McLaren, B.M., Timms, M., Weihnacht, D., & Brenner, D. (2012). Exploring the assistance dilemma in an inquiry learning environment for evolution theory. In the *Proceedings of the Workshop on Intelligent Support for Exploratory Environments 2012: Exploring, Collaborating and Learning Together* at the *11th International Conference on Intelligent Tutoring Systems* (ITS 2012).

Exploring the Assistance Dilemma in an Inquiry Learning Environment for Evolution Theory

Bruce M. McLaren[1], Michael Timms[2], Doug Weihnacht[3], Daniel Brenner[4]

[1] Carnegie Mellon University, Pittsburgh, PA, U.S.A.
[2] Australian Council for Educational Research, Melbourne, Australia
[3] MW Productions, San Francisco, California, U.S.A.
[4] WestEd, San Francisco, California, U.S.A.
{bmclaren@cs.cmu.edu}

Abstract. A central question to learning science is: How much assistance is the right amount to provide to students as they learn with educational technology? Providing students with help allows them to proceed when they are struggling, yet can depress their motivation to learn on their own. Assistance withholding, on the other hand, encourages students to try to learn for themselves, yet can also lead to frustration when they are stuck. We are investigating this question in a project in the area of inquiry learning in science. We have developed a web-based computer program, *Voyage to Galapagos (VTG)*, which helps individual students "follow" the steps of Darwin through a simulation of the Galapagos Islands, guiding the student toward learning the theories of natural selection and evolution. Students are encouraged to explore the islands, take pictures of finches and iguanas, evaluate the animals' characteristics, and use scientific methodology and analysis to "discover" evolution as they explore the simulated Galapagos Islands. We have designed a study in which we will examine five levels of assistance: (1) no support, (2) error flagging only, (3) error flagging and text feedback on errors, (4) error flagging, text feedback on errors, and hints, and (5) preemptive hints with error flagging, error feedback, and hints. Middle and high school students will participate in the study, which will take place later in 2012. In this paper we discuss the design of the software and our plans for varying assistance in the experiment.

Keywords: the assistance dilemma, inquiry learning, science learning

1 Introduction

A key problem in the learning sciences and educational technology is the *assistance dilemma:* How much assistance is the right amount to provide to students as they learn with educational technology? (Koedinger & Aleven, 2007). While past research with, for instance, inquiry learning environments clearly points toward *some* guidance being necessary (Geier et al 2008) it doesn't fully answer the assistance dilemma (which has also been investigated under the guise of "desirable difficulty" (Schmidt & Bjork, 1992) and "productive failure" (Kapur, 2009). Essentially the issue is to find the right balance between, on the one hand, full support and, on the other hand, allowing students to make their own decisions and, at times, mistakes. There are benefits and costs associated with both ends of this spectrum. *Assistance giving* allows students to move forward when they are struggling and experience success, yet can lead to shallow learning and the lack of motivation to learn on their own. On the other hand, *assistance withholding* encourages students to think and learn for themselves, yet can lead to frustration and wasted time when students are unsure of what to do. Advocates of direct instruction point to the many studies that show the advantages of assistance giving (Kirschner, Sweller, & Clark, 2006; Mayer, 2004), but this still does not address the subtlety of exactly how and when instruction should be made available, particularly in light of differences in domains and learners (Klahr, 2009).

Research in the area of scientific inquiry learning, where students tackle non-trivial scientific problems by investigating, experimenting, and exploring in relatively wide-open problem spaces, has provided various results about how different types of guidance support students. Researchers have built on inquiry learning theory (Edelson, 2001; Quintana et al., 2004) and have developed and experimented with simulations, cognitive tools, and microworlds to support inquiry learning in science. Systems of this kind that have demonstrated learning benefits have included BGUILE (Sandoval & Reiser, 2004), the WISE system (Linn & Hsi, 2000; Slotta, 2004), the Co-Lab collaborative learning system (van Joolingen et al., 2005), and a chemistry virtual laboratory (Tsovaltzi et al., 2009). A large study by Geier et al (2008) that involved over 1,800 middle-school students in the experimental condition versus more than 17,000 students in the control, showed that the students who were given scaffolded tools for performing inquiry learning exercises (in earth, physical and life science) did significantly better on standardized exams than students who did not use the tools.

Thus, there is some evidence that supporting and guiding students in inquiry learning is beneficial. Yet questions still remain: How much support is the right amount? How should assistance vary according to different levels of prior knowledge? To explore these questions we have developed (and continue to develop) a web-based inquiry learning system called *Voyage to Galapagos (VTG)*[1] and will experiment with the software in a systematic manner intended to uncover how much help is necessary for students to learn about the theories of natural selection and evolution. *Voyage to Galapagos* is software that guides students through a simulation of Darwin's journey through the Galapagos

---

[1] *Voyage to Galapagos* was originally designed and developed as a non-web-based program by the third author of this paper, Weihnacht, under National Science Foundation Award # 9618014.

Islands, where Darwin collected data and made observations that helped him develop his theories. The program provides students with the opportunity to do simulated science field work in Galapagos, including data collection and data analysis during investigation of the key biological principles of variation, function, and natural selection.

In typical inquiry learning fashion, the *VTG* program also provides a wide range of actions that a student can take. For instance, as they travel on the virtual paths of individual islands, students can snap pictures of a variety of animals, some of which are relevant to the analysis and understanding of evolution, and some of which are not relevant. This variety of action implies that there are also many possibilities to guide – or not guide – students as they learn and work through the program. Such variety also means that *VTG* is a rich environment to experiment with the assistance dilemma and different amounts and types of guidance. In our planned study of the assistance dilemma, which will be conducted later this year, we will look at two variables of assistance: *frequency of intervention* (values: never, when student is struggling, always) and *level of support* (values: error flagging, text feedback on errors, and hints). This leads to a 3 x 3 matrix that we have reduced to five distinct levels of assistance (and conditions) that we will experiment with: (1) no support, (2) error flagging only, (3) error flagging and text feedback on errors, (4) error flagging, text feedback on errors, and hints, and (5) preemptive hints with error flagging, error feedback, and hints. We will randomly assign between 300 and 450 students to these five conditions and run an experiment in which we will compare conditions and determine which level of assistance leads to the best learning outcomes, both overall and per different levels of prior knowledge.

2 Misconceptions about the Theories of Natural Selection and Evolution

Misconceptions that students have about evolution originate from multiple sources, all of which are related to prior knowledge, beliefs, and conceptions about the world (Alters & Nelson, 2002):

1. *From-experience misconceptions* – Misconceptions that arise from the everyday experience of students. For example, "mutations" are always detrimental to the fitness and quality of an organism (since the word "mutation" in everyday use typically implies an unwanted outcome)
2. *Self-constructed misconceptions* – Misconceptions that arise from students trying to incorporate new knowledge into an already incorrect concept. For example, students who think that evolution is somehow "progressive", always moving toward more "positive" variations.
3. *Taught-and-learned misconceptions* – Misconceptions that arise from informally learned and unscientific "facts." For example, watching movies with dinosaurs and humans can lead students to the mistaken idea that these species lived at the same time (and, of course, they did not).

4. *Vernacular misconceptions* – Misconceptions that arise from the everyday use versus scientific use of words. For example, "theory" in everyday use means an unsubstantiated idea; the scientific use of "theory" means an idea with substantial supporting evidence.
5. *Religious and myth-based misconceptions* – Ideas that come from religious or mythical teaching that, when transferred to science education, become factually inaccurate. For example, the belief that the Earth is too young for evolution (given the Bible's dating of the Earth at 10,000 years).

We met with a focus group of seven experienced middle and high school teachers from diverse institutions in June 2011 to determine which misconceptions they observe most frequently in their students. The teachers ranked how frequently they encountered a set of 11 common evolution misconceptions in their classrooms. The set of misconceptions was derived from a literature review (e.g., AAAS, 2011; Anderson, Fisher, & Norman, 2002; Bishop & Anderson, 1990; Lane, 2011) and identification of the misconceptions that are relevant to *VTG*. The rankings ranged from, at the top, "Natural selection involves organisms 'trying' to adapt" to the bottom ranking of "Sudden environmental change is required for evolution to occur."

In order to develop educational technology to help students learn about evolution, it is important to understand the types of prior knowledge and misconceptions they have, such as those mentioned above. If prior knowledge is not directly engaged, students may have trouble grasping the new concepts (Bransford, Brown, & Cocking, 2000). Inquiry learning is one method of engaging prior knowledge and overcoming misconceptions. Prior work has shown that good scientific inquiry – both in general and in the use of educational technology – involves systematic steps such as formulating questions, developing hypotheses, designing experiments, analyzing data, drawing conclusions, and reflecting on acquired knowledge. Essentially, students who mimic (or are guided towards) the cognitive processes of scientific experts are most likely to benefit from inquiry (De Jong & van Joolingen, 1998). In addition, while undertaking these scientific steps, students are likely to reveal and/or act upon their misconceptions, which, in turn, can be directly addressed by the feedback and guidance provided by an educational technology system.

3 *Voyage to Galapagos* – Software to Help Students Explore and Learn About the Theories of Natural Selection and Evolution

Our approach to overcoming misconceptions about evolution is to have students work with *Voyage to Galapagos,* a web-based, inquiry-driven instructional program that mirrors Darwin's pathway to the development of the theories of natural selection and evolution. The program, which is largely implemented but still under development, encourages the student to follow the steps of good scientific inquiry, e.g., developing hypotheses, analyzing data, drawing conclusions, and reveals the basic principles of evolution theory to the students. Darwin's early ideas about the possibility of an ongoing process of continual modification of flora and fauna were initially inspired by his observations in the Galapagos islands, where he noted the patterns of species distribution on the archipelagos. Students working with *Voyage to Galapagos* have the opportunity to "follow" Darwin's steps and observe and analyze differences among island fauna. This occurs through a virtual exploration of six Galapagos islands where students take photographs of different animals and then perform various analyses.

Figure 1 shows a screen shot of *VTG* in which the student is located on the Santa Fe Galapagos Island and has a land iguana captured in her camera viewfinder (see box at the bottom of the photograph). An overall view of the Galapagos Islands is shown in the upper right, and a close-up view of a portion of a selected island (in this case Santa Fe) is shown in the lower right. The student can follow the virtual path on the selected island, by selecting individual steps that are in the close-up view of that island. When a step is selected, a picture of the view from that point on the island is shown (Note: the pictures in the program are authentic, taken by the third author of this paper during a visit to the Galapagos).

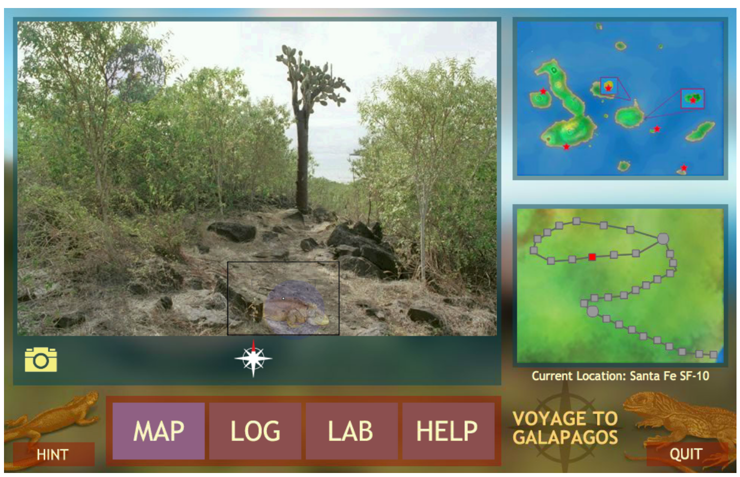

Fig. 1: The *VTG* inquiry-learning program. Here the student is about to take a picture of an iguana (see the bottom of the photograph)

As the student takes pictures of animals, they are stored in her Logbook, the central repository (as well as organizing tool) for the student's inquiry (see Figure 2). Students are instructed to collect animals (i.e., finches or iguanas) that have as much variation between them as possible. The student is placed on either the "iguana" path of the program, in which case they should try to take pictures exclusively of iguanas, or the "finch" path, in which case they should try to take pictures exclusively of finches. They can take up to 12 photographs in an attempt to cover as wide a variety as possible of characteristics of the species of interest.

The Lab is the place where students perform various analyses on the data they have collected. The Schematometer is a tool that allows the student to examine, measure, and categorize characteristics (e.g., size, shape, color) of the collected animals. The Trait Tester is a tool that allows the student to test a hypothesis about the function of a trait variation. The Chart Creator is a data analysis tool that allows the student to plot the various measured characteristics of the collected animals for comparison purposes.

There is considerable "student action" variability within *VTG*; that is, there are many degrees of freedom and opportunities for students to make mistakes. For instance, as shown in Figure 2, the student can take pictures of irrelevant species when they are supposed to focus on either iguanas or finches. The student might take pictures of a single species, say iguanas, but not capture enough trait variation. The student can visit islands that have little useful data to collect or compare traits that will not be useful in learning about variation. This potential variability of student actions – and student errors – allows for a wide variety of assistance, and the ability to either intervene after those actions are taken with help – or not. This provides the foundation for our experimental test bed.

Fig. 2: The Logbook of the *VTG* inquiry-learning program. Here the student has taken 3 pictures – 2 iguanas and 1 sea lion. It is a mistake – but still permitted by the program – for the student to photograph a sea lion, since he or she is on the "iguana" path, concentrating on evaluating iguanas

A Bayesian Network is used to collect data about student actions, assign probabilities of students having made certain errors, and make decisions about error feedback and hints to provide students (provided students are in the conditions to receive such assistance). The Bayes Net has three top layers that range from the most general to most specific – the Knowledge, Skills, and Abilities (KSA) Layer, the Error Evaluation Layer, and the Error Diagnosis Layer. The specific nodes at each of these layers have associated error feedback and hints that are triggered when the nodes at the associated level reach a certain threshold. Whether a student receives the feedback or hints is configurable according to (a) what condition they are in and (b), in the case of hints, whether they request help. By allowing the assistance to be configured in this way, we are able to create the conditions of assistance that are the focus of our experimental design, which is discussed next.

4 Planned Study to Explore the Assistance Dilemma in *VTG*

We have two research questions to answer with our planned quasi-experimental design:

1. How much assistance do students who learn with *VTG* require to achieve the highest learning gains and maximize their inquiry-learning skills?
2. Which mode of assistance is optimal for students with high, medium and low levels of prior science knowledge?

Our goal is to find the right balance between, on the one hand, full support and, on the other hand, allowing students to make their own decisions and, at times, mistakes. Our approach conceives of this as a spectrum of assistance, driven by two orthogonal variables. The first, *Frequency of Intervention,* is important because it determines *how often* a student is given assistance and, therefore, the total amount of help provided over an entire session with *VTG*. *Level of Support* is important because it refers to *how much* and *what type* of assistance the student receives. At the low end of the spectrum this is simply indicating the errors a student has made. The next point on the spectrum is to both flag the errors and explain the nature of the errors. At the high end of this spectrum, the assistance involves flagging the error, explaining its nature and providing (on demand) help about how to correct it.

Figure 3 shows the research design that results from crossing these two variables. This 3 x 3 matrix provides a maximum possible set of 9 assistance conditions, but we have combined some of the cells and will not be testing two others. First, the "Never" value of the *Frequency of Intervention* variable essentially means that no assistance will ever be provided, so the *Level of Support* variable is not applicable in such a case. Thus, we combined all three cells of the first column of Figure 3 to create Condition 1 (*NoSupport*). Second, we want to have a relatively wide mid-range of assistance; this is achieved by having variations of "When Struggling", varying over the Level of Support variable (i.e., error flagging only (Condition 2 - *Flagging*), error flagging and error feedback (Condition 3 - *Flagging&Feedback*), and error flagging, error feedback, and hints (Condition 4 – *Flagging&Feedback&Hints*)). Note that in all three of these mid-level conditions the provision of assistance is predicated on the current value of nodes in the Bayes Net. Finally, we wanted both to include the most extreme level of assistance (i.e., always providing assistance, providing all three levels of assistance, and also providing pre-emptive assistance), yet also limit the total number of conditions in the experiment. Thus, we will include the most extreme form of assistance (Condition 5 - *FullSupport*), while excluding the somewhat less extreme forms of assistance (those in the upper right).

Level of Support		Frequency of Intervention: 1. Never	Frequency of Intervention: 2. When Struggling	Frequency of Intervention: 3. Always
	1. Error Flagging + **2. Error Feedback +** **3. Hints**		**Condition 2** **(*Flagging*)**. When the student exceeds a probability threshold in the Bayes Net for a particular error type: 1. Flag all errors from this point on Cease assistance when the probability drops below threshold.	***Skipped Condition***. Begin with a preemptive hint. From the very beginning: 1. Flag all errors
	1. Error Flagging + **2. Error Feedback**	**Condition 1:** **(*NoSupport*)**. Base *VTG* System with no support during investigation tasks	**Condition 3** **(*Flagging&Feedback*)**. When the student exceeds a probability threshold in the Bayes Net for a particular error type: 1. Flag all errors from this point on 2. Provide error feedback on nature of error Cease assistance when the probability drops below threshold.	***Skipped Condition***. Begin with a preemptive hint. From the very beginning: 1. Flag all errors 2. Provide error feedback on nature of error
	1. Error Flagging		**Condition 4** **(*Flagging&Feedback&Hints*)**. When the student exceeds a probability threshold in the Bayes Net for a particular error type: 1. Flag all errors from this point on 2. Provide error feedback on nature of error 3. Make hints available for student selection (hints will range from principles to concrete) Cease assistance when the probability drops below threshold.	**Condition 5** **(*FullSupport*)**. Begin with a preemptive hint. From the beginning and throughout the entire learning session: 1. Flag all errors 2. Provide error feedback on nature of error 3. Make hints available for student selection (hints will range from principles to concrete)

Fig. 3: The Experimental Design, crossing two variables of assistance

With respect to our first research question (i.e., “How much assistance do students who learn with *VTG* require to achieve the highest learning gains and maximize their inquiry-learning skills?”), our hypothesis is that one of the middle conditions – *Flagging*, *Flagging&Feedback*, or *Flagging&Feedback&Hints* – will lead to the best domain and inquiry learning outcomes for the overall student population. These conditions all trade off between assistance giving (such as what is provided by Condition 5) and assistance withholding (such as what is provided by Condition 1). With respect to our second research question (i.e., "Which mode of assistance is optimal for students with high, medium and low levels of prior science knowledge?"), we hypothesize that Condition 1 (no assistance) will be most beneficial to higher prior knowledge learners and Condition 5 (high assistance) will be most beneficial to lower prior knowledge learners. Our theory is that higher prior knowledge students are more likely to benefit by struggling a bit and exploring without guidance, while lower prior knowledge students, those who are more likely to experience too much cognitive load (Paas, Renkl, & Sweller, 2003) if left on their own, are more likely to benefit by being strongly supported.

5 Conclusion

The assistance dilemma is a fundamental challenge to learning scientists and educational technologists. Until we better understand how much guidance students need as they learn – and how to cater guidance to the prior knowledge level of students – we won't be able to appropriately design software to best support student learning. This is especially so in domains and with software that are open ended, i.e., those that encourage exploration and inquiry. The *VTG* software, a web-based inquiry-learning environment for learning about the theory of evolution, will allow us to experiment with different types of instructional support and provide an important data point in answering the assistance dilemma. We are in the process of completing implementation of this program and will soon conduct the experiment described in this paper.

Acknowledgements: This research is supported by a U.S. Department of Education, Institute Of Education Science Grant (Number R305A110021). We would like to thank the other members of the *VTG* project team, including Donna Winston, Kim Luttgen, Dave Brown, Nick Matzke, and Russell Almond, for their contributions to the development of *VTG*.

6 References

AAAS (2011). Project 2061: AAAS Science Assessment. *AAAS: Advancing Science, Serving Society*. http://assessment.aaas.org/

Alters, B. and Nelson, C.E. (2002). Teaching evolution in higher education. *Evolution: International Journal of Organic Evolution*, 56(10), 1891-1901.

Anderson, D.L., Fisher, K.M., & Norman, G.J. (2002). Development and evaluation of the conceptual inventory of natural selection. *J. of Research in Sci Teaching*, *39*(10), 952–978.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.) (2000). *How People Learn*. Washington, D.C.: National Academy Press.

Bishop, B. A., & Anderson, C.W. (1990). Student conceptions of natural selection and its role in evolution. *J. of Research in Science Teaching*, *27*(5), 415–427.

De Jong, T. & van Joolingen W.R. (1998). Scientific Discovery Learning with Computer Simulations of Conceptual Domains. *Review of Educational Research*, 68(2), 179-201.

Edelson, D. (2001). Learning-for-use: A framework for the design of technology-supported inquiry activities. *Journal of Research in Science Teaching 38*(3), 355-385

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., et al. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *J. of Research in Science Teaching, 45*(8) 922–939.

Kapur, M. (2009). Productive failure in mathematical problem solving. *Instructional Science*. doi: 10.1007/s11251-009-9093-x.

Klahr, D. (2009). "To every thing there is a season, and a time to every purpose under the heavens"; What about direct instruction? In S. Tobias and T. M. Duffy (Eds.) Constructivist Theory Applied to Instruction: Success or Failure? Taylor and Francis.

Kirschner, P.A., Sweller, J., and Clark, R.E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 75—86.

Koedinger, K. R. and Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*, 239–264.

Lane, E. (2011). U.S. Students and Science: AAAS Testing Gives New Insight on What They Know and Their Misconceptions. *AAAS: Advancing Science, Serving Society* http://www.aaas.org/news/releases/2011/0407p2061_assessment.shtml

Linn, M.C. and Hsi, S. (2000). *Computers, Teachers, Peers: Science Learning Partners*. Hillsdale, NJ: Erlbaum.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, 59, 14–19.

Paas, F.G., Renkl, A., and Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38:1-4.

Quintana, C., Reiser, B.J., Davis, E.A., Krajcik, J., Fretz, E., Duncan, R.G. et al. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, 13, 337-387.

Sandoval, W. A. and Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, 88. 345-372.

Schmidt, R.A. and Bjork, R.A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3(4), 207-217.

Slotta, J. (2004). The web-based inquiry science environment (WISE): Scaffolding knowledge integration in the science classroom. In: M. Linn, E.A. Davis, & P. Bell (Eds), *Internet environments for science education* (pp. 203-233). Mahwah, NJ: Erlbaum.

Tsovaltzi, D., Rummel, N., McLaren, B.M., Pinkwart, N., Scheuer, O., Harrer, A. and Braun, I. (2010). Extending a virtual chemistry laboratory with a collaboration script to promote conceptual learning. *International Journal of Technology Enhanced Learning*, Vol. 2, Nos. 1/2, 91–110.

van Joolingen, W..R., de Jong, T., Lazonder, A.W., Savelsbergh, E., and Manlove, S. (2005). Co-Lab: Research and development of an on-line learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21*, 671-688.