WIP: Exploratory Learning Before Instruction on Python Error Messages: Looking Beyond the Learning Outcomes

Marci S. DeCaro
Department of Psychological and Brain Sciences
University of Louisville
Louisville, KY, USA
marci.decaro@louisville.edu

Angela Thompson
Department of Engineering Fundamentals
University of Louisville
Louisville, KY, USA
angela.thompson@louisville.edu

Lianda Velić
Department of Psychological and Brain Sciences
University of Louisville
Louisville, KY, USA
lianda.velic@louisville.edu

Cenetria Crockett
Department of Anthropology
University of Louisville
Louisville, KY, USA
cenetria.crockett@louisville.edu

Campbell Bego
Department of Engineering Fundamentals
University of Louisville
Louisville, KY, USA
campbell.bego@louisville.edu

***Abstract*—This research work in progress research paper examines student perceptions after completing an exploratory learning lesson before instruction on an introductory programming concept. During exploratory learning activities, students explore a novel concept prior to instruction—the reverse of typical instruct-then-practice methods. Exploratory learning before instruction can help students activate prior knowledge, become aware of their knowledge gaps, and discern important problem features to improve conceptual understanding. Students in a first-year engineering course (*N*=402) learned about Python error messages in one of two conditions. In the *explore-first condition*, students completed a collaborative activity prior to instruction. In the *instruct-first condition*, students received instruction prior to the activity. Following the activity and instruction, students completed a survey to assess their perceptions of the activities. Survey items (e.g. cognitive load, self-efficacy, belonging, knowledge gaps) were chosen as potential factors that could explain learning outcomes between the two conditions. In prior work, we found higher posttest scores in the instruct-first compared to explore-first condition, contrary to the majority of previous studies. Cognitive load and knowledge gaps were higher in the explore-first condition than the instruct-first condition. Self-efficacy and competence were lower in the explore-first condition. No other significant differences were found. Exploring before instruction might disrupt learning and perceived efficacy and competence if the activity is too challenging, or if the instruction does not fully resolve gaps in students' knowledge.**

Keywords—exploratory learning, productive failure, engineering education, programming

I. Introduction

Engineering students are introduced to a wide variety of topics in their first year, many of which are foundational for more advanced topics in subsequent courses. One way that instructors can help students better learn these threshold concepts is to implement evidence-based teaching methods at strategic points in the semester. Exploratory learning before instruction (explore-instruct), for example, is a two-phase method of active learning. In exploratory learning, a concept is first introduced with a new activity or problem to solve (i.e., an activity phase). Then, following the activity, the correct problem-solving methods and underlying concepts are explained through instruction (i.e., instructional phase). This order of instruction is opposite a traditional lesson in which the activity follows instruction.

This reversal can invoke several mechanisms that support learning. For example, students are likely to first attempt to solve a novel problem with their existing knowledge [1]. When they realize that they cannot solve the problem, students have productively discovered gaps within their current knowledge [2]. This awareness of knowledge gaps can motivate students to focus more during the instructional phase. In addition, exploration can help students identify critical problem features that aid in understanding the new concept during instruction [3]. Together, these mechanisms can result in stronger conceptual knowledge and transfer [1], [2], [4], [5], [6]. Thus, although exploring a novel concept can be difficult, this challenge is often considered to be "productive failure" [1].

One recent study in an undergraduate programming course examined student emotion and type of reasoning during exploration [7]. Students collaborated with a partner online while working on a novel programming task (sorting numbers in a list). Conversations were audio recorded and transcribed, and students responded to questionnaires following the activity and following instruction. The questionnaire and qualitative coding of transcripts focused on participants' emotional states (e.g., anxiety, boredom, confusion, and enjoyment) and types of reasoning (e.g., constructive, active, see [8]). There were no significant correlations between emotion frequency and posttest scores. However, constructive reasoning during collaboration

This work was supported by the NSF Division for Undergraduate Education under grant number 2012342. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the University of Louisville.

demonstrated a positive correlation with posttest scores. Although not testing the difference between instruct-first and explore-first conditions, this study indicates that constructive processes during exploration can increase learning in programming.

However, in another recent experiment in an introductory engineering programming course, the learning benefits were reversed [29]. Students were randomly assigned to either a traditional instruct-then-practice condition (instruct-first condition) or an exploratory learning condition (explore-first condition). Students were provided instruction and an activity on types of error messages, followed by an assessment to evaluate their learning of the topic. Contrary to expectations, students who received the traditional instruct-first approach scored higher than students in the explore-first condition on the assessment.

During the study, a survey was administered to the students following the activity and instruction. The survey questions examined several factors that could potentially explain learning outcomes. These included cognitive load (i.e., perceived mental effort exerted during the learning session; [9]), self-efficacy, competence, flow state, as well as several aspects of motivation. These also included two social-psychological factors assessing perceived threat in the learning environment (i.e., security, belonging), and students' awareness of knowledge gaps for the topic learned in the class session.

The current paper examined these survey responses to provide additional insights into the instruct-first learning benefit from the previous programming study. Our research question was as follows:

RQ: How do reported perceptions of the learning session differ between students who completed the instruct-first versus explore-first conditions?

Differences in survey measures, such as for cognitive load, self-efficacy, or motivation, might help to explain why student learning outcomes did not benefit from exploration. We hypothesized that the exploratory learning activity we created was challenging, increasing cognitive load. If cognitive load is too high, exploratory learning could have the opposite effect than expected [10], [11], [12], [13].

II. Methodology

This study was reviewed and approved by the university institutional review board.

A. Participants

Participants (N = 402) were first-year engineering undergraduate students at the University of Louisville, enrolled in an introductory engineering course in Fall 2022. Participants were included in analyses if they attended class on the day of the experiment and completed all phases of the experiment.

B. Procedures

Students were assigned to one of two conditions based on course section: the explore-first condition and the instruct-first condition. There were six course sections, with three instructors leading two sections each. Condition was counterbalanced across instructors (each instructor was assigned one explore-first and one instruct-first section). One instructor was unavailable on the day of the experiment, and a second instructor covered their sections. Thus, one instructor led four sections, and another led two.

In the *instruct-first condition,* students were provided instruction on interpreting error messages in Python programs, followed by an activity to apply the new concepts. In the *explore-first condition,* students received the same materials in reverse order. First, they explored the novel activity, which was then followed by instruction. After both instruction and activity phases, students completed a survey to assess student attitudes about the activity, followed by an assessment to evaluate their learning of both concepts and procedures taught during the class session.

Students were encouraged to collaborate with peers on the exploration activity, and were instructed to work independently on the assessment. Students received participation credit for attempting the assessment, irrespective of their performance, and were reminded to do their best. All phases of the study (instruction, exploration activity, survey, assessment) were completed within one 50-minute class period.

C. Materials

The class session focused on three Python language-based errors: Name Errors, Syntax Errors, and Type Errors.

Instruction. The three most common error types were described and illustrated with code examples that would generate these error messages. The composition of the error message was discussed, including hints about the type and location of the error in the code (e.g., Name errors can occur when variables are misspelled; Syntax errors frequently result from incorrect punctuation; Type errors indicate an invalid datatype.)

Activity. Students were provided a complete, working Python program that performed a basic task (i.e., the program converted a numerical input from degrees Fahrenheit to Celsius and determined whether the temperature was freezing, cool, warm, or hot based on simple range criteria). Students were instructed to "break" the code by adding, altering, or deleting text to generate varied error messages. Students recorded their findings on group worksheets identifying the error types and causes. Students were encouraged to collaborate on the activity in groups of 2–3 at tables of 4–5.

Survey. Information about each survey measure is included in Table 1. The cognitive load item was taken directly from [15] and assessed on a 9-point scale, ranging from 1 (*very, very low mental effort*) to 9 (*very, very high mental effort*). The other scales were closely adapted or shortened for our study. Items were interspersed and rated on a 5-point Likert scale (1=*strongly disagree*, 5=*strongly agree*).

Assessment. The assessment included 12 questions (multiple choice and short answer). Students were asked to identify the error type generated by the provided sample code or scenarios. Students were presented with code as well as the generated error message and were prompted to identify the error cause. Three of the 12 questions tested students' conceptual understanding, considering plausible explanations

TABLE 1: INFORMATIONAL DETAILS ABOUT SURVEY ITEMS AND MEANS BY CONDITION.

Factor	Number of Items	Cronbach's Alpha	Sample Item	Reference	Order of Instruction	
					Explore-First *M* (*SD*)	Instruct-First *M* (*SD*)
Cognitive Load	1	-	Please indicate how much mental effort you invested when completing the learning activities.	[18]	5.93 (1.54)	5.59 (1.67)
Self-Efficacy	3	.63	I feel confident in my ability to learn these kinds of topics.	[19]	3.77 (.61)	3.96 (.65)
Situational Interest	2	.79	I enjoyed working on these activities	[20]	3.57 (.78)	3.66 (.88)
Flow state	4	.69	I was totally absorbed in what I was doing.	[21]	3.53 (.64)	3.58 (.70)
Security	4	.83	I felt secure to express my ideas.	[22]	3.98 (.59)	4.02(.64)
Belonging Uncertainty	4	.80	After working on today's activities, I feel like I don't belong.	[23]	2.01 (.73)	1.58 (.70)
Competence	2	.76	Thanks to today's learning activities, I feel more competent in this topic area.	[22]	3.58 (.72)	3.74 (.74)
Curiosity	3	.81	I wanted to know more about what I was working on.	[24]	3.49 (.75)	3.47 (.83)
Constructive Engagement	3	.70	I tried to explain key concepts to myself.	[25], [26]	3.84 (.60)	3.82 (.66)
Insight	2	.75	I had a moment of insight.	[27]	3.39 (.84)	3.25 (.90)
Knowledge Gaps	3	.66	I do not feel very knowledgeable about the topic we learned today.	[28]	2.51 (.72)	2.30 (.72)
Behavioral Intentions of Future Performance	2	.79	If given the choice, I would do an activity on this topic again.	[15]	3.50 (.85)	3.56 (.87)

for a given error message, and understanding when and why python generates an error message.

D. Analysis

We assessed differences in survey responses by conducting one-way, between-subjects, analyses of variance (ANOVA) on survey measures (average rating of item responses per measure) with an independent factor of *order* (explore-first, instruct-first). All analyses were conducted using IBM SPSS Statistics with statistical significance set at $p < 0.05$.

III. RESULTS

Descriptive statistics for the survey responses are presented in Table 1. ANOVA results revealed a significant effect of order of instruction on student cognitive load, $F(1, 400) = 4.46$, $p = .035$, $\eta_p^2 = .011$. Students in the explore-first condition reported exerting greater mental effort during the activity compared to those in the instruct-first condition. Means for the explore-first condition were close to 6 on the 9-point scale, indicating "rather high mental effort."

As shown in Fig. 1, students in the explore-first condition also reported higher awareness of knowledge gaps, $F(1, 400) = 8.81$, $p = .003$, η_p^2=.022. Students in the explore-first condition reported lower self-efficacy. $F(1, 400) = 8.85$, $p = .003$, η_p^2=.022, and competence $F(1, 400) = 4.48$, $p = .035$, η_p^2=.011, than students in the instruct-first condition. There were no significant differences on any other survey measures (insight: $F(1, 400) = 2.96$, $p = .086$, all other Fs<1, ps > .338).

IV. DISCUSSION

A number of previous studies have shown learning benefits for students who explore a novel activity prior to instruction [2], [14]. However, a few studies show the opposite—higher learning outcomes when receiving instruction before the activity compared to exploring first (e.g., [15], [16]). This study sought to evaluate differences in student perceptions across conditions that could help to explain or augment learning outcomes in one such prior study [29].

Cognitive load was higher in the explore-first condition than in the instruct-first condition, and indicated "rather high mental effort." This result could suggest that the activity and instruction, paired together, were mentally taxing for students in the exploratory learning condition. Kapur (2016) suggests that exploratory learning activities should be designed so that they are challenging, but not so challenging such that students might give up. Consistent with this idea, some research suggests that exploratory learning activities can be less effective when the cognitive load is too high [10], [11], [12]. The current results align with this idea.

Consistent with the possibility that students who explored were more challenged, students in the explore-first condition also reported feeling less knowledgeable about the topic (higher knowledge gap score) than those in the instruct-first condition. Knowledge gap differences are expected based on prior research. Students who explore-first are thought to become better aware of the existing gaps in their understanding, and seek to fill these gaps through the instruction phase [2], [11], [17]. However, these knowledge gap differences are typically found following the activity, prior to receiving instruction. This

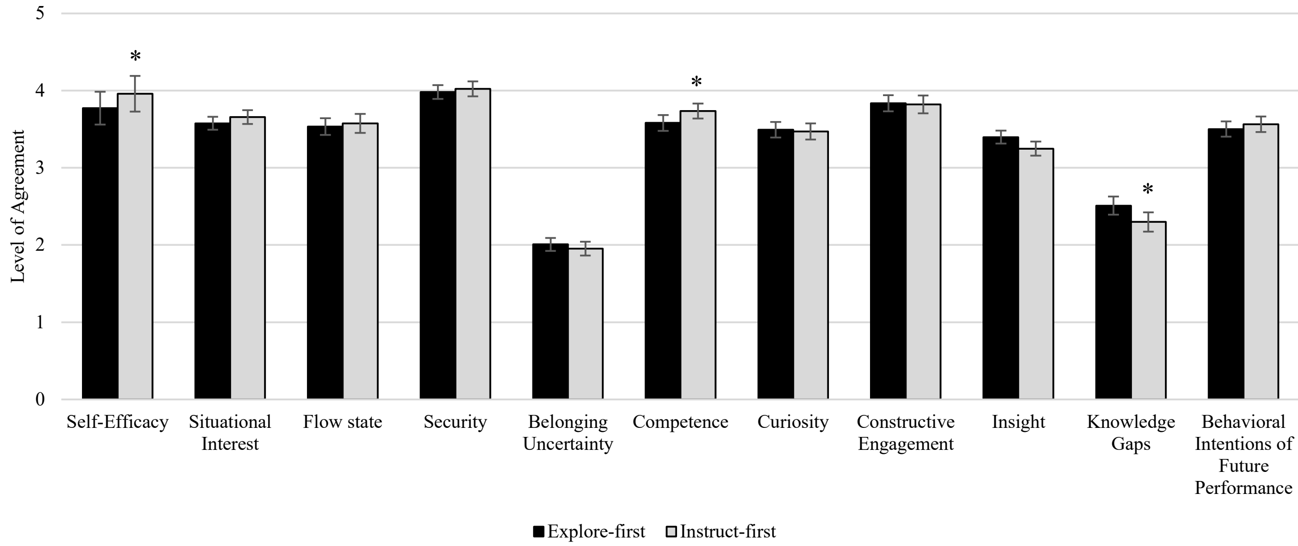

Fig. 1. Mean survey responses for students in each condition (explore-first and instruct-first). Error bars represent the 95% confidence interval from the mean difference comparison. * indicates factors with statistically significant differences in the means for each condition.

gap in knowledge should be filled following the instruction. The survey in this study was given at the end of the learning session, after both the activity and instruction were completed. Additionally, the survey came right after the instruction for students in the explore-first condition. Thus, we would expect students in the explore-first condition to perceive knowledge gaps equally to those in the instruct-first condition by this point. However, we found a significant difference in this study. This finding suggests that the instruction did not help students feel as knowledgeable overall at the end of the learning session, when given exploration followed by instruction. One possibility is that the exploration activity encouraged questions that were unanswered by the instruction. Another possibility is that students in the explore-first condition felt more challenged overall, reducing their confidence in their knowledge.

Consistent with these ideas, students in the explore-first condition reported feeling lower self-efficacy and competence. Items on these scales reflect students' confidence in their ability to learn the topics (self-efficacy), and perceptions that they are competent and able to take on learning challenges in the topic area.

Taken together, the explore-instruct sequence was perceived as challenging to students, and affected their perceptions of their ability. These perceptions likely translated into lower learning scores, although it is also possible that they were the result of lower perceived learning due to other reasons. The challenge may have been due to the topic itself (e.g., it might have involved too many interacting elements [13]). The instruction might have been insufficient to address all the knowledge gaps that arose during instruction. It is also possible that students simply needed more time to really engage with the important problem features. Several students noted in an open-ended comment that they wished they had been given more time.

Importantly, survey responses did not differ between conditions for the other factors measured, including belonging uncertainty, security, flow, constructive engagement, insight, interest, curiosity and behavioral intentions of future performance. These findings suggest that exploring before instruction did not dampen other aspects of students' motivation, nor did students find the activities threatening overall. Given that we asked students to attempt a novel activity before providing any direction about the underlying concepts, these results are assuring.

A. Limitations and Future Work

This study was conducted during one classroom session on one topic. More work is needed to determine if these results generalize to other topics or domains.

Currently, the survey results and learning outcomes are considered in parallel, but a causal link cannot be made between them. Future research should directly manipulate the cognitive load of the materials, to determine if reducing the challenge or number of elements, or adding more time, increases learning from exploration beyond that of the traditional instruction condition.

REFERENCES

[1] M. Kapur, "Productive failure in learning the concept of variance," *Instr Sci*, vol. 40, pp. 651–672, 2012, doi: 10.1007/s11251-012-9209-6.

[2] K. Loibl, I. Roll, and N. Rummel, "Towards a theory of when and how problem solving followed by instruction supports learning," *Educ Psychol Rev*, vol. 29, pp. 693–715, 2017, doi: 10.1007/s10648-016-9379-x.

[3] D. L. Schwartz, C. C. Chase, M. A. Oppezzo, and D. B. Chin, "Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer.," *J Educ Psychol*, vol. 103, pp. 759–775, 2011, doi: 10.1037/a0025140.

[4] M. S. DeCaro and B. Rittle-Johnson, "Exploring mathematics problems prepares children to learn from instruction," *J Exp Child Psychol*, vol. 113, pp. 552–568, 2012, doi: 10.1016/j.jecp.2012.06.009.

[5] D. L. Schwartz and T. Martin, "Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction," *Cogn Instr*, vol. 22, no. 2, pp. 129–184, 2004, doi: http://dx.doi.org/10.1207/s1532690xci2202_1.

[6] M. Kapur and K. Bielaczyc, "Designing for productive failure," *Journal of the Learning Sciences*, vol. 21, pp. 45–83, 2012, doi: 10.1080/10508406.2011.591717.

[7] Z. M. Savelson and K. Muldner, "How do students feel and collaborate during programming activities in the productive failure paradigm?," *Computer Science Education*, 2023, doi: 10.1080/08993408.2023.2237365.

[8] M. T. H. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," *Educ Psychol*, vol. 49, no. 4, pp. 219–243, 2014, doi: 10.1080/00461520.2014.965823.

[9] J. Sweller, P. Ayres, and S. Kalyuga, "Cognitive load theory," *Psychology of learning and motivation.*, vol. 55, pp. 37–76, 2011, doi: 10.1007/978-1-4419-8126-4.

[10] G. Ashman, S. Kalyuga, and J. Sweller, "Problem-solving or explicit instruction: Which should go first when element interactivity is high?," *Educ Psychol Rev*, vol. 32, no. 1, pp. 229–247, 2020, doi: 10.1007/s10648-019-09500-5.

[11] P. M. Newman and M. S. DeCaro, "Learning by exploring: How much guidance is optimal?," *Learn Instr*, vol. 62, 2019, doi: 10.1016/j.learninstruc.2019.05.005.

[12] M. Kapur, "Examining productive failure, productive success, unproductive failure, and unproductive success in learning," *Educ Psychol*, vol. 51, no. 2, pp. 289–299, 2016, doi: 10.1080/00461520.2016.1155457.

[13] P. A. Kirschner, J. Sweller, and R. E. Clark, "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, problem-based, experiential and inquiry-based teaching," *Educ Psychol*, vol. 41, no. 2, pp. 75–86, 2006, doi: 10.1207/s15326985ep4102_1.

[14] A. Darabi, T. L. Arrington, and E. Sayilir, "Learning from failure: A meta-analysis of the empirical studies," *Educational Technology Research and Development*, vol. 66, pp. 1–18, 2018, doi: 10.1007/s11423-018-9579-9.

[15] K. Loibl, M. Tillema, N. Rummel, and T. van Gog, "The effect of contrasting cases during problem solving prior to and after instruction," *Instr Sci*, vol. 2020, pp. 1–22, 2020, doi: 10.1007/s11251-020-09504-7.

[16] E. R. Fyfe, M. S. DeCaro, and B. Rittle-Johnson, "An alternative time for telling: When conceptual instruction prior to problem solving improves mathematical knowledge," *British Journal of Educational Psychology*, vol. 84, no. 3, pp. 502–519, 2014, doi: 10.1111/bjep.12035.

[17] I. Glogger-Frey, C. Fleischer, L. Grüny, J. Kappich, and A. Renkl, "Inventing a solution and studying a worked solution prepare differently for learning from direct instruction," *Learn Instr*, vol. 39, pp. 72–87, 2015, doi: 10.1016/j.learninstruc.2015.05.001.

[18] F. G. Paas, "Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach.," *J Educ Psychol*, vol. 84, pp. 429–434, 1992, doi: 10.1037/0022-0663.84.4.429.

[19] P. R. R. Pintrich, D. Smith, T. Garcia, and W. McKeachie, "A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)," Ann Arbor, MI, 1991. doi: ED338122.

[20] J. I. Rotgans and H. G. Schmidt, "Situational interest and learning: Thirst for knowledge," *Learn Instr*, vol. 32, 2014, doi: 10.1016/j.learninstruc.2014.01.002.

[21] F. Rheinberg, R. Vollmeyer, and S. Engeser, "Flow short scale," *Motiv Emot*, vol. Database record, 2003, doi: 10.1037/t47787-000.

[22] K. M. Sheldon, A. J. Elliot, Y. Kim, and T. Kasser, "What is satisfying about satisfying events? Testing 10 candidate psychological needs," *J Pers Soc Psychol*, vol. 80, no. 2, 2001, doi: 10.1037/0022-3514.80.2.325.

[23] G. M. Walton and G. L. Cohen, "A brief social-belonging intervention improves academic and health outcomes of minority students," *Science (1979)*, vol. 331, no. 6023, pp. 1447–1451, 2011, doi: 10.1126/science.1198364.

[24] F. D. Naylor, "A state-trait curiosity inventory," *Aust Psychol*, vol. 16, no. 2, 1981, doi: 10.1080/00050068108255893.

[25] K. McEldoon, "Supporting novice learning through focused processing of worked examples and explanations," 2014.

[26] M. T. H. Chi *et al.*, "Translating the ICAP theory of cognitive engagement into practice," *Cogn Sci*, no. 2018, pp. 1–56, 2018, doi: 10.1111/cogs.12626.

[27] A. H. Danek and J. Wiley, "What causes the insight memory advantage?," *Cognition*, vol. 205, 2020, doi: 10.1016/j.cognition.2020.104411.

[28] L. R. Flynn and R. E. Goldsmith, "A short, reliable measure of subjective knowledge," *J Bus Res*, vol. 46, no. 1, pp. 57–66, 1999, doi: 10.1016/S0148-2963(98)00057-5.

[29] C. Bego *et al.*, " Exploratory learning in engineering programming.," in *2023 IEEE Frontiers in Education Conference (FIE)*, College Station, TX, Oct. 2023. doi: 10.1109/FIE58773.2023.10343300.