visualize the regeneration of cells?" Video

Directions: Rate how familiar you are with each of the listed terms by placing a check in each row of the table below. The rating scale ranges from 1 to 5. Checking 1 means you are not at all familiar (have never heard the word), while checking a 5 means you are very familiar (understand the word and can use the concept in your thinking).

	Not at all familiar		*Somewhat familiar*		*Very familiar*
Term	**1**	**2**	**3**	**4**	**5**
Stem cells					
Regeneration					
Fluorescence					
BrdU					
DNA					
Antibodies					

Directions: In the space below, define each of the following terms in your own words to the best of your ability.

Stem cells:

Regeneration:

Fluorescence:

BrdU:

DNA:

Antibodies:

Directions: Based on watching the video ("How do scientists visualize the regeneration of cells?") please answer the following questions as thoroughly as possible. Feel free to use the back of the paper if you need additional space.

1. Please write down a list of as many steps you remember of how scientists visualize the regeneration of cells.

2. What could you do to increase the intensity of fluorescence while visualizing cells?

3. Suppose you do not see any fluorescence when you go to visualize cells. List as many ideas as you can think of for why you might not see any fluorescence.

4. What is the purpose for mixing BrdU with pureed beef liver?

5. What causes stem cells to fluoresce or glow green?

APPENDIX C: Scoring Rubrics

Scoring Guide for Student Definitions of “BrdU”

The criteria outlined below establish the minimum components required to assign a score to each response. Adding positive features of responses cannot raise a score if the criteria outlined are otherwise not met (in many cases additional components of a response would qualify it for a higher score based on the criteria). However, negative features of responses such as misconceptions can lower a score by **one point** (e.g., If a response would have been scored a 5, except BrdU was referred to as “a cell,” then the response should be recorded as a 4.). The score is lowered one point regardless of the number of misconceptions that are included. When assigning scores, keep track of any point deductions. Do not penalize students for personification (e.g., saying “your” rather than relating to planarians). Do not penalize students for misspellings that do not affect meaning (e.g., spelling protein as “protien”).

Please read through the criteria for each score several times before beginning to score student responses.

Score of 5 = Response references both the structure of BrdU and the purpose for using BrdU.

Examples of acceptable references to structure can include one or more of the following: replacement for thymine, replacement for thymidine, [chemical, base, nucleotide, or nucleoside] integrated in DNA (binds in DNA or attaches to DNA are also acceptable).

Examples of acceptable references to purpose can include one or more of the following: label stem cells, locate stem cells, visualize or see stem cells, track where stem cells go, allow scientists to visualize or watch regeneration, allows antibodies to target stem cells.

Score of 4 = Response only references either the structure *or* purpose of BrdU.

See above for examples.

Score of 3 = Response is vague or includes limited details. These responses are generally brief and may have been scored higher if the student elaborated.

Examples: A chemical. Observe changes in planarians. Gives a better understanding of stem cells. Scientists put it in cells to study regeneration. It goes into their stem cells. Planarians eat it. They feed it to planarians. Attaches to antibodies. Bromodeoxyuridine.

Score of 2 = Response includes mostly misconceptions.

Examples include: a cell, an antibiotic, a protein, an amino acid, a sequence or strip of DNA code, a gene, a dye, is food or nutrients for planarians, BrdU is injected, BrdU *attaches* on the *outside* of the cell, attaches to RNA, is fluorescent or glows green, freezes or kills planarians, directs stem cells where to go, stops the growth of proteins, changes your proteins.

Score of 1 = Response is blank, or the student explicitly states “I don’t know.” Responses that do not include any relevant details should also be scored in this category.

Scoring Guide for Student Definitions of “Antibodies”

The criteria outlined below establish the minimum components required to assign a score to each response. Adding positive features of responses cannot raise a score if the criteria outlined are otherwise not met (in many cases additional components of a response would qualify it for a higher score based on the criteria). However, negative features of responses such as misconceptions can lower a score by **one point** (e.g., If a response would have been scored a 5, except antibodies were referred to as “a cell,” then the response should be recorded as a 4.). The score is lowered one point regardless of the number of misconceptions that are included. When assigning scores, keep track of any point deductions. Do not penalize students for personification (e.g., saying “your” rather than relating to planarians). Do not penalize students for misspellings that do not affect meaning (e.g., spelling protein as “protien”).

Please read through the criteria for each score several times before beginning to score student responses.

Score of 5 = Response references both the structure of antibodies and the function of antibodies.

Examples of acceptable references to structure can include one or more of the following: Y-shaped protein, U-shaped protein, chemical created by the immune system.

Examples of acceptable references to function can include one or more of the following: specifically targets/attaches to/"grab"/"hold on to" cells, [proteins, foreign matter, bacteria, viruses, etc.], attach/match up/link to/connect to/contact BrdU, attach fluorescent dyes to BrdU.

Score of 4 = Response only references either the structure *or* function of antibodies.

See above for examples.

Score of 3 = Response is vague or includes limited details. These responses are generally brief and may have been scored higher if the student elaborated.

Examples: helps fight [cells that don’t belong, foreign matter, bacteria, viruses, etc.], “thing”, help with proteins, fight off disease, good things that prevent disease, Y figures, make stuff glow, chemicals,

Score of 2 = Response includes mostly misconceptions.

Examples include: a cell, an antibiotic, medicine, bacteria, microbes, organisms, produced from the planarians.

Score of 1 = Response is blank, or the student explicitly states “I don’t know.” Responses that do not include any relevant details should also be scored in this category.

Scoring Guide for Student Definitions of “Stem Cells”

The criteria outlined below establish the minimum components required to assign a score to each response. Adding positive features of responses cannot raise a score if the criteria outlined are otherwise not met (in many cases additional components of a response would qualify it for a higher score based on the criteria). However, negative features of responses such as misconceptions can lower a score by **one point** (e.g., If a response would have been scored a 5, except for a misconception, then the response should be recorded as a 4.). The score is lowered one point regardless of the number of misconceptions that are included. When assigning scores, keep track of any point deductions. Do not penalize students for personification (e.g., saying “your” rather than relating to planarians). Do not penalize students for misspellings that do not affect meaning (e.g., spelling protein as “protien”).

Please read through the criteria for each score several times before beginning to score student responses.

Score of 5 = Response references both the description of Stem Cells and the function of Stem Cells.

Examples of acceptable references to description can include one or more of the following: unspecialized/undesignated cells; cells with no specific job; “blank” cells; cells without a defined role;

Examples of acceptable references to function can include one or more of the following: help in regeneration/regrowth of lost limbs; help repair damaged tissues; can potentially differentiate into/become/turn into other types of cells;

Score of 4 = Response only references either the description *or* purpose of Stem Cells.

See above for examples.

Score of 3 = Response is vague or includes limited details. These responses are generally brief and may have been scored higher if the student elaborated.

Examples: cells that can turn into other things; cells that regenerate; special cells; cells that have yet to receive a purpose (function would be appropriate); cells that help the growth of specific parts;

Score of 2 = Response includes mostly misconceptions.

Examples include: can become anything; cells *with* a defined role; can duplicate cells nearby; helps cancer; make copies of cells; cells all connected in some way; cells that hold water to support plant stems;

Score of 1 = Response is blank, or the student explicitly states “I don’t know.” Responses that do not include any relevant details should also be scored in this category.

Scoring Guide for Student Definitions of "DNA"

The criteria outlined below establish the minimum components required to assign a score to each response. Adding positive features of responses cannot raise a score if the criteria outlined are otherwise not met (in many cases additional components of a response would qualify it for a higher score based on the criteria). However, negative features of responses such as misconceptions can lower a score by **one point** (e.g., If a response would have been scored a 5, except DNA was referred to as "a cell," then the response should be recorded as a 4.). The score is lowered one point regardless of the number of misconceptions that are included. When assigning scores, keep track of any point deductions. Do not penalize students for personification (e.g., saying "your" rather than relating to planarians). Do not penalize students for misspellings that do not affect meaning (e.g., spelling protein as "protien").

Please read through the criteria for each score several times before beginning to score student responses.

Score of 5 = Response references both the structure of DNA and the function of DNA.

Examples of acceptable references to structure can include one or more of the following: molecule/chemical (in cells); nucleic acid; strand of nucleotides/acids; sequence of genes;

Examples of acceptable references to purpose can include one or more of the following: provides genetic instructions for creating protein products; codes for proteins/genes/RNA; influences traits/characteristics/features such as some appearances and some behaviors

Score of 4 = Response only references either the structure *or* function of DNA.

See above for examples.

Score of 3 = Response is vague or includes limited details. These responses are generally brief and may have been scored higher if the student elaborated.

Examples: Genetic code/ material/information/makeup/blueprints; **What your genes are made of**; In every living thing; makes you, you/ makes us, us; makes everyone different/makes an organism unique; double stranded; ATGC; Deoxyribo**nucleic acid**; Tells our cells what to do; The data of a given organism; Where cell info is stored; Contains traits;

Score of 2 = Response includes mostly misconceptions.

Examples include: makes up everything; describes everything about who we are; makes up a human; make up your body; amino acids; proteins; directed by RNA; mixed with RNA; doubled during mitosis; Cells divide then new DNA is created; contains uracil; Contains protons, neutrons, electrons; contains neurons;

Score of 1 = Response is blank, or the student explicitly states "I don't know." Responses that do not include any relevant details should also be scored in this category.

Scoring Guide for Student Definitions of “Fluorescence”

The criteria outlined below establish the minimum components required to assign a score to each response. Adding positive features of responses cannot raise a score *if the criteria outlined are otherwise not met* (in many cases additional components of a response would qualify it for a higher score based on the criteria). However, negative features of responses (misconceptions from the score=2 category) can lower a score by **one point** (e.g., If a response would have been scored a 5, except there was a misconception, then the response should be recorded as a 4.). The score is lowered one point regardless of the number of misconceptions that are included. When assigning scores, keep track of any point deductions in the right column of the scoring sheet. Do not penalize students for personification (e.g., saying “your” rather than relating to planarians). Do not penalize students for misspellings that do not affect meaning (e.g., spelling protein as “protien”).

Please read through the criteria for each score several times before beginning to score student responses.

Score of 5 = Response references both the description of Fluorescence and the purpose for using Fluorescence.

Examples of acceptable references to description can include one or more of the following: giving off light/glowing green under a specialized/fluorescent/UV light/microscope;

Examples of acceptable references to purpose can include one or more of the following: used to locate stem cells/regeneration/mitosis/BrdU

Score of 4 = Response only references either the description *or* purpose of Fluorescence.

See above for examples.

Score of 3 = Response is vague or includes limited details. These responses are generally brief and may have been scored higher if the student elaborated.

Examples: dye; glowing; light; green pigment; help see planarians; determine location of cells

Score of 2 = Response includes mostly misconceptions for how Fluorescence is used in this context.

Examples include: light bulb; light fixture; black light; BrdU; effect of BrdU; food for planarians; changes DNA sequence; fluoride; a virus; unresponsive cell;

Score of 1 = Response is blank, or the student explicitly states “I don’t know.” Responses that do not include any relevant details should also be scored in this category.

Scoring Guide for Student Definitions of "Regeneration"

The criteria outlined below establish the minimum components required to assign a score to each response. Adding positive features of responses cannot raise a score *if the criteria outlined are otherwise not met* (in many cases additional components of a response would qualify it for a higher score based on the criteria). However, negative features of responses (misconceptions from the score=2 category) can lower a score by **one point** (e.g., If a response would have been scored a 5, except regeneration was referred to as "initial growth of an organism," then the response should be recorded as a 4.). The score is lowered one point regardless of the number of misconceptions that are included. When assigning scores, keep track of any point deductions in the right column of the scoring sheet. Do not penalize students for personification (e.g., saying "your" rather than relating to planarians). Do not penalize students for misspellings that do not affect meaning (e.g., spelling protein as "protien").

Please read through the criteria for each score several times before beginning to score student responses.

Score of 5 = Response references both the description of Regeneration and the mechanism of Regeneration.

Examples of acceptable references to description can include one or more of the following: process in which damaged tissue regrows

Examples of acceptable references to mechanism can include one or more of the following: results from the activity of stem cells

Score of 4 = Response only references either the description *or* mechanism of Regeneration.

See above for examples.

Score of 3 = Response is vague or includes limited details. These responses are generally brief and may have been scored higher if the student elaborated.

Examples: production of new cells; replacing old cells; process of new body parts forming; when cells form into something missing; bringing new life into something already dead; mitosis; to generate something again; bringing new life to something dead;

Score of 2 = Response includes mostly misconceptions.

Examples include: initial growth of organism

Score of 1 = Response is blank, or the student explicitly states "I don't know." Responses that do not include any relevant details should also be scored in this category.