



#### **Exploring Novice Programming Behavior over Time** Joe Dillane **Ioannis Karvelas Brett A. Becker University College Dublin University College Dublin University College Dublin** Dublin, Ireland Dublin, Ireland Dublin, Ireland ioannis.karvelas@ucdconnect.ie joe.dillane@itcarlow.ie brett.becker@ucd.ie

## Abstract

This work focuses on the effect that programming time has on novice programmers' interaction with two versions of the BlueJ programming environment that differ in compilation mechanism and error message presentation. We utilize programming process data from users of both BlueJ versions, and those who used each one exclusively. Results indicate there is a threshold of approximately twenty hours of programming below and above which compilation and error message frequency changes. This indicates that being exposed to an environment for prolonged time can influence the programming interaction between novices and environments. This phenomenon is not yet understood.

## **1. Introduction**

This work is part of an ongoing investigation [1, 2] into how novices interact with two versions of the BlueJ pedagogical environment that differ substantially in the mechanisms they employ. Here, we focus on the effect that programming time has on compilation-related interaction of users with BlueJ 3 and BlueJ 4. BlueJ 3 features a click-to-compile mechanism and enforced first-error-message only standard presentation. BlueJ 4 facilitates automatic background compilation and on-demand error message display for all errors present in code (see Figure 1). Our aim is to determine how these environmental differences impact novice programming behaviour.





Figure 1. BlueJ 3 (left) and BlueJ 4 (right) output panes. BlueJ 4 pops up an error message when users hover over an offending code area, whereas BlueJ 3 shows only the first error message at the bottom of the pane.

## 2. Methodology

We analyzed Blackbox programming process data from users belonging to five distinct groups. These groups include users who:

- exclusively used BlueJ 3
- 2. exclusively used BlueJ 4
- 3. started using BlueJ 3 and moved to BlueJ 4
- 4. started using BlueJ 4 and moved to BlueJ 3

Figure 3. Line plot describing Displayed Compiler Error Messages per Hour (DCEMpH) and manual Compilations per Hour (CpH) of all user cohorts.

# 4. Discussion & Moving Forward

#### **Version - specific observations**

### In BlueJ 3:

- 1. Students see half the error messages than they click compile. Their compilation success rate is stable at around 50% regardless of the amount of programming time they have spent using BlueJ.
- Gradual decline in both displayed error messages and compilations for <20 hours of programming for all groups. For >20 hours, this decline is negligible.
- Hypothesis: If there was an improvement in the students' activity regarding syntax errors,

Max amount of CEMs (equals the number of compilations)

5. repeatedly switched between the BlueJ 3 and BlueJ 4

We measure two features to investigate user interaction with BlueJ: **Displayed Compiler Error Messages per Hour (DCEMpH)** and **manual Compilations per Hour** (CpH). We categorized users into total programming time subgroups ranging from 0 to 50 hours and investigated how the mean feature values change with elapsed programming time.

## **3. Results**

Figure 2 describes mean manual compilation and displayed error messages over time for users in the five different cohorts based on their use of BlueJ versions. Figure 3 provides an aggregate picture of all user cohorts.



there should be a gradual decrease in displayed compiler error messages over time as shown in Figure 4. In the beginning, there would be a match between error messages and compilations (since students can't see more messages than they compile), and gradually that difference would increase, until the error message line eventually dropped close to 0. The closer these two lines are, the more the compilation success rate decreases. If these two are farther apart, then it means that students encounter errors less frequently. However, this is not what is happening in



Programming Time **Figure 4.** Error decline hypothesis graph.

# In BlueJ 4:

BlueJ 3.

- 1. Users who transitioned between BlueJ versions show a synchronization between shown error messages and manual compilations from the beginning of their programming activity, whereas those who used only one of the two versions show this behaviour after 20 hours.
- 2. There is a decrease in manual compilations and shown messages as time goes by, similar to BlueJ 3.
- 3. For users who exclusively used one of the two versions or kept switching between versions, the drop of error messages in BlueJ 4 at 20 hours is more vivid.

#### **General Discussion**

Figure 2. Line plots describing Displayed Compiler Error Messages per Hour (DCEMpH) and manual Compilations per Hour (CpH) of each user cohort for the two BlueJ versions. The discontinuous BlueJ 4 lines in the bottom right graph are the result of the 25–30-hour user subgroup being empty.

Results indicate that in BlueJ 3, users generally manually compile more frequently than in BlueJ 4 as shown in Figures 1 & 2. Also, as time progresses towards an approximate threshold of 20 hours, compilations and displayed error messages drop and stabilize for all users. Additionally, for times <20 minutes, BlueJ 4 DCEMpH is higher than BlueJ 3. However, for times >20 minutes, they are approximately equal. This effect is more pronounced for users who programmed only in BlueJ 4 compared to other groups, suggesting that this is related to whether users interact with both versions or not. Future work will investigate these observations to uncover additional evidence on the role of programming environment design in influencing novice programming behavior.

## Acknowledgments

The work presented in this poster is funded by the Irish Research Council under grant agreement GOIPG/2020/1660.

## References

[1] Ioannis Karvelas, Joe Dillane, and Brett A. Becker. 2020. Compile Much? A Closer Look at the Programming Behavior of Novices in Different Compilation and Error Message Presentation Contexts. In United Kingdom & Ireland Computing Education Research Conference. (Glasgow, United Kingdom) (UKICER '20). Association for Computing Machinery, New York, NY, USA, 59–65. https://doi.org/10.1145/3416465.3416471

[2] Ioannis Karvelas, Joe Dillane, and Brett A. Becker. 2020. Compiler Error Messages: Their Content and Accessibility in Novice Programming Environments. In Proceedings of the 51st ACM Technical Symposium on Computer Science Education (Portland, OR, USA) (SIGCSE '20). Association for Computing Machinery, New York, NY, USA, 1310. https://doi.org/10.1145/3328778.3372617