Exploratory Data Analysis of Reflective Writing in Scratch Projects

Paulina Haduong, Harvard Graduate School of Education, haduong@g.harvard.edu

**Abstract:** Getting Unstuck was a July 2018 summer professional learning experience for K–12 educators teaching with Scratch, designed to support teachers in developing greater familiarity with Scratch. In addition to creating Scratch projects, teachers were also asked to share them to the online community and submit written reflections in the project notes. In this paper, I first describe Getting Unstuck and the role of reflection in writing. I then describe some of the exploratory data analysis conducted on these Scratch projects. Finally, I suggest some of the future directions for the work.

# Introduction

Demand for K-12 computer science (CS) is at an all-time high, and yet, nationally, we are experiencing a shortage of prepared K-12 CS teachers (Blikstein, 2018). In response to this demand, many popular K-12 introductory CS curriculum providers, such as Code.org or Google CSFirst have produced curriculum that asks little from the classroom teacher, instead positioning the computer as the teacher of individual students. Many teachers, however, have expressed a desire to improve their own computational fluency, requesting access to more resources and professional development opportunities (Blikstein, 2018). Through *Getting Unstuck*, a summer 2019 online professional learning experience for K-12 educators, we looked to offer an opportunity for teachers to deepen their own computational fluency while building support networks and connecting with other teachers. But what kinds of supports are effective for K-12 educators? How does deepening their own learning support teachers in supporting students?

In this paper, I describe some of the quantitative methods used to explore data generated from participating in *Getting Unstuck*—specifically, the writing associated with Scratch projects participants made and shared in the Scratch online community. I first describe the motivation for and the design of *Getting Unstuck* before summarizing some of the ways reflective practice can support learning and some of the quantitative methods used to analyze reflective writing. I then describe some of the exploratory data analysis conducted on *Getting Unstuck* Scratch projects before concluding with some suggestions for future directions for data analysis.

# Background

*Getting Unstuck* was a free 21-day email-based learning experience for K–12 teachers interested in developing greater familiarity and fluency with Scratch, a block-based programming language and online community (Resnick et al., 2009). This professional development experience featured daily messages that invited teachers to: (1) create a Scratch project in response to a creative prompt that focused on a particular programming concept or Scratch feature, (2) share their work in the Scratch online community, and (3) reflect on their learning (Haduong & Brennan, 2018). Over 1900 participants signed up, and over 400 participated in at least one day (where participation is understood as submitting at least one Scratch project to a daily studio). Feedback at the end of the program, collected via an anonymous survey, was generally extremely positive, with many teachers reporting that they felt that they now knew more about Scratch. Some teachers also reported that participating in Getting Unstuck and supported their understanding of the emotional aspects of the creative process in computer programming, such as the feeling of satisfaction when a project worked, or the feeling of frustration when project bugs were left unresolved. As designers and facilitators of the 21-day experience, we (Dr. Karen Brennan and I) certainly *felt* that participants were making incredible, joyful, and creative projects. But what did participants actually *learn*?

For the purposes of this analysis, I decided to focus on the text which often accompanies a Scratch project shared in the online community. In addition to the code (e.g., block-based scripts) and the media assets (e.g., sprite costumes, sounds, and backgrounds), Scratch users often write “Instructions” and “Notes and Credits,” offering their audience more information about how to use the program or about their design process. For *Getting Unstuck*, the daily email often asked participants to write about what was challenging or surprising about their process, encouraging participants to share struggles they were experiencing. From a learning design perspective, we hoped that reflecting would deepen a learner’s understanding of what they learned, why they learned it, and how that learning took place (Dewey, 1933; Mezirow, 1994). While the reflection could be done after the project was finalized, Scratch projects can always be edited, offering opportunities for reflection-in-action, which can support teachers in considering not only what they made, but also how their own students might feel in these moments (Schön, 1987) .

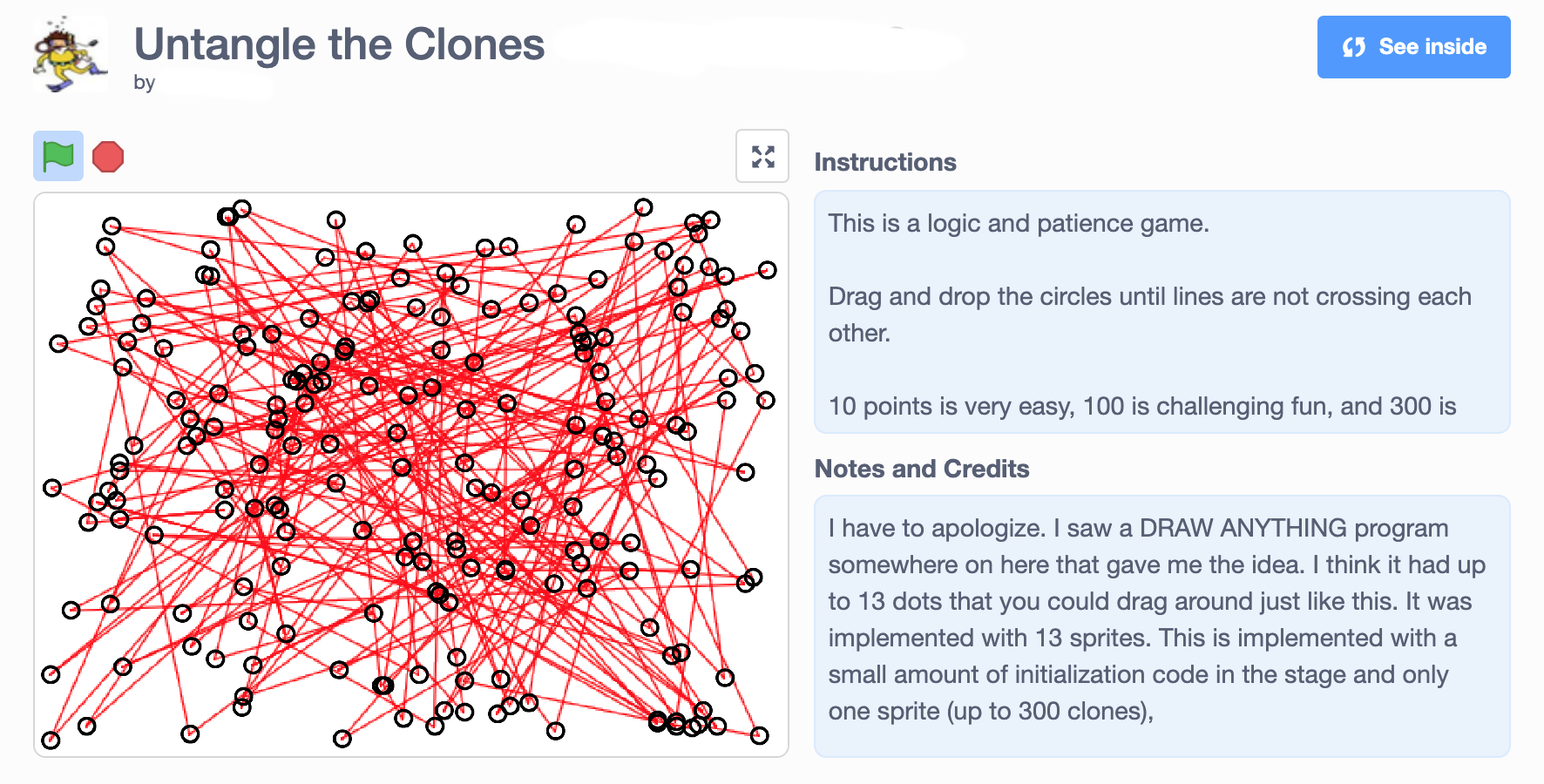


Figure 1. A prototypical Scratch project.

Through an analysis of the reflective texts associated with *Getting Unstuck* Scratch projects, I aspired to explore the following research questions:

* How much and how often did participants write in their Scratch projects?
* How often was that writing reflective?
* What kinds of things did participants write about?

All of these questions were connected to key questions about the relationship between reflection and learning, particularly in the *Getting Unstuck* context.

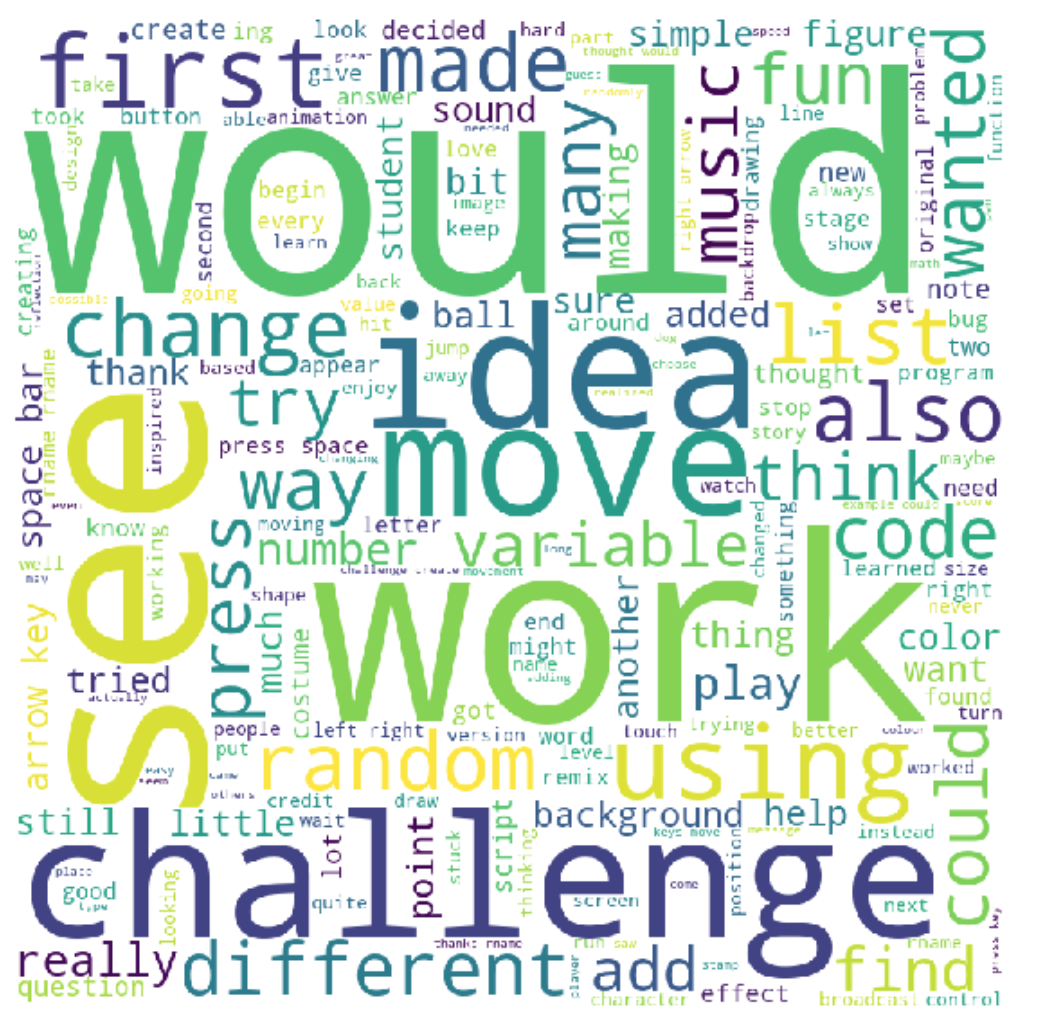
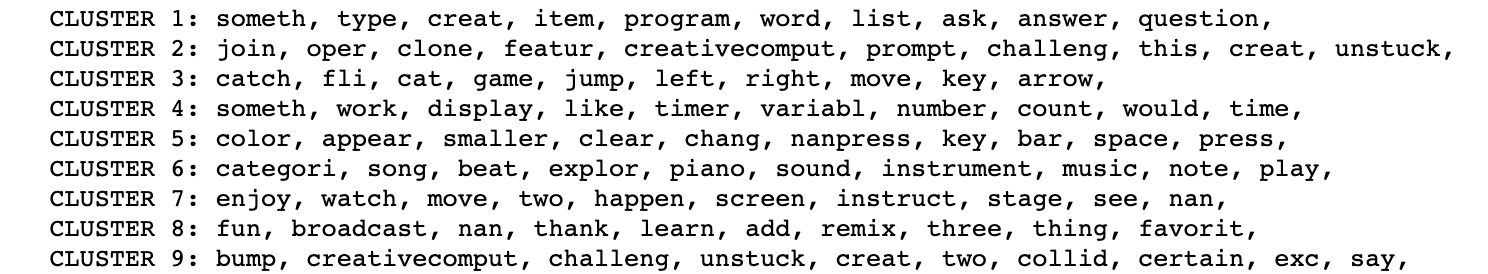
# Methods

Because there were so many projects, a quantitative approach was adopted for initial analysis. Scratch maintains a well-documented (if fairly limited) API,[[1]](#footnote-1) which was used to pull data about each project (though not the project files themselves). This data included the author ID, project ID, project title, “Instructions”, “Notes and Credits,” as well as other information not used for this analysis, such as the number of comments or number of times others had starred the project (indicating that it was a crowd favorite). Data was scraped by Dr. Karen Brennan from each Getting Unstuck studio (n=21) in August 2018, using Perl.

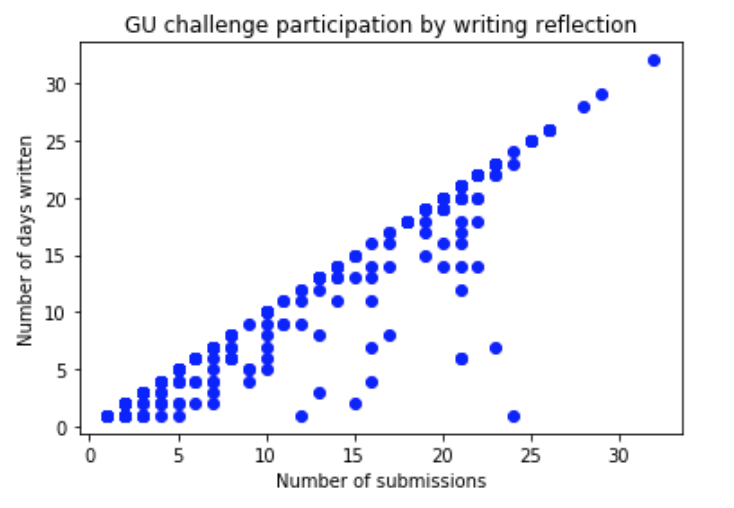
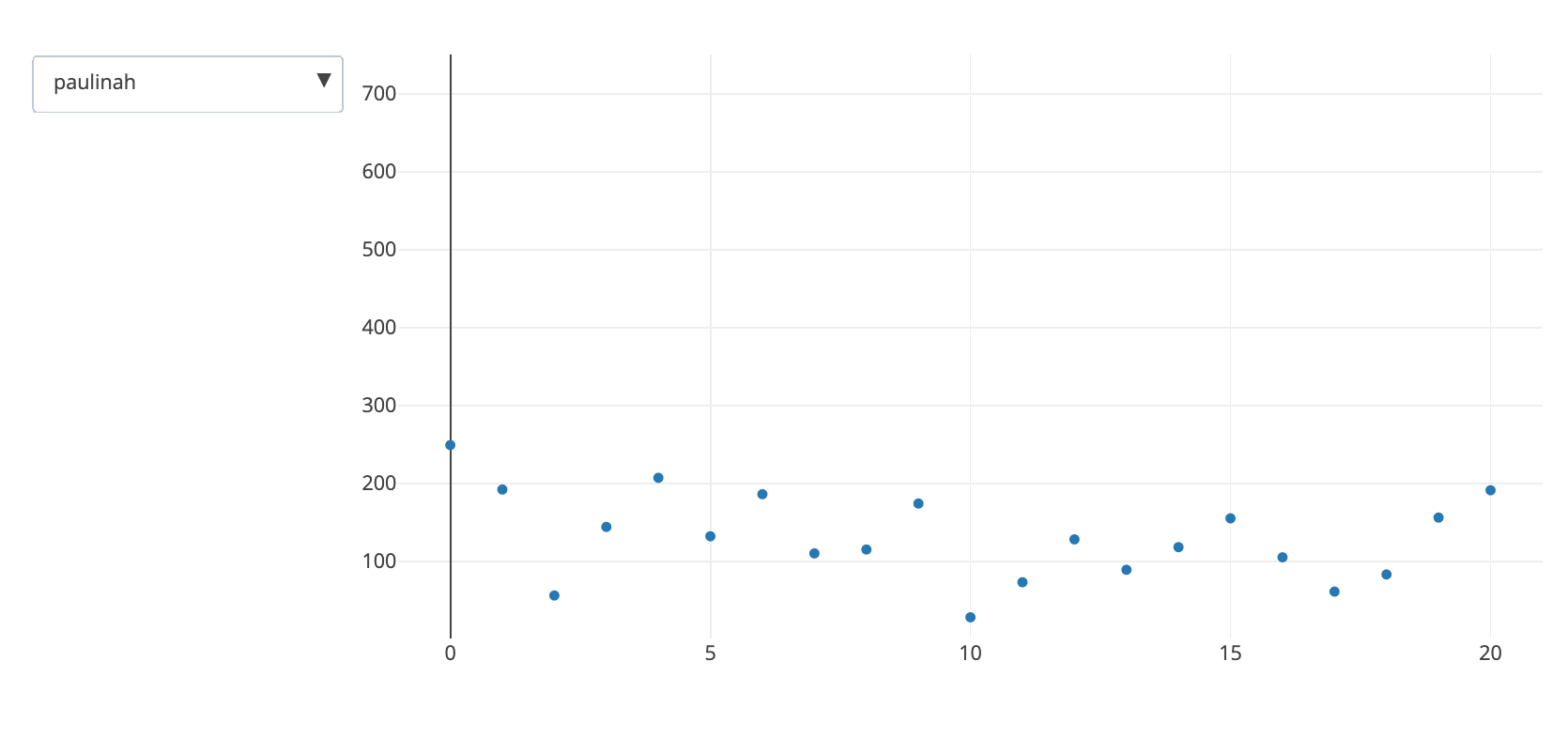
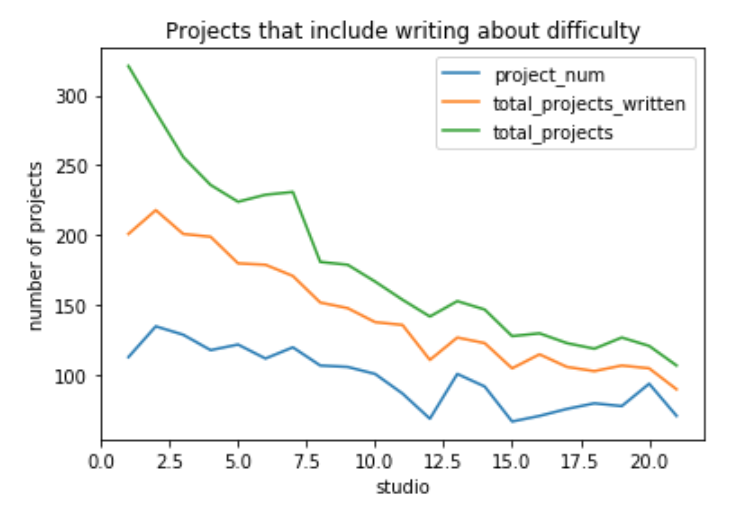
* Tools
  + Jupyter, Python3, Pandas
  + Everything stored in jupyter notebook at this link.
* Scraping/acquisition
  + Data was scraped from each Getting Unstuck studio in Aug 2018 (using Perl) – Dr. Karen Brennan scraped this
  + Author ids are not meaningfully connected to usernames; scraped by Emily Hong, an undergraduate member of the research team
  + I merged the CSV files
* Topic modeling (clusters) & word frequency
  + Why topic modeling
    - sklearn
    - Shirin’s paper
* Line graphs to look at trends over time
  + Reflection keyword frequency
    - Ullmann
  + Plotly, matplotlib,

# What did participants write about?

* K-means clustering
  + Meaningless
* Wordcloud



# How often and how much did participants write?



# Discussion

* Use keyword frequency as a way to segment the data & identify data subsets for qualitative coding.
* In Getting Unstuck 2019; could run a daily difficulty keyword analysis to attempt to find projects where learners were feeling challenged (not easy to do otherwise)

# Limitations

* This was from the synchronous experience, not asynchronous; studios are still running; could re-run the data analysis on newly collected data
* This analysis included all projects that participants submitted, which is why some of the x-ranges go past 21. Some participants submitted multiple projects for an individual day. One way to handle duplicates might be to drop duplicates from each studio, only analyzing data from a particular participant’s project.

# Conclusion & future directions

* Ngrams?
* Sentiment analysis
  + Two questions of pride/joy

# References (use Heading 1)

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In G. Sykes & L. Darling-Hammond (Eds.), Teaching as the learning profession: Handbook of policy and practice (pp. 3-32). San Francisco: Jossey Bass.

Brennan, K. (2013). *Best of both worlds: Issues of structure and agency in computational creation, in and out of school.* (Doctoral thesis). Massachusetts Institute of Technology, Cambridge, MA.

Gibson, A., Kitto, K., & Bruza, P. (2016). Towards the discovery of learner metacognition from reflective writing. *Journal of Learning Analytics*, 3(2), 22-36.

Haduong, P., & Brennan, K. (2019). Helping K–12 Teachers Get Unstuck with Scratch: The Design of an Online Professional Learning Experience. *Proceedings of the 2019 50th ACM Technical Symposium on Computer Science Education.* Minneapolis, MN.

Kovanović, V., Joksimović, S., Mirriahi, N., Blaine, E., Gašević, D., Siemens, G., & Dawson, S. (2018, March). Understand students' self-reflections through learning analytics*. In Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 389-398). ACM.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? Educational Researcher, 4-15.

Race, P. (2002). Evidencing reflection: putting the “w” into reflection. Retrieved May, 30, 2013.

Raelin, J. A. (2001). Public Reflection as the Basis of Learning. Management Learning, 32(1), 11-30. doi:10.1177/1350507601321002

Resnick, M., Maloney, J., Monroy-Hernandéz, A., Rusk, N., Eastmond, E., Brennan, K., . . . Kafai, Y. (2009). Scratch: programming for all. Commun. ACM, 52(11), 60-67. doi:10.1145/1592761.1592779

Schön, D. A. (1987). Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.

Ullmann, T. D. (2017, March). Reflective writing analytics: empirically determined keywords of written reflection. *In Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 163-167). ACM.

# Acknowledgments

I would like to thank Karen Brennan, Jess Eng, and Emily Hong for their ongoing support and feedback in the Getting Unstuck project. I would also like thank Bertrand Schneider and Will Yao for their guidance through the S435 class.

# General notes (Style “Heading 1” — Arial, 12 pt, bold, 12pt. space before, only first word capitalized except after a colon)

These are the formatting requirements for the papers that will be published in the CSCL 2019 proceedings. Your paper must conform to these guidelines so that each paper in the proceedings is uniform. Submitting your paper following these guidelines is necessary for insuring your paper’s inclusion in the proceedings. It is recommended that you use the MS Word styles in this document as a template for your document.

## Body text formatting (Style “Heading 2” — Arial, 12 pt, regular, 10 pt. space before, only first word capitalized except after a colon)

The initial paragraph following a header is in style “Body Text.” Use Times or Times New Roman 10 point. “Full justify” paragraph text, i.e., left and right justify. Do not indent or leave a blank line following headings. The margins for your entire paper should be 1 inch on the right and left and 1.25 inches on the top and bottom, using A4 paper.

Subsequent paragraphs are in style “Body Text First Indent.” Do not leave blank lines between paragraphs. Indent the first line of each paragraph .5 inches from the left margin.

*DO NOT use page numbers, running heads, or footnotes.* If you must use notes, please use endnotes, and place them immediately before the reference list. Refer to endnotes in text using a standard full-sized numeral inside parentheses without superscripting (1). For *emphasis* in your text, use *italics*.

Leave one line before each section header. This will be accomplished automatically if you used the MS Word styles in this document. Do not use more than three levels of headers.

### Page length (Style Heading 3 — Arial, 10 pt, regular, underlined, 10 pt. space before, only first word capitalized except after a colon)

Full papers may use a *maximum* of 8 pages; posters may use a maximum of 2 pages; short papers may use a maximum of 4 pages; symposia may use a maximumof 8 pages. *You may not exceed the maximum*. Refer to the conference call for papers for details.

### Blind review for papers and posters

Papers and posters will be reviewed blind, so it is important to prepare submissions so that reviewers do not know the names of the authors. Please prepare your paper in a way that preserves author anonymity. *Bulleted and enumerated lists are in 10 pt. Times New Roman font and having a 3 pt. space before and after each line. They are indented a quarter-inch on the left and have a quarter-inch hanging indent.*

* Do not include author names and institutions under the title in your submissions.
* Avoid using phrases like "our previous work" when referring to earlier publications by the authors.
* Remove information that may identify the authors in the acknowledgements (*e.g.,* grant IDs).
* Avoid providing web links to sites or other supplementary materials that may identify the authors

### Extended quotations

For extended quotes from source material, use style “Block Text,” Times or Times New Roman 10 pt., and indent the quote .5 inch from both the left and right margins. “Full justify” the text for the extended quote. The extended quote should be preceded and followed by one 10 pt. blank line.

Lorem ipsum dolor sit amet, consectetaur adipisicing elit, sed do eiusmod temporincididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrudexercitation ullamco laboris nisi ut aliquip ex ea commodo. (Euripedes, 1999, p. 213)

### Figures and tables

All figures and tables must be referred to in your text (see Table 1). Color figures may be included (see Figure 1). All figures and tables should be centered. Table captions are underlined and aligned left *above* the table. Figure captions are centered and placed *below* the figure.

Table 1: An example of a table for the ICLS proceedings

|  |  |  |  |
| --- | --- | --- | --- |
|  | Heading 1 | Heading 2 | Heading 3 |
| Row 1 | Cell 1 | Cell 2 | Cell 3 |
| Row 2 | Cell 4 | Cell 5 | Cell 6 |



Figure 1. The ISLS logo.

# References (use Heading 1)

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In G. Sykes & L. Darling-Hammond (Eds.), Teaching as the learning profession: Handbook of policy and practice (pp. 3-32). San Francisco: Jossey Bass.

Brennan, K. (2013). *Best of both worlds: Issues of structure and agency in computational creation, in and out of school.* (Doctoral thesis). Massachusetts Institute of Technology, Cambridge, MA.

Gibson, A., Kitto, K., & Bruza, P. (2016). Towards the discovery of learner metacognition from reflective writing. *Journal of Learning Analytics*, 3(2), 22-36.

Haduong, P., & Brennan, K. (2019). Helping K–12 Teachers Get Unstuck with Scratch: The Design of an Online Professional Learning Experience. *Proceedings of the 2019 50th ACM Technical Symposium on Computer Science Education.* Minneapolis, MN.

Kovanović, V., Joksimović, S., Mirriahi, N., Blaine, E., Gašević, D., Siemens, G., & Dawson, S. (2018, March). Understand students' self-reflections through learning analytics*. In Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 389-398). ACM.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? Educational Researcher, 4-15.

Race, P. (2002). Evidencing reflection: putting the “w” into reflection. Retrieved May, 30, 2013.

Raelin, J. A. (2001). Public Reflection as the Basis of Learning. Management Learning, 32(1), 11-30. doi:10.1177/1350507601321002

Resnick, M., Maloney, J., Monroy-Hernandéz, A., Rusk, N., Eastmond, E., Brennan, K., . . . Kafai, Y. (2009). Scratch: programming for all. Commun. ACM, 52(11), 60-67. doi:10.1145/1592761.1592779

Schön, D. A. (1987). Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.

Ullmann, T. D. (2017, March). Reflective writing analytics: empirically determined keywords of written reflection. *In Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 163-167). ACM.

# Acknowledgments (MOVE THIS BELOW)

I would like to thank Karen Brennan, Jess Eng, and Emily Hong for their ongoing support and feedback in the Getting Unstuck project. I would also like thank Bertrand Schneider and Will Yao for their guidance through the S435 class.

# References

Blikstein, P. (2018). *Pre-College Computer Science Education: A Survey of the Field*. Mountain View, CA: Google LLC.

Dewey, J. (1933). How we think: A restatement of the relation of reflective thinking to the educational process. *Lexington, MA: Heath, 35*(64), 690-698.

Haduong, P., & Brennan, K. (2018). Getting Unstuck: New Resources for Teaching Debugging Strategies in Scratch (Abstract Only). *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 1092-1092. doi:10.1145/3159450.3162248

Mezirow, J. (1994). Understanding transformation theory. *Adult education quarterly, 44*(4), 222-232.

Resnick, M., Maloney, J., Monroy-Hernandéz, A., Rusk, N., Eastmond, E., Brennan, K., . . . Kafai, Y. (2009). Scratch: programming for all. *Commun. ACM, 52*(11), 60-67. doi:10.1145/1592761.1592779

Schön, D. A. (1987). Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.

1. <https://github.com/LLK/scratch-rest-api/wiki> [↑](#footnote-ref-1)