

Investigating learners' Cognitive Engagement in Python Programming using ICAP framework

Daevesh Singh
Indian Institute of Technology Bombay
daeveshsingh@iitb.ac.in

Ramkumar Rajendran
Indian Institute of Technology Bombay
ramkumar.rajendran@iitb.ac.in

ABSTRACT

Introductory programming courses suffer a fate of high failure. Also, research has shown a significant relationship between engagement and academic success. Therefore, reliable and real-time engagement measures could help identify students who need help and provide them with personalised instruction. The present methods of measuring engagement rely primarily on self-reports that have reliability issues due to self-bias and poor recall and, at the same time, do not provide real-time and high granularity data. Interestingly, the online environments allow for real-time capturing of fine-grained interactions, which could be used to measure students' engagement and overcome the issues of self-report measures. The focus of our work will be cognitive engagement, which is a less explored dimension of engagement. To achieve this, we have developed an online learning environment called PyGuru for teaching-learning of Python. We propose to collect learner interaction data, classify these actions into different levels of cognitive-engagement, and study the impact of these different levels on their learning. We present the initial work done in this direction regarding the system developed and the data collected. We intend to seek advice on the validity and reliability of our approach to measuring cognitive engagement.

Keywords

Programming, Cognitive Engagement, PyGuru

1. INTRODUCTION

Research has demonstrated a significant relationship between engagement and educational outcomes like persistence, completion and achievement [7, 9, 18]. While student engagement is crucial to any learning experience, it is particularly important for domains like computer programming since introductory programming courses suffer a fate of high failure and dropouts and engagement as a construct can help in addressing this [2, 21].

D. Singh and R. Rajendran. Investigating learners' cognitive engagement in Python programming using ICAP framework. In A. Mitrovic and N. Bosch, editors, *Proceedings of the 15th International Conference on Educational Data Mining*, pages 789–794, Durham, United Kingdom, July 2022. International Educational Data Mining Society.

© 2022 Copyright is held by the author(s). This work is distributed under the Creative Commons Attribution NonCommercial NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

<https://doi.org/10.5281/zenodo.6852960>

The construct “engagement” has multiple unique conceptualisations. Kuh (2007) defines engagement as “participation in educationally effective practices, both inside and outside the classroom, which leads to a range of measurable outcomes” . Krause and Coates (2008) define it as “the extent to which students are engaging in activities that higher education research has shown to be linked with high-quality learning outcomes” [13, 14].

Engagement as a construct not only requires involvement but also feelings and sense-making [15] as a result, many theorists argue for a multidimensional definition of engagement. In literature, the various dimensions of engagement include academic, behavioural, emotional, affective, psychological, social, cognitive, and agentic dimensions [4, 8, 9, 11]. The focus of our work will be Cognitive Engagement(CE), because of two reasons. Firstly, as it has established research and a theoretical base that supports its significance in learning [3, 9, 10]. Secondly, CE has been less explored in the domain of computer programming [12]. Cognitive engagement is defined as the extent to which students are willing to invest in working on the task [5], and how long they persist [17, 19]. Similarly, as per Fredricks et al.(2004) “*cognitively engaged students would be invested in their learning, would seek to go beyond the requirements, and would relish challenge.*” [9].

Our research uses the framework developed by Chi & Wylie (2014). This framework classifies students' overt behaviours (like taking notes, asking questions etc.) into four different modes of engagement, namely interactive, constructive, active, and passive(ICAP)[3]. ICAP model, which was initially developed for classroom learning, has been extended to online learning as well [6, 16, 20, 22]. However, most of the existing studies in the online environment explore components such as video watching, DF, in pieces (i.e. they apply ICAP to only one of the online learning components like video watching, discussion forums, etc). This implies a lacuna of studies in a Computer-Based Learning Environment (CBLE) that study all these components together and hence, lacks an understanding of how students' behaviour in CBLE is related to their CE. Also, most studies that explore engagement in computer programming mainly focus on the motivational and behavioural aspect and not the cognitive aspect of engagement[12]. Hence, there is a need of analysing actions learners' perform in CBLE with all the existing components to better understand cognitive engagement and its impact on learning .

To fill the above gaps, we developed a CBLE called PyGuru for teaching-learning of Python programming that consists of four components: book-reader, video player, Integrated Development Environment(IDE), and discussion forum. This learning environment captures fine-grained data related to learning like highlights and annotations in the book-reader. Play, pause and seek behaviour and response to in-video questions in a video player. Views, comments, and likes in the discussion forum and their performance in IDE using test cases for each program.

In our research, we propose to use the interaction data generated to classify students' actions into different modes of cognitive engagement, namely Interactive, Constructive, Active and Passive (ICAP). We will then study the impact of these different engagement levels on students learning of computer programming. These results will help us develop models to measure learners' cognitive engagement and provide them feedback to prevent failure and dropouts.

In this paper, Section II presents the background and related works. The description of the learning environment and study design are provided in section III. Section IV offers a proposed solution, and the closing section contains questions for which we intend to seek advise.

2. BACKGROUND AND RELATED WORK

To analyze learners' interaction in the system, we will be using the ICAP framework developed by Chi Wylie (2014). This section consists of two subsections. The first one provides a brief introduction to the ICAP framework and the second subsection involves the related works that use the ICAP framework.

2.1 ICAP Framework

ICAP is a hierarchical framework that defines engagement in terms of students' overt behaviours and tries to distinguish these overt behaviours into four modes of engagement, namely interactive, constructive, active and passive[3].

They define passive mode as *"learners receiving information without overtly doing anything related to learning"* (Chi Wylie, 2014, p. 221). In this mode, learners simply attend to the information (without performing any actions like note-taking) and store it in episodic form rather than integrating it with prior knowledge.

Active engagement occurs when the learner's information acquisition is accompanied by certain physical or motoric actions that support their learning. This includes taking notes in the classroom, pausing and recap of the videos being watched or highlighting the text while reading.

The cognitive process involved during active engagement demands the activation of prior knowledge and integrating the new information into the existing one. Constructive engagement happens when the learners attempt to produce some artefacts using their prior knowledge and the information available in the environment. The characteristic feature of these artefacts is that they use information that goes beyond the available information. This includes "elaborating, comparing and contrasting, generalizing, reflecting on, and explaining how something works" (Chi Wylie, 2014, p. 228).

Interactive engagement occurs when learners engage in interaction, and during these interactions, both partners must be constructively engaged and there must be sufficient turn-taking.

2.2 Studies based on ICAP

Yogev et al. (2018) examined students' CE in reading material using Nota Bene annotation platform[22]. They firstly analyzed CE anchored in the text by manually labelling students' annotations and then developed an interactive decision tree to automate this process. They found that different sections of the text elicit different levels of CE. For instance, low CE corresponds to definitions provided in the text. They also developed a visualization tool to show the distribution of varying levels of CE over the reading material.

Acknowledging the absence of a framework for active viewing(active learning to describe students' behaviours while learning from video), Dodson et al. (2018) developed a framework to classify students' different video watching behaviours as per the ICAP framework which classified behaviours as passive(playing video content), active(replay, pause, seek specific information), constructive(taking notes, highlighting), and interactive (cooperating and collaborating with others) [6]. To fully support video-based learning, they used ViDeX in this study which supported a broader set of behaviours, as mentioned previously.

Wang et al. (2015) investigated students' contributions in the discussion forum participating in a MOOC[20]. They studied the relationship between the kind of posts students write and their learning gain. For this purpose, they considered each post as a sampling unit and coded each post into nine categories: active (note-taking, repeat, paraphrase), constructive (compare or connect, ask novel questions, provide justification or reason), and interactive (building on partner's contribution, acknowledging partner's contribution, defend or challenge). Their study shows a significant association between the discourse in the discussion forum and learning gains.

The study by Atapattu et al. (2019) tried to automate this process of classifying the post in the discussion forum into the active and constructive modes of engagement[1]. Their research included finding semantic similarities between each post and related course materials using cosine similarity. The posts significantly different from the learning materials were classified as constructive, and the more similar ones were labelled as active.

To summarise, we have presented how ICAP is used in book-reader, video-watching and discussion forums. These studies have influenced the design of our system PyGuru. Moreover, the results presented in these studies are promising and indicate a strong link between students' CE and their performance. It is, therefore, crucial to investigate how overall learning behaviour (actions in book reader, video player etc.) impact students' performance in Python programming. To understand this, we present PyGuru- a learning environment for learning Python programming capable of logging the user actions. The following section offers details about the system, study design, and the proposed solution.

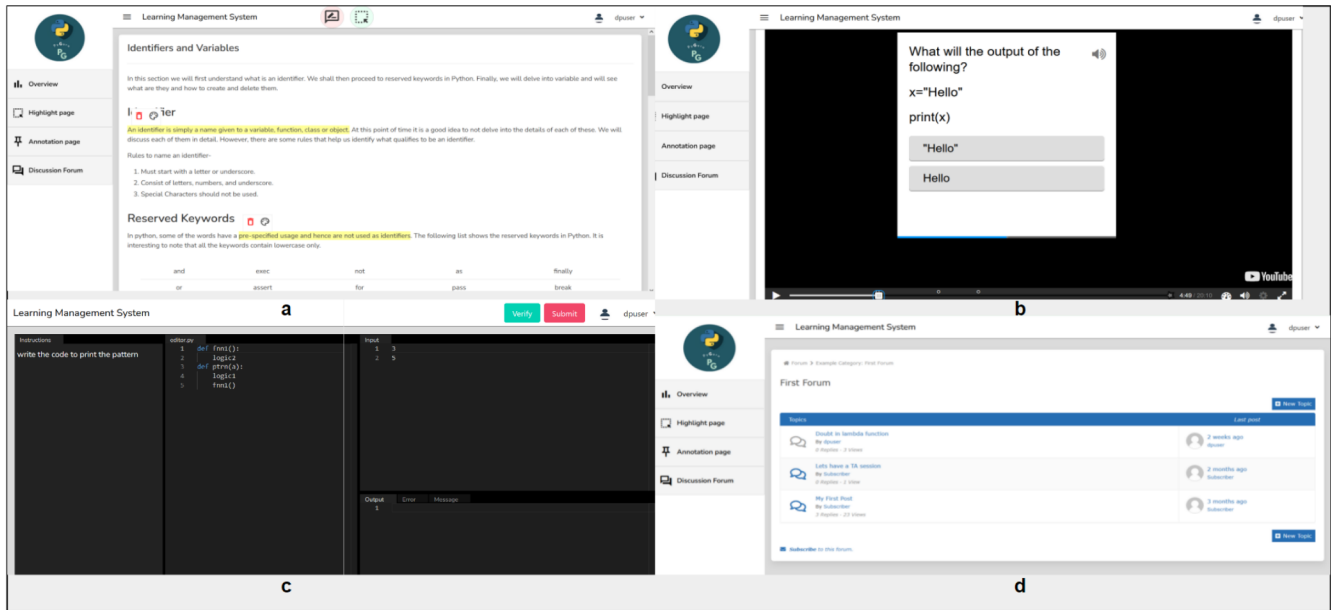


Figure 1: PyGuru Environment (a) Book Reader (b) Video Player (c) IDE (d) Discussion Forum

3. LEARNING ENVIRONMENT

PyGuru¹ is a computer-based learning environment for teaching and learning Python programming skills. PyGuru has four major components: 1) Book Reader 2) Video Player 3) Integrated Development Environment (IDE) 4) Discussion Forum.

A book reader (shown in Fig 1. a) in PyGuru allows readers to highlight and annotate the text. Highlight in the digital context consists of selecting a text and colouring it. The annotating feature in the learning environment comprises selecting a text, commenting on that text, and providing a tag to that text.

PyGuru has an interactive video watching platform (Fig. 1.b). The students can interact with the video using basic video player features like enhancing the speed of the video and performing other actions like play, pause and seek. More advanced interactive features are embedded into the system that allows the instructor to add questions within the video. The video automatically stops and waits for the learner's response.

We also have integrated an IDE (Fig 1. c) in PyGuru. Currently, the IDE evaluates the learners' code against the test cases. The "verify" button in the IDE allows learners' to check their program for errors and test cases before submitting. The IDE consist of four panels 1) Instruction Panel -To provide details about the problem or algorithm 2) Input Panel - To provide the test cases for the problem, and these test cases are tested against these inputs 3) Coding Panel - The coding panel is where the student is expected to write the code, and the instructor can present some partial codes as well 4) Output Panel - To display the output once the program is run and provide information about the number

¹<https://tinyurl.com/2p8h6zpw>

of test cases passed and failed.

PyGuru has a discussion Forum (Fig 1.d). The learners can use this discussion forum to put forth their queries which the instructor or their peers can then answer. They can also respond and hit like on the post made by their peers or instructor.

4. METHODOLOGY

The study is focused on investigating learners' cognitive engagement in a CBLE, designed for teaching and learning of Python programming. We decided to use ICAP framework not only to design the system but will also use it to analyze learners' interactions in PyGuru. The aim is to answer the following questions: 1.1 Is there a relationship between log data of student activity in CBLE and self-reported student engagement survey scores? 1.2 Is there a relationship between actions that are believed to represent higher cognitive engagement and learning gains? 1.3 How do learners' actions having different learning gains differ in a CBLE while learning computer programming?

4.1 Study Design

The study participants will be students from undergraduate colleges or diploma colleges located in developing countries (India and Malaysia). The duration of the study will be 3-5 weeks. During this time the learners will be interacting with the system to learn programming by reading content, watching videos, and coding on IDE. One week prior to the study, learners will be introduced to the system. Along with this a consent form, demographic survey and a pre-test will be administered. This pre-test will be designed by the instructor and on similar lines a post-test will be designed to measure learning gains. Student engagement will be measured using the 7-item activity-level student engagement survey (Henrie et al. 2016). From this survey,

we will only consider items corresponding to cognitive engagement. Log data from the PyGuru will be collected for all participants. The log data consists of the user id, session id, timestamp, page id, and other details like the action performed (e.g. highlighted or annotated) and the context information (the text highlighted).

4.2 Proposed Solution

We propose to investigate cognitive engagement in CBLE using log-data. At this point, there are two directions. One is to do a correlation analysis with the engagement scores obtained through engagement surveys and identify whether the data captured from logs could be used as a proxy for cognitive engagement. The second one is to use a data-driven approach to identify the characteristics of students with higher learning gains using descriptive and diagnostic analytics. Further, using the findings obtained from descriptive and diagnostic analytics to develop regression models to measure cognitive engagement.

5. ADVICE SOUGHT

Q1. Are the system's design and the data captured suitable for measuring engagement?

Q2. Since we plan to develop regression models for measuring engagement using learning gains. Is this proposed approach valid? Also, what are other ways to validate our measures other than self-report or teachers' ratings?

Q3. What are the general suggestions regarding the study design in terms of duration of the study, the data collected from the environment, etc.?

6. ACKNOWLEDGMENTS

I would like to thank Ashwin TS, Deepak Pathak and Rumana Pathan for useful discussions and their support. This paper is based upon work supported by the grant from Council of Scientific and Industrial Research(CSIR).

7. REFERENCES

- [1] T. Atapattu, M. Thilakaratne, R. Vivian, and K. Falkner. Detecting cognitive engagement using word embeddings within an online teacher professional development community. *Computers & Education*, 140:103594, 2019.
- [2] C. Beise, L. VanBrackle, M. Myers, and N. Chevli-Saroq. An examination of age, race, and sex as predictors of success in the first programming course. 2003.
- [3] M. T. Chi and R. Wylie. The icap framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist*, 49(4):219–243, 2014.
- [4] S. Christenson, A. L. Reschly, C. Wylie, et al. *Handbook of research on student engagement*, volume 840. Springer, 2012.
- [5] L. Corno and E. B. Mandinach. The role of cognitive engagement in classroom learning and motivation. *Educational psychologist*, 18(2):88–108, 1983.
- [6] S. Dodson, I. Roll, M. Fong, D. Yoon, N. M. Harandi, and S. Fels. An active viewing framework for video-based learning. In *Proceedings of the fifth annual ACM conference on learning at scale*, pages 1–4, 2018.
- [7] J. D. Finn and J. Owings. *The adult lives of at-risk students the roles of attainment and engagement in high school—statistical analysis report*. DIANE Publishing, 2006.
- [8] J. D. Finn and K. S. Zimmer. Student engagement: What is it? why does it matter? In *Handbook of research on student engagement*, pages 97–131. Springer, 2012.
- [9] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris. School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74(1):59–109, 2004.
- [10] C. R. Henrie, R. Bodily, R. Larsen, and C. R. Graham. Exploring the potential of lms log data as a proxy measure of student engagement. *Journal of Computing in Higher Education*, 30(2):344–362, 2018.
- [11] C. R. Henrie, L. R. Halverson, and C. R. Graham. Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90:36–53, 2015.
- [12] G. Kanaparan, R. Cullen, D. Mason, et al. Effect of self-efficacy and emotional engagement on introductory programming students. *Australasian Journal of Information Systems*, 23, 2019.
- [13] K.-L. Krause and H. Coates. Students' engagement in first-year university. *Assessment & Evaluation in Higher Education*, 33(5):493–505, 2008.
- [14] G. D. Kuh. What student engagement data tell us about college readiness. 2007.
- [15] S. J. Quaye, S. R. Harper, and S. L. Pendakur. *Student engagement in higher education: Theoretical perspectives and practical approaches for diverse populations*. Routledge, 2019.
- [16] M. Raković, Z. Marzouk, A. Liaqat, P. H. Winne, and J. C. Nesbit. Fine grained analysis of students' online discussion posts. *Computers & Education*, 157:103982, 2020.
- [17] J. C. Richardson and T. Newby. The role of students' cognitive engagement in online learning. *American Journal of Distance Education*, 20(1):23–37, 2006.
- [18] V. Trowler. Student engagement literature review. *The higher education academy*, 11(1):1–15, 2010.
- [19] C. O. Walker, B. A. Greene, and R. A. Mansell. Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and individual differences*, 16(1):1–12, 2006.
- [20] X. Wang, D. Yang, M. Wen, K. Koedinger, and C. P. Rosé. Investigating how student's cognitive behavior in mooc discussion forums affect learning gains. *International Educational Data Mining Society*, 2015.
- [21] C. Watson and F. W. Li. Failure rates in introductory programming revisited. In *Proceedings of the 2014 conference on Innovation & technology in computer science education*, pages 39–44, 2014.
- [22] E. Yogev, K. Gal, D. Karger, M. T. Facciotti, and M. Igo. Classifying and visualizing students' cognitive engagement in course readings. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, pages 1–10, 2018.

Presentation Letter

Name: Daevesh Singh

Institute: Indian Institute of Technology Bombay (IITB), Mumbai, India

Supervisor: [Prof. Ramkumar Rajendran](#)

Title: Measuring learner's Cognitive Engagement in Python Programming using log-data

Introduction: I am Daevesh Singh, a third year research scholar at the Educational Technology Department, IITB. My research area is educational data analytics. All my research work till now has been focused on using log-data from learning environments and their analysis. Some of my published works include-

1. Nishane, I., Sabanwar, V., Lakshmi, T. G., **Singh, D.**, & Rajendran, R. (2021, July). Learning about learners: Understanding learner behaviours in software conceptual design TELE. In *2021 International Conference on Advanced Learning Technologies (ICALT)* (pp. 297-301). IEEE.
2. Satavlekar, S., Nath, D., Priyadarshini, R., Prasad, P., **Singh, D.**, & Rajendran, R. (2021, July). Unraveling Learner Interaction Strategies in VeriSIM for Software Design Diagrams. In *2021 International Conference on Advanced Learning Technologies (ICALT)* (pp. 308-310). IEEE.
3. **SINGH, D.**, PATHAN, R., BANERJEE, G., & RAJENDRAN, R. From Hello to Bye-Bye: Churn Prediction in English Language Learning App.

About Thesis: My PhD thesis is focussed on measuring cognitive engagement using log-interaction data. The domain includes introductory programming. The study will involve diploma or undergraduate students that are learning Python as a part of their curriculum from the Indian Subcontinent (Mostly India and Malaysia) . For this purpose, we have developed a Computer Based Learning Environment called PyGuru. PyGuru consists of four components, namely the book-reader, video-player, Integrated Learning environment(IDE), and discussion Forum and the environment logs all the interactions made by the learner. We plan to use this data to come-up with meaningful measures of cognitive engagement.

EDM researchers:

1. Neil Hefferman - Neil's research involves computer-based learning environments and engagement.

2. Tanja Mitrovic - Tanja's work in active video watching could be helpful since my environment also has an interactive video player.
3. Sharon Hsiao - Sharon's work includes the use of learning analytics in the domain of computer programming.