In Proceedings for the 18th International Conference on Intelligent Tutoring Systems (ITS), Bucharest, Romania, June 29 - July 1 2022

# Effects of Guidance on Learning about Ill-defined Problems

Sungeun An[1], Emily Weigel[2], and Ashok K. Goel[1]

[1] School of Interactive Computing, Georgia Institute of Technology, Atlanta GA 30308, USA
[2] School of Biological Sciences, Georgia Institute of Technology, Atlanta GA 30332, USA
sungeun.an@gatech.edu

**Abstract.** We present a study that examines the effects of guidance on learning about addressing ill-defined problems in undergraduate biology education. Two groups of college students used an online laboratory named VERA to learn about ill-defined ecological phenomena. While one group received guidance, such as giving the learners a specific problem and instruction on problem-solving methods, the other group received minimal guidance. The results indicate that, while performance in a problem-solving task was not different between groups receiving more vs. minimal guidance, the group that received minimal guidance adopted a more exploratory strategy and generated more interesting models of the given phenomena in a problem-solving task.

**Keywords:** Self-directed learning · Guided learning · Modeling and simulation · Ill-defined Problems · Online laboratory.

## 1 Introduction

Interactive Learning Environments (ILEs) enables learners to explore ill-defined problems such as modeling complex systems [3][4]. Many studies have found that while guided learning in ILEs leads to more efficient learning [6][10][15], it can limit critical thinking and creative problem solving [7][11]. In contrast, while self-directed exploration can be inefficient [15], it often results in more varied and interesting solutions [18].

However, many studies exploring the effects of guidance in ILEs have focused on K-12 education with well-defined goals, assessments, and outcomes [6][14][19][17][10]. The precise role of guidance in the context of ill-defined problem-solving is not yet well understood. For example, Chen (2007) found that problem-solving prompts did not have a positive effect on solving ill-structured problems [5]. This contradicted Ge & Land's (2003) study indicating that prompts had significantly improved students' problem-solving [8]. Koedinger & Aleven (2007) called balancing the amount of guidance the "assistance dilemma" to achieve the optimal learning [16]. To solve the assistance dilemma, it is necessary to understand the trade-offs in providing guidance for learning on ill-defined problems as well.

We present a study on the use of an ILE called VERA (for Virtual Experimentation Research Assistant) for learning about ill-defined problems in undergraduate biology education. VERA is a web-based online laboratory that enables users to construct conceptual models of ecological systems and run interactive agent-based simulations of these models [1][2] (`vera.cc.gatech.edu`). This allows learners to explore multiple hypotheses about ecological phenomena and perform "what if" experiments to either explain an ecological phenomenon or attempt to predict the outcomes of changes to an ecological system. Our goal in this study was to answer two research questions. *RQ1: Does more guidance indeed make learning more efficient in ill-defined problem-solving contexts? RQ2: Does minimal guidance indeed result in more varied and interesting models?*

To answer these questions, we conducted an experiment to analyze the relative effects of guidance on the learning process and outcomes. Two groups of students engaged in two different but related learning activities: one with guidance such as giving the learners a specific problem and instruction on problem-solving methods, and the other with minimal guidance. After the learning activities, a new problem-solving task was given to students to examine which learning context enabled the students to solve the task effectively and creatively.

## 2 A Brief Description of the VERA Online Laboratory

In VERA, learners build conceptual models of complex phenomena, evaluate them through simulation, and revise the models as needed. VERA enables learners to construct Component-Mechanism-Phenomena (CMP) models [13] of ecological phenomena that originate from our earlier work on Structure-Behavior-Function models of complex systems [9]. A CMP model specifies the biotic and abiotic components participating in an ecological phenomenon, the relationships among the components, and the processes that arise through the interactions among the components. VERA automatically translates a CMP model into an agent-based simulation on the NetLogo platform [12] (`ccl.northwestern.edu/netlogo`). Using VERA system involves three high-level activities: (1) model construction (e.g., adding biotic/abiotic components and relationships), (2) parameterization (e.g. selecting a set of simulation parameters for each component/relationship), and (3) execution of the agent-based simulation. Parameterization is further divided into four categories: (1) *t-parameters* (trivial parameter) that directly manipulates an initial population (e.g., initial population, minimum population), (2) *b-parameters* related to the biological properties (e.g., lifespan, reproductive rate), (3) *a-parameters* related to the abiotic substances (e.g., amount, growth rate), and (4) *i-parameters* that adjust the relationship between two components (e.g., interaction probability, consumption rate).

## 3 Experimental Design

The goal of this study was to compare the relative effects of more guidance and minimal guidance. Figure 1 illustrates a schematic diagram of the study. Independent variable is the amount of guidance in Phase 1 (problem representation) and Phase 2 (problem solving). Dependent variables are outcomes (performance, efficiency, model quality) in Phase 3.

### 3.1 Participants

The participants were college students in an undergraduate biology class at a large public R1 institution located in the southeastern US. This class teaches scientific methods and skills that address population ecology. Students in the class used VERA in the accompanying laboratory sections. The duration of this class was 2 hours and 45 minutes. Our study consisted of six lab sections (2 hr 45 min each) with 16 students for each section (N=96). The students were assigned to either Group A (guided) or Group B (self-exploration) based on their sections.

### 3.2 Procedure

As shown in Figure 1, the students in three lab sections were assigned to Group A (N=48); The students in the remaining three sections were assigned to Group B (N=48). As students arrived in the classroom, two teaching assistants played a pre-recorded introduction video giving a high-level description of VERA and how to use the system (5 minutes). All students were then asked to complete a consent form and indicate their major and year. The modeling task consisted of three phases: problem representation (Phase 1), problem solving (Phase 2), and a problem-solving task (Phase 3). In Phase 1 and 2, Group A explored a given problem with guidance in using VERA (e.g., representing a problem, solving a problem, as described in detail below). Meanwhile, Group B explored a problem of their choice with minimal guidance. In Phase 3, Group A and Group B were given the same problem to solve.

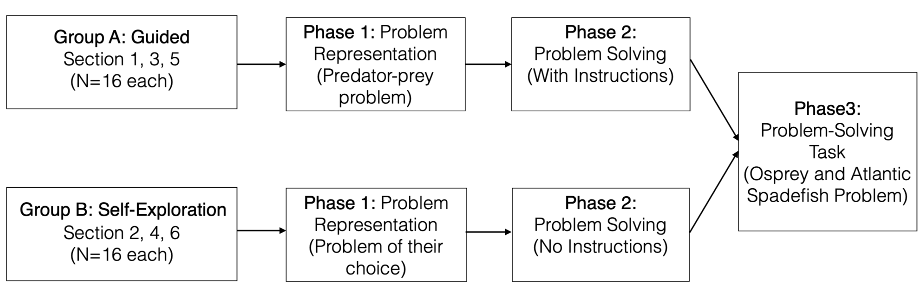


**Fig. 1.** Schematic Diagram of Experimental Design.

### 3.3 Group A: Guided Learning

The students in Group A were given some heuristics guidance during Phase 1 and Phase 2 (N=48). In Phase 1, the students were asked to create a model that represents a common interaction in ecology: a predator-prey relationship in which the predator consumes the prey and the prey gradually decreases. In Phase 2, the students were asked to adjust their models by reducing the predator population so that the prey can survive longer. In this way, the students were given a problem representation as well as instructions on how to solve it. The instruction stated "there are many ways to reduce the population of the predators. One way is to directly reduce the predator's initial population. Another way is to indirectly reduce the predator's population by adjusting various parameter values of the predator."

### 3.4 Group B: Self-Exploration

The students in Group B were given minimum guidance during Phase 1 and Phase 2 (N=48). In Phase 1, the students were asked to create a model that represents any problem in ecology which includes an abnormal increase or decrease in the population of a target species. In Phase 2, the students were asked to adjust their models to mitigate the problem in their previous model by adding/deleting components and/or adjusting various parameter values. Since the students in Group B received more general instructions, they had to define a problem as well as explore various hypothetical solutions to solve the problem.

### 3.5 Problem-Solving Task

In Phase 3 (Problem-Solving Task), the students both in Group A and Group B were given a new problem to solve. The pre-built model that represents a food web between Osprey (*Pandion haliaetus*) and Atlantic Spadefish (*Chaetodipterus faber*) was given to both groups of students (see 2 (A)). Their goal was to adjust this model to create a stable ecosystem where both species can survive. The students were asked to write a list of initial thoughts of how they would change the model. Then, they tested their hypothetical solutions in VERA and compared them in terms of effectiveness, cost, practicality, etc.

## 4 Data Analysis

### 4.1 Outcome Measures

The students' outcomes during Phase 3 (problem-solving task) were measured using the students' models, their log data, and survey answers on Qualtrics based on the following metrics.

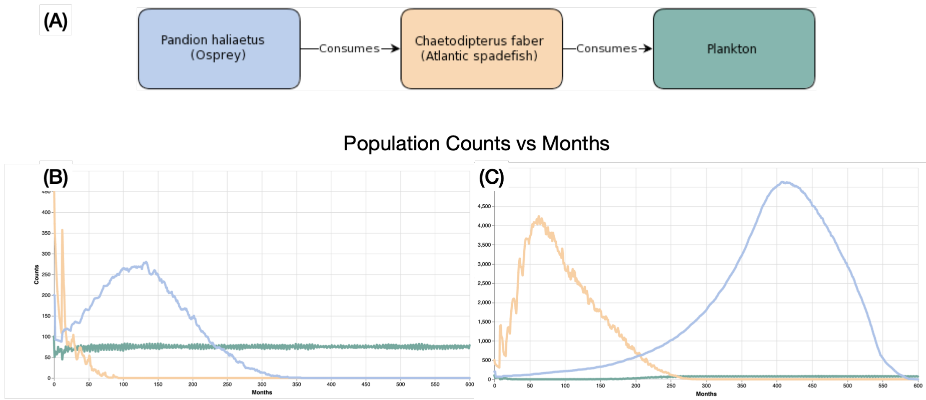


**Fig. 2.** (A) The Original Model Used in the Task. (B) The Simulation Output of the Original Model. (C) The Simulation Output of the Successfully Adjusted Model (Both Osprey and Atlanta Spadefish Survived Longer than the Original Model.)

- ***O1* Success in making models according to specification**: The simulation outcomes between the original model and the students' adjusted models were compared. The score is 1 (correct) when both species (Osprey and Atlantic Spadefish) survived longer than the original model (for example, Figure 2 (C)); 0.5 (partially correct) when only one species survived longer; and 0 (incorrect) when none of the species survived longer.
- ***O2* The number of hypothesis**: The number of hypothetical solutions and corresponding models generated by the students were measured.
- ***O3* Time taken to making models according to specification**: The time taken in the problem-solving task (Phase 3) was measured based on the time the students submitted their answers on Qualtrics.
- ***O4* Model complexity (depth)**: Model complexity (the total number of components and relationships that exist within a conceptual model) was measured. If a student doesn't add or delete any components (nodes) and relationships (edges), the model complexity remains 5 as the original model (e.g., see Figure 2).
- ***O5* Model variability (breadth)**: Model variability measures how many distinct simulation parameter categories were used. It is calculated by the percentage of the parameter categories changed within a conceptual model. Since there was no abiotic substance in the original model, *t-parameters, b-parameters, and i-parameters* were used.

### 4.2 Results

A total of 79 students consented to the study (A=39; B=40): Participants included 53 female, 23 male, 3 prefer not to say; 19-23 years of age, mean 20.2 years. Group A created a total of 81 models for the task; Group B created 84 models. A Shapiro-Wilk test was performed in each outcome and showed a significant departure from normality. Based on this outcome, a non-parametric test, Mann-Whitney U test, as well as descriptive statistics, were used to summarize the results.

**Performance and Efficiency (O1, O2, O3)** Figure 3 shows comparative outcomes *O1 (Task score), O2 (The number of hypotheses), and O3 (Completion Time)* between the two groups. The mean of the task score was higher for the guided group ($M = 0.91, SD = 0.17$) and the self-exploration group ($M = 0.84, SD = 0.20$) with the same median and the interquartile range ($Mdn = 1.0, IQR = 0.25$). The number of hypotheses generated during the problem-solving task was also higher for the guided group ($M = 2.20, SD = 0.57$) than the exploration group ($M = 2.25, SD = 0.58$) with the same median and the interquartile range ($Mdn = 2.0, IQR = 0.0$). On the other hand, the mean and the median of completion time was lower for the guided group ($M = 842.55, SD = 511.19, Mdn = 708.34, IQR = 355.56$) than the exploration group ($M = 957.49, SD = 565.41, Mdn = 832.20, IQR = 422.67$).

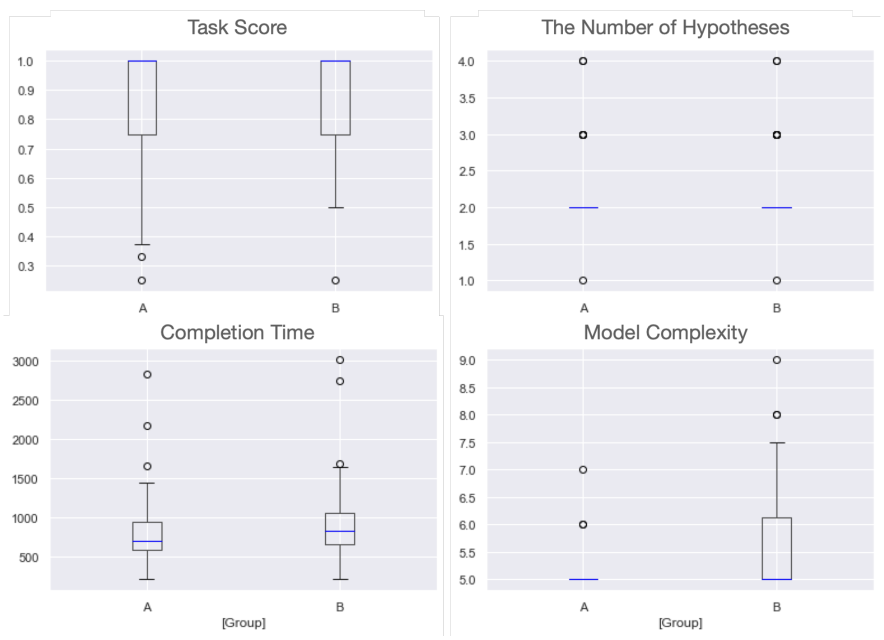


**Fig. 3.** Comparative Outcomes: (O1) Task Score. (O2) The Number of Hypotheses. (O3) Completion Time. (O4) Model Complexity.

**Model Quality (O4, O5)** We measured model quality using two proxies: *model complexity* and *model variability*. Figure 3 shows comparative model complexity between the two groups. The mean of the measured model complexity was higher in the exploration group B ($M = 8.02, SD = 1.45, Mdn = 5.0, IQR = 1.25$) than the guided group A ($M = 7.09, SD = 0.74, Mdn = 5.0, IQR = 0.0$). This difference was statistically significant as determined by the Mann-Whitney U test ($U = 503.0, p < .0005$).

Figure 4 shows comparative model variability between the two groups. The most frequently used parameter for the guided group was the t-parameter of the predator (Osprey) ($M = 23.32, SD = 21.76$), followed by b-parameter of the prey (Atlantic Spadefish) ($M = 21.07, SD = 25.65$) and b-parameter of the predator ($M = 15.74, SD = 23.92$). On the contrary, the most frequently used parameter for the self-exploration group was i-parameter, the interaction parameter, ($M = 18.08, SD = 17.82$), followed by t-parameter of the predator ($M = 17.67, SD = 24.56$) and b-parameter of the prey ($M = 17.44, SD = 24.02$).

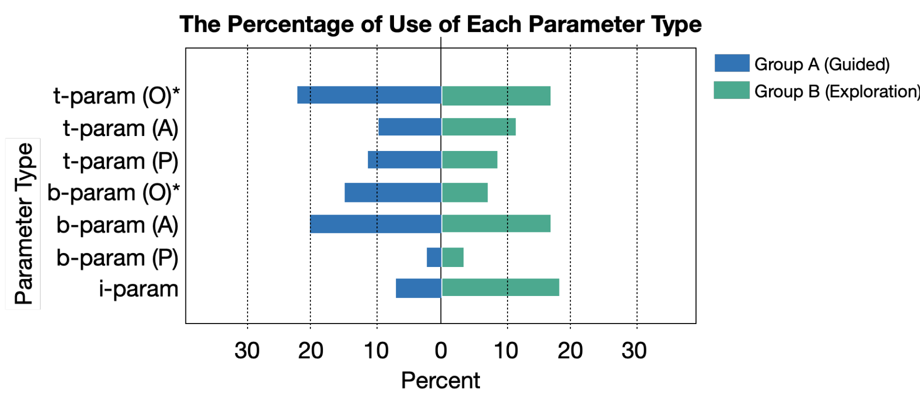


**Fig. 4.** The Percentage of Use of Each Parameter Type. *Parameters for Osprey (Predator). O:Osprey. A:Atlantic Spadefish. P:Plankton.

## 5 Discussion

Our data concur with earlier findings in well-defined problem contexts (e.g., [6]), specifically that giving more information or assistance leads to higher accuracy and efficiency of learning [6][10]. In our study, the students who were given guidance in solving an ill-defined problem solved the second ill-defined problem in Phase 3 slightly more efficiently in terms of completion time (O3), but the difference was not statistically significant. Irrespective of guidance, there does not seem to be much difference in the ability to understand the problem, devise hypothetical solutions, and test and validate one's own hypotheses.

However, our data adds to what is known about guidance in relation to the complexity of models (O4). While the students in the guided group typically made small changes by adjusting parameter values, some students in the self-exploration group also added more components to the model. Given that the performance between the two groups was similar (e.g., task score, the number of hypotheses, and completion time), the way they approached the problem showed interesting differences. In addition to manipulating simulation parameters, adding new components and relationships as we observed in our data likely represents more constructivist learning and discovery learning described in other studies [18][6][20][15]. Many studies have shown that minimally guided learning can foster students' creative thinking by having them independently explore broader issues [21][18]. Similarly in our study, the students solved the task by adding new components beyond manipulating existing variables.

The students in the guided group focused on adjusting the parameters of the components, especially the parameters of the predator (see Figure 4). This may be due to the fact that the students were asked to adjust their models by reducing the predator population in Phase 2. On the contrary, the students in the self-exploration group approached the problem in more varied ways. In addition to adding new components and relationships, they commonly used i-parameters such as interaction probability and consumption rate, which were not frequently used by the guided group. This suggests that providing a specific example or instruction on problem-solving methods has the potential to affect students' learning as it might restrict what learners try to do in learning.

## 6 Conclusion

In this paper, we presented a study in which two groups of college students used the online laboratory called VERA to learn about ecological phenomena: while one group received more guidance (e.g., heuristics), the other group received minimal guidance (e.g., process constraints). The data indicates that the group that received minimal guidance adopted a more exploratory strategy on the problem-solving task and generated more varied and complex models of the given phenomena. However, the group that received more guidance did not show significant benefits in efficiency and accuracy, which was unexpected. These preliminary results invite further research on the proper level of guidance in learning about ill-defined problems in undergraduate science education.

## Acknowledgement

This research was supported by US NSF grant #1636848. We thank members of the VERA project, especially Spencer Rugaber, Luke Eglington, and Stephen Buckley. We also thank the TAs and the students in the biology lab. This research was conducted in accordance with IRB protocol #H18258.

## References


1. An, S., Bates, R., Hammock, J., Rugaber, S., Weigel, E., Goel, A.: Scientific modeling using large scale knowledge. In: International Conference on Artificial Intelligence in Education. pp. 20–24. Springer (2020)
2. An, S., Broniec, W., Rugaber, S., Weigel, E., Hammock, J., Goel, A.: Recognizing novice learner's modeling behaviors. In: International Conference on Intelligent Tutoring Systems. pp. 189–200. Springer (2021)
3. Brown, A.L., Campione, J.C.: Interactive learning environments and the teaching of science and mathematics. Toward a scientific practice of science education pp. 111–139 (1990)
4. Brown, A.L., Campione, J.C., Webber, L.S., McGilly, K.: Interactive learning environments: A new look at assessment and instruction. In: Changing assessments, pp. 121–211. Springer (1992)
5. Chen, C.H., Bradshaw, A.C.: The effect of web-based question prompts on scaffolding knowledge integration and ill-structured problem solving. Journal of research on Technology in Education **39**(4), 359–375 (2007)
6. Cohen, M.T.: The effect of direct instruction versus discovery learning on the understanding of science lessons by second grade students (2008)
7. Dabbagh, N.: Scaffolding: An important teacher competency in online learning. TechTrends **47**(2), 39 (2003)
8. Ge, X., Land, S.M.: Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. Educational technology research and development **51**(1), 21–38 (2003)
9. Goel, A.K., Rugaber, S., Vattam, S.: Structure, behavior, and function of complex systems: The structure, behavior, and function modeling language. Ai Edam **23**(1), 23–35 (2009)
10. González-Calero, J.A., Arnau, D., Puig, L., Arevalillo-Herráez, M.: Intensive scaffolding in an intelligent tutoring system for the learning of algebraic word problem solving. British Journal of Educational Technology **46**(6), 1189–1200 (2015)
11. Joolingen, W.v., Jong, T.d.: Model-based diagnosis for regulative support in inquiry learning. In: International handbook of metacognition and learning technologies, pp. 589–600. Springer (2013)
12. Joyner, D.A., Goel, A.K., Papin, N.M.: Mila–s: generation of agent-based simulations from conceptual models of complex systems. In: Proceedings of the 19th international conference on intelligent user interfaces. pp. 289–298 (2014)
13. Joyner, D.A., Goel, A.K., Rugaber, S., Hmelo-Silver, C., Jordan, R.: Evolution of an integrated technology for supporting learning about complex systems. In: 2011 IEEE 11th International Conference on Advanced Learning Technologies. pp. 257–259. IEEE (2011)
14. Kim, M.C., Hannafin, M.J.: Scaffolding 6th graders' problem solving in technology-enhanced science classrooms: A qualitative case study. Instructional science **39**(3), 255–282 (2011)
15. Klahr, D., Nigam, M.: The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. Psychological science **15**(10), 661–667 (2004)
16. Koedinger, K.R., Aleven, V.: Exploring the assistance dilemma in experiments with cognitive tutors. Educational Psychology Review **19**(3), 239–264 (2007)
17. Lazonder, A.W., Harmsen, R.: Meta-analysis of inquiry-based learning: Effects of guidance. Review of educational research **86**(3), 681–718 (2016)

18. Rahman, M.H.: Using discovery learning to encourage creative thinking. International Journal of Social Sciences & Educational Studies **4**(2), 98 (2017)
19. Rittle-Johnson, B., Koedinger, K.R.: Designing knowledge scaffolds to support mathematical problem solving. Cognition and Instruction **23**(3), 313–349 (2005)
20. Schunk, D.H.: Learning theories an educational perspective sixth edition. Pearson (2012)
21. Veermans, K., de Jong, T., van Joolingen, W.R.: Promoting self-directed learning in simulation-based discovery learning environments through intelligent support. Interactive learning environments **8**(3), 229–255 (2000)