Designing a Real-Time Intervention to Address Negative Self-Assessments While Programming

Melissa Chen[1] and Eleanor O'Rourke[2]

[12] Computer Science, Northwestern University [2] Learning Sciences, Northwestern University

[1] melissac@u.northwestern.edu [2] eorourke@northwestern.edu

Motivation & Design

We propose a **real-time, IDE-based intervention** that **delivers feedback** to students and seek to **understand their perceptions of the tool**

Unrealistic expectations of the programming process [1]

↓

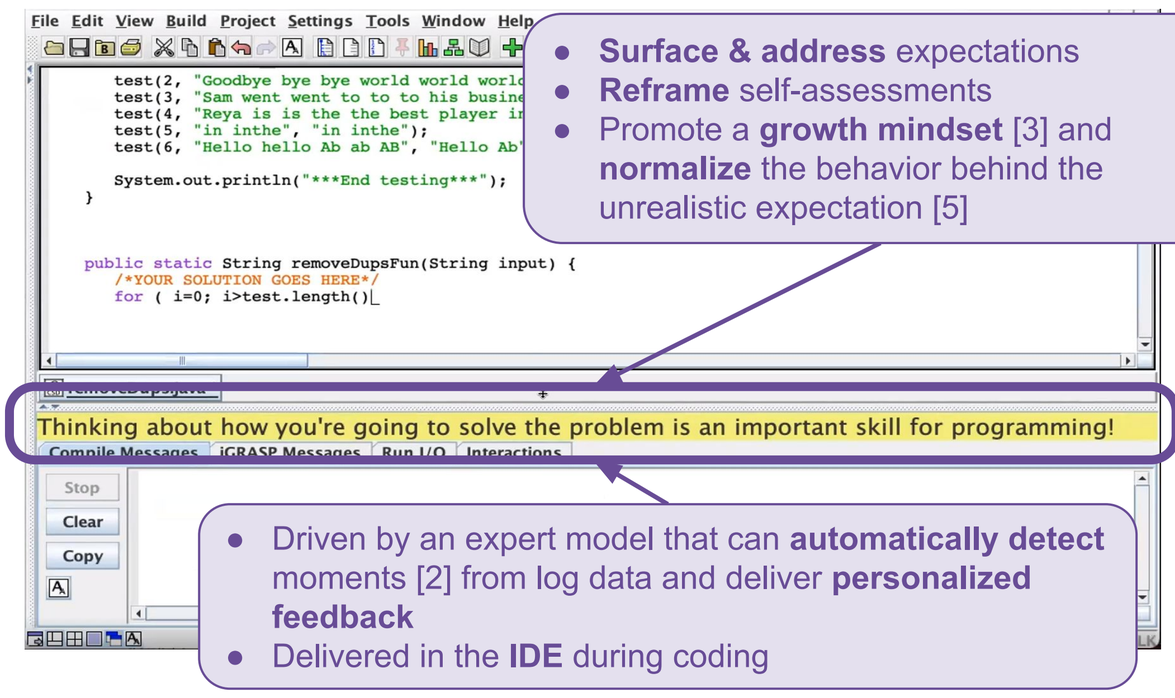
Frequent and self-critical self-assessments [1]

↕

Lower self-efficacy [1]

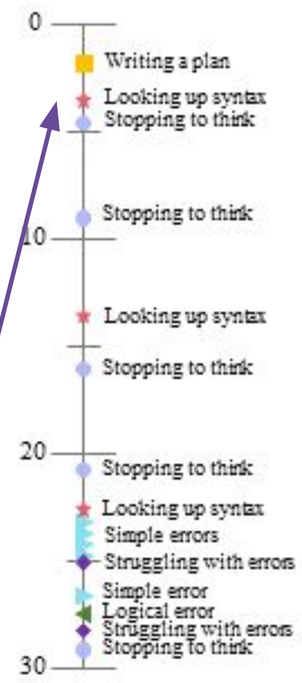
↕

Lower persistence in undergrad programs [4]



- **Surface & address** expectations
- **Reframe** self-assessments
- Promote a **growth mindset** [3] and **normalize** the behavior behind the unrealistic expectation [5]

- Driven by an expert model that can **automatically detect** moments [2] from log data and deliver **personalized feedback**
- Delivered in the **IDE** during coding



"*Yeah I definitely feel bad when I have to spend time planning and can't start programming right away* [...] **we should implement it and we should write it whatever way we feel and by running the program, we can get to know the error instead of wasting time in the beginning.**"

– P8 from Gorson and O'Rourke's study demonstrates a critical self-assessment (*italicized*) based on unrealistic expectations about the programming process (**bolded**) [1]

Fig. 1: *a timeline of a student's self-assessment moments during a 30-minute programming session (right), an example of a self-assessment (bottom), and an example feedback message in the jGRASP IDE addressing said assessment (right)*

Formative User Study Design & Results

Takeaway: students **perceive** that the feedback was **supportive** but **many prefer content-oriented feedback over emotional support** from an AI feedback system.

- 7 CS1 and 11 CS2 students from a large public university in western United States
- Semi-structured interviews
  - Pre/post-test
  - 20 minute coding session
  - Video-guided retrospective interview
- Retrospective was transcribed and open-coded

Surprisingly: The pretest showed most participants **didn't have unrealistic expectations** of the programming process!

Even though previous research suggested that it is common [1]

Future work: understanding how students' expectations change with experience

Students **perceived** that the messages were **designed to be supportive**
(14 students)

"Maybe [the message was shown] **to motivate a student** working on the thing and not, you know, **even if you're struggling, it's fine to struggle**. Maybe that's why."
- P12 (CS2 student) with few negative self-assessment moments

Fewer negative self-assessments → feedback resonated less
(16 students)

"I saw that there was a message trying to encourage me and I was like, okay. **And then I ignored it.**"
- P9 (CS2 student) with few negative self-assessment moments

More negative self-assessments or unproductive struggle → feedback resonated more because they were **perceived to be actionable**
(2 students)

"I think **I slowed down a little bit** because I realized that I did jump into the code pretty fast"
- P17 (CS1 student) with many negative self-assessment moments

Students are **surprised** by the AI's capabilities and perceive both the detection & messages as **"human-like"**
(9 students)

"It did kind of seem almost **human-like** at some points, which I thought was pretty interesting."
- P5 (CS2 student)

They believe **emotional support is less beneficial** coming from an AI system
(5 students)

Instead, they want **content-focused feedback**
(5 students)

"I feel like it's **slightly less beneficial coming from a robot** just because, you know, it's a robot."
- P17 (CS1 student)

"I think I generally find **a message about like specifics in the code versus just kind of being a general like encouragement** to be more useful."
- P15 (CS1 student)

Also, they report that the **immediacy & constancy** of the system felt **off-putting**
(4 students)

"I was just looking up something and then I [saw the message about] looking up syntax and important skills, so I was like, hmm, **that's a little bit creepy**."
- P10 (CS2 student)

Future Work

- Run user studies with students who have more unrealistic expectations
- Include feedback about the programming process
- Summative evaluation through a long-term deployment to measure change in self-assessments

Acknowledgements

This project is supported by the National Science Foundation under grant ISS-2045809. We thank Larry Barowski from jGRASP for his technical support and the Delta Lab for their valuable feedback.

References

[1] Jamie Gorson and Eleanor O'Rourke. 2020. Why do CS1 Students Think They're Bad at Programming?: Investigating Self-efficacy and Self-assessments at Three Universities. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, ACM, Virtual Event New Zealand, 170–181. DOI:https://doi.org/10.1145/3372782.3406273 [2] Jamie Gorson, Nicholas LaGrassa, Cindy Hsinyu Hu, Elise Lee, Ava Marie Robinson, and Eleanor O'Rourke. 2021. An Approach for Detecting Student Perceptions of the Programming Experience from Interaction Log Data. In *Artificial Intelligence in Education*, Ido Roll, Danielle McNamara, Sergey Sosnovsky, Rose Luckin and Vania Dimitrova (eds.). Springer International Publishing, Cham, 150–164. DOI:https://doi.org/10.1007/978-3-030-78292-4_13 [3] John Hattie and Helen Timperley. 2007. The Power of Feedback. *Review of Educational Research* 77, 1 (March 2007), 81–112. DOI:https://doi.org/10.3102/003465430298487 [4] Colleen M. Lewis, Ken Yasuhara, and Ruth E. Anderson. 2011. Deciding to major in computer science: a grounded theory of students' self-assessment of ability. In *Proceedings of the seventh international workshop on Computing education research* (ICER '11), Association for Computing Machinery, New York, NY, USA, 3–10. DOI:https://doi.org/10.1145/2016911.2016915 [5] Valerie J. Shute. 2008. Focus on Formative Feedback. *Review of Educational Research* 78, 1 (March 2008), 153–189. DOI:https://doi.org/10.3102/0034654307313795