REFERENCES

[1] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. 2004. An integrated theory of the mind. *Psychological review* 111, 4 (2004), 1036.

[2] John R Anderson, Frederick G Conrad, and Albert T Corbett. 1989. Skill acquisition and the LISP tutor. *Cognitive Science* 13, 4 (1989), 467–505.

[3] Austin Cory Bart, Javier Tibau, Eli Tilevich, Clifford A Shaffer, and Dennis Kafura. 2017. BlockPy: An Open Access Data-Science Environment for Introductory Programmers. *Computer* 50, 5 (2017), 18–26.

[4] Brett A Becker. 2015. *An exploration of the effects of enhanced compiler error messages for computer programming novices.* Master's thesis. Dublin Institute of Technology.

[5] Benjamin S Bloom. 1984. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher* 13, 6 (1984), 4–16.

[6] Gary M Brosvic and Beth D Cohen. 1988. The horizontal-vertical illusion and knowledge of results. *Perceptual and motor skills* 67, 2 (1988), 463–469.

[7] Ricardo Caceffo, Steve Wolfman, Kellogg S. Booth, and Rodolfo Azevedo. 2016. Developing a computer science concept inventory for introductory programming. In *Proceedings of the 47th ACM Technical Symposium on Computer Science Education.* ACM, 364–369.

[8] Jacob Cohen. 1988. Statistical power analysis for the behavioral sciences. 2nd.

[9] Albert T Corbett and John R Anderson. 2001. Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 245–252.

[10] Tyne Crow, Andrew Luxton-Reilly, and Burkhard Wuensche. 2018. Intelligent tutoring systems for programming education: a systematic review. In *Proceedings of the 20th Australasian Computing Education Conference.* ACM, 53–62.

[11] Roberta E Dihoff, Gary M Brosvic, and Michael L Epstein. 2003. The role of feedback during academic testing: The delay retention effect revisited. *The Psychological Record* 53, 4 (2003), 533–548.

[12] Luke Gusukuma, Austin Cory Bart, Dennis Kafura, Jeremy Ernst, and Katherine Cennamo. 2018. Instructional Design+ Knowledge Components: A Systematic Method for Refining Instruction. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education.* ACM, 338–343.

[13] Georgiana Haldeman, Andrew Tjang, Monica Babeş-Vroman, Stephen Bartos, Jay Shah, Danielle Yucht, and Thu D Nguyen. 2018. Providing Meaningful Feedback for Autograding of Programming Assignments. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education.* ACM, 278–283.

[14] Andrew Head, Elena Glassman, Gustavo Soares, Ryo Suzuki, Lucas Figueredo, Loris D'Antoni, and Björn Hartmann. 2017. Writing Reusable Code Feedback at Scale with Mixed-Initiative Program Synthesis. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale.* ACM, 89–98.

[15] Lisa Kaczmarczyk, Elizabeth Petrick, J Philip East, and Geoffrey L Herman. 2010. Identifying student misconceptions of programming. In *Proceedings of the 41st ACM technical symposium on Computer science education.* ACM, 107–111.

[16] Dennis Kafura, Austin Cory Bart, and Bushra Chowdhury. 2015. Design and Preliminary Results From a Computational Thinking Course. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '15).* ACM, 63–68.

[17] Donald E Knuth. 1969. The art of computer programming. Vol. 1: Fundamental algorithms. Second printing.

[18] Kenneth R Koedinger, Vincent Aleven, Neil Heffernan, Bruce McLaren, and Matthew Hockenberry. 2004. Opening the door to non-programmers: Authoring intelligent tutor behavior by demonstration. In *International Conference on Intelligent Tutoring Systems.* Springer, 162–174.

[19] Kenneth R Koedinger, Albert T Corbett, and Charles Perfetti. 2012. The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive science* 36, 5 (2012), 757–798.

[20] Einari Kurvinen, Niko Hellgren, Erkki Kaila, Mikko-Jussi Laakso, and Tapio Salakoski. 2016. Programming misconceptions in an introductory level programming course exam. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education.* ACM, 308–313.

[21] Nguyen-Thinh Le. 2016. A classification of adaptive feedback in educational systems for programming. *Systems* 4, 2 (2016), 22.

[22] David J Nicol and Debra Macfarlane-Dick. 2006. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education* 31, 2 (2006), 199–218.

[23] Chris Piech, Jonathan Huang, Andy Nguyen, Mike Phulsuksombati, Mehran Sahami, and Leonidas Guibas. 2015. Learning program embeddings to propagate feedback on student code. *arXiv preprint arXiv:1505.05969* (2015).

[24] Thomas W Price, Yihuan Dong, and Tiffany Barnes. 2016. Generating Data-driven Hints for Open-ended Programming.. In *EDM.* 191–198.

[25] Thomas W Price, Yihuan Dong, and Dragan Lipovac. 2017. iSnap: Towards Intelligent Tutoring in Novice Programming Environments. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education.* ACM, 483–488.

[26] Kelly Rivers and Kenneth R Koedinger. 2017. Data-driven hint generation in vast solution spaces: a self-improving python programming tutor. *International Journal of Artificial Intelligence in Education* 27, 1 (2017), 37–64.

[27] Takayuki Sekiya and Kazunori Yamaguchi. 2013. Tracing quiz set to identify novices' programming misconceptions. In *Proceedings of the 13th Koli Calling International Conference on Computing Education Research.* ACM, 87–95.

[28] Valerie J Shute. 2008. Focus on formative feedback. *Review of educational research* 78, 1 (2008), 153–189.

[29] Teemu Sirkiä and Juha Sorva. 2012. Exploring programming misconceptions: an analysis of student mistakes in visual program simulation exercises. In *Proceedings of the 12th Koli Calling International Conference on Computing Education Research.* ACM, 19–28.

[30] Juha Sorva and Teemu Sirkiä. 2011. Context-sensitive guidance in the UUhistle program visualization system. In *Proceedings of the 6th Program Visualization Workshop (PVWâĂŹ11).* 77–85.

[31] JC Spohrer and Elliot Soloway. 1986. Alternatives to construct-based programming misconceptions. In *Acm sigchi bulletin*, Vol. 17. ACM, 183–191.

[32] Marieke Thurlings, Marjan Vermeulen, Theo Bastiaens, and Sjef Stijnen. 2013. Understanding feedback: A learning theory perspective. *Educational Research Review* 9 (2013), 1–15.