

Impact of children encouraging other children while programming

Problem & Motivation

- ◆ Scratch is a visual, block-based programming language for children (Figure 1).
- ◆ Scratch also functions as a social network [3], facilitating interactions between users through comments on each others projects.
- ◆ Many comments contain encouragements, suggestions and tips [4].
- ◆ Users that receive negative and disparaging comments due to cultural or ethnic racism seem to leave the platform (hypothesis) [5] (figure 3).

- ◆ This raises the research question: To what extent is there a relation between the **sentiment of comments** on a Scratch project and the **creator's inclination to continue producing projects**?

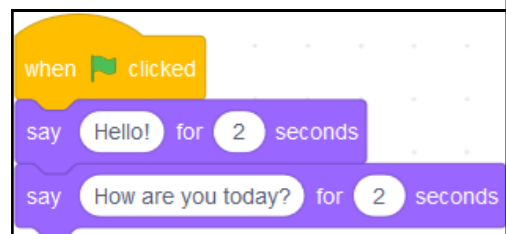


Figure 1: An example of block-based programming.

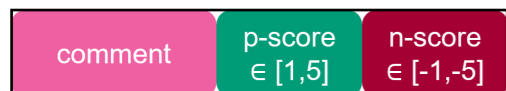


Figure 2: Visualization of comment data.

Methodology

- ◆ We conducted seven experiments to find a correlation, each a combination of **independent** and **dependent** variables:
- ⇒ **Number of projects created by a user**, possibly normalized by the number of projects he created before

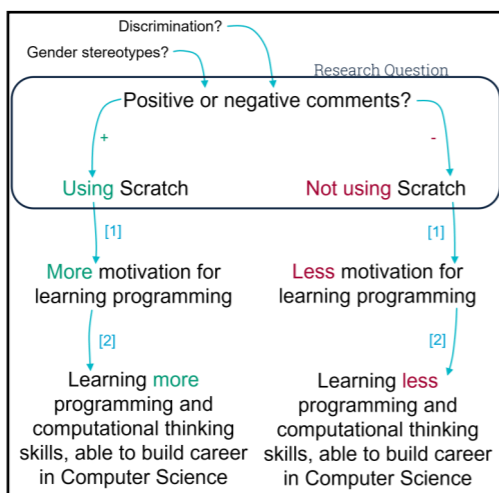


Figure 3: Diagram demonstrating the research context.

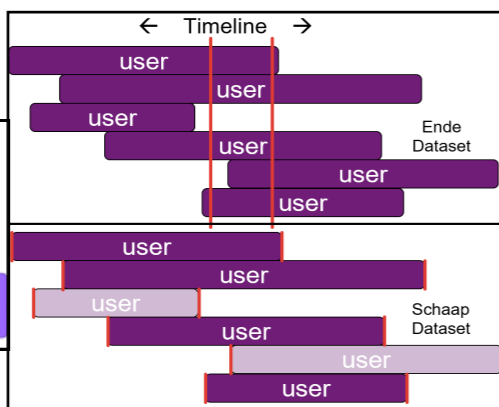


Figure 4: Differences between datasets. Ende contains users' projects and their comments during only a specific time period, namely 2019, for a large group of users. Schaap contains users' projects and their comments during the whole time users are active, but for a smaller group of users.

⇒ **Total time that a user is active** on the Scratch platform

⇒ **Relative or absolute positiveness or negativeness** of all comments received

⇒ **Percentage of comments received categorized** as positive, negative or neutral

- ◆ For each comment, we retrieved a value for **positiveness** from 1 to 5 and for **negativeness** from -1 to -5 (Figure 2).

- ◆ We used two datasets: The Ende dataset contains projects and their comments from 2019 for specific users. The Schaap dataset contains all projects and their comments for specific users (Figure 4).

Experiment and Dataset	Statistical Test	Independent Variable	Dependent Variable	Results
Experiment 1 Ende	Pearson	Relative or absolute number of comments that have a positive or negative score (above a threshold) in timeframe 1	Number of projects created per month in timeframe 2 over the number of projects created per month in timeframe 1	No correlation
Experiment 2 Schaap	Pearson	Relative or absolute number of comments that have a positive or negative score (above a threshold)	Active time, from the first project creation date to the last project modification date	Low correlation
Experiment 3 Schaap	Pearson	Ratio of the number of positive, negative or neutral comments over the total number of comments in active time	Active time, from the first project creation date to the last project modification date	Low correlation
Experiment 4 Ende	Pearson	Ratio of the number of positive, negative or neutral comments over the total number of comments in timeframes 1 and 2 together	Number of projects created in timeframes 1 and 2 together	No correlation
Experiment 5 Schaap	Pearson, Spearman, Kendall Tau	Average number of comments per project	Number of projects created	Low correlation
Experiment 6 Schaap	Pearson, Spearman, Kendall Tau	Ratio of the number of positive, negative or neutral comments over the total number of comments in the first three projects	Number of projects coming afterwards	No correlation
Experiment 7 Schaap	Pearson, Spearman, Kendall Tau	Negative or positive score for each comment	Number of projects created afterward	No correlation

Table 1: Data Analysis performed on different combinations of variables and their results.

Findings & Discussion

- ◆ For some combinations of variables, we found a low correlation. For other combinations of variables, we found no correlation (Table 1). A high correlation was not found.
- ◆ A low correlation was found between the number of comments received and the number of projects created (Figure 5).
- ◆ The correlation was not always clearly visible from a graph (Figure 6).
- ◆ **Active time** should be calculated differently for future research, now there is too much weight on users who create many projects (Figure 8).

- ◆ In experiment 1, it seems that many active users create on average less than 1 projects per month (Table 2), which makes that result unreliable.

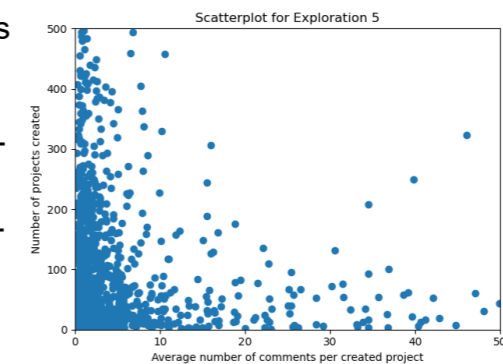


Figure 5: Low correlation between number of comments received and number of projects created.

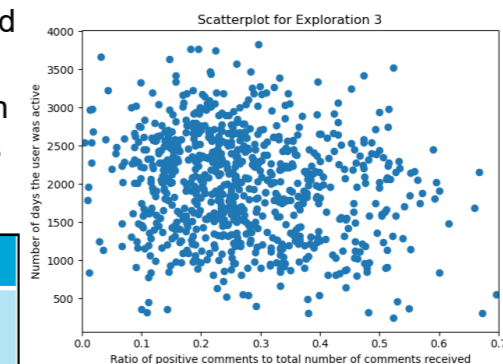


Figure 6: Low, but hardly visible, correlation between ratio of positive comments and active time.

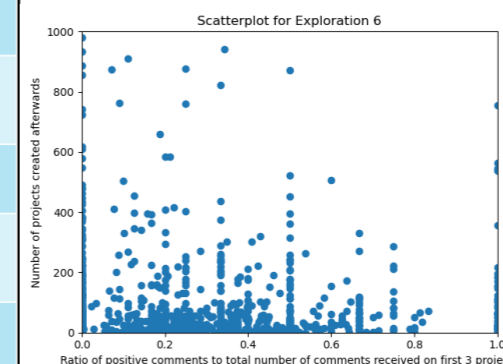


Figure 7: Spikes in one of the scatterplots revealing no correlation.

- ◆ In experiment 6, "spikes" are visible throughout the graph (Figure 7), which makes that result unreliable.
- ◆ Negative comments might already have been reported before we scraped the data.

- ◆ When scraping friends of friends, the way the Ende dataset and the Schaap dataset were originally created, the dataset might not include hated projects.
- ◆ Children might make more spelling mistakes than adults, or they might use more unofficial language.

Conclusion

- ◆ The effect of **comment sentiment** alone on the **creator's inclination to continue producing projects** is very limited, no correlation was found in 4 out of 7 experiments and a low correlation was found in 3 out of 7 experiments.

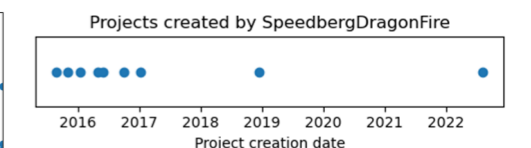


Figure 8: Example of a user who has been inactive a long time and then got active again.

Min.	Q1	Q2	Q3	Max.
0.00	0.00	0.50	1.16	48.00

Table 2: Spread of number of projects created in May 2019 over number of projects created per month in January-April 2019.

References

[1]: Fu-Hsiang Wen, Tienhua Wu, and Wei-Chih Hsu. Toward improving student motivation and performance in introductory programming learning by scratch: The role of achievement emotions. *Science Progress*, 106(4):00368504231205985, 2023.

[2]: Ad Zeevaarders and Efthimia Aivaloglou. Exploring the programming concepts

practiced by scratch users: an analysis of project repositories. In *2021 IEEE Global Engineering Education Conference (EDUCON)*, pages 1287–1295. IEEE, 2021.

[3]: Isabella Graßl and Gordon Fraser. Scratch as social network: topic modeling and sentiment analysis in scratch projects. In *Proceedings of the 2022 ACM/IEEE 44th International Conference on Software Engineering: Software Engineering in Society*, pages 143–148, 2022.

[4]: Deborah A Fields, Katarina Pantic, and Yasmin B Kafai. "i have a tutorial for this" the

language of online peer support in the scratch programming community. In *Proceedings of the 14th International Conference on Interaction Design and Children*, pages 229–238, 2015.

[5]: Gabriela T Richard and Yasmin B Kafai. Blind spots in youth diy programming: Examining diversity in creators, content, and comments within the scratch online community. In *Proceedings of the 2016 CHI conference on Human Factors in Computing Systems*, pages 1473–1485, 2016