

Individual versus computer-supported collaborative self-explanations: how do their writing analytics differ?

Laia Albó, Marc Beardsley, Ishari Amarasinghe, Davinia Hernández-Leo
Department of Information and Communications Technologies
Universitat Pompeu Fabra
Barcelona, Spain
{laia.albo, marc.beardsley, ishari.amarasinghe, davinia.hernandez-leo}@upf.edu

Abstract— Researchers have demonstrated the effectiveness of self-explanations (SE) as an instructional practice and study strategy. However, there is a lack of work studying the characteristics of SE responses prompted by collaborative activities. In this paper, we use writing analytics to investigate differences between SE text responses resulting from individual versus collaborative learning activities. A Coh-Metrix analysis suggests that students in the collaborative SE activity demonstrated a higher level of comprehension. Future research should explore how writing analytics can be incorporated into CSCL systems to support student performance of SE activities.

Keywords: *Self-explanations, writing analytics, Coh-Metrix, computer-supported collaborative learning*

I. INTRODUCTION

Having students generate explanations is a recommended instructional practice [1] and is also an effective study strategy for students to adopt [2]. Explaining to oneself in an attempt to make sense of new information is considered beneficial as Self-Explanations (SE) promote greater processing of causal information, conceptual relationships [3][4] and more coherent mental representations of text [5][6]. Yet, despite the benefits of SE, not all learners are able to effectively engage in SE activities and, most of the time, the SE process needs to be supported or prompted [7]. Research has demonstrated that assisting SE prompts (e.g. requesting learners to process the to-be-learned contents in a specific way) are more effective than simple open prompts (e.g. open questions inducing SE) [7].

Although SE content varies widely [1][3], SE tasks and assisting SE prompts are commonly characterized for being of an individual nature. Yet, collaborative learning approaches have been shown to benefit learner comprehension [8]. Further, peer assessment approaches in which students review and assess, either summatively or formatively, the work of their peers, has been shown to positively benefit student learning [9]. Thus, including collaborative tasks during a SE activity, should foster learner comprehension and assisting SE prompts can facilitate student performance of such tasks.

In the domain of collaborative learning, carefully designed collaboration scripts are being employed to structure the flow of collaboration [10]. In the context of Computer-Supported Collaborative Learning (CSCL), Collaborative Learning Flow Patterns (CLFPs) such as jigsaw and pyramid patterns, formulate the essence of the script structures. CLFPs have been proven effective in multiple educational situations [11]. Collaboration scripts not only help to maintain a pedagogical method structure throughout a learning activity but also can be used to trigger beneficial social and cognitive interactions among activity participants. In comparison to spontaneous collaboration,

these facilitated interactions increase the learning opportunities for participants [10]. Moreover, engaging in argumentative knowledge construction processes during scripted collaborative learning activities has been shown to enhance the domain-specific knowledge acquisition of activity participants [12].

In the case of SE, the benefits of a collaborative SE approach should be evident in the written responses of participants. Research related to assessing comprehension in written text suggests that certain properties of text (e.g. cohesion) correlate with enhanced levels of comprehension [5]. Further, the way individuals make connections across a text can provide important information about successful discourse processing strategies [5][14]. These different connection types can be examined through automated cohesion analyses [13]. Recent work has explored how cohesion relates to SE prompts [3] and CSCL participation [14]. However, much less is known about the characteristics of SE texts prompted by Computer-Supported Collaborative Self-Explanation (CSCSE) activities.

Therefore, the objective of this research is to determine whether, and how, individual and CSCSE differ in their textual properties. This analysis also aims to shed light on how the incorporation of writing analytics in CSCL systems could be useful for monitoring student performance and facilitating student self-regulation of learning, e.g. by providing real-time feedback in the context of SE activities.

II. METHODOLOGY

This research was conducted using a quasi-experimental design with post-test only [15] as we studied whether the instructional condition (individual versus collaborative activity) influenced the cohesion of students' SE responses. Data analysis results and data excerpts are available in Zenodo [16].

A. Participants and procedure

Participants of the study were first-year engineering degree students from a public university enrolled in an Introduction to Engineering Studies academic course. The study was conducted in a 2-hour face-to-face transversal skills lesson on the Science of Learning in which students learned about cognitive theories on learning and efficient learning strategies such as retrieval practice, distributed practice and spaced learning [17]. The lesson was designed to include two reflection activities: one on retrieval practice and the other about distributed practice. The first reflection activity occurred after a third of the lesson had been completed. The second reflection activity occurred after two-thirds of the lesson had been completed. The reflection activities were designed as either individual or collaborative self-explanation tasks (Fig. 1). The individual self-

explanation task asked students to (1) explain a key concept in writing (e.g. retrieval practice); (2) define keywords (e.g. neuroplasticity, encoding); and (3) rewrite the explanation of the key concept (e.g. retrieval practice) while including concepts from the keyword definitions. The collaborative self-explanation task asks students to (1) explain the key concept in writing (e.g. retrieval practice); (2) consider the provided keywords (e.g. neuroplasticity, encoding) when reading and rating the answers provided by peers; and (3) rewrite the explanation of the key concept (e.g. retrieval practice) to include concepts from the keyword definitions.

B. Instrumentation, data collection and analysis

An online Google Form was used for each step of the individual SE path and the third step of the collaborative one (Fig. 1). PyramidApp, a web-based tool [18] that facilitates the deployment of Pyramid CLFP activities, was used for enacting the collaborative activity in steps (1) and (2) of the collaborative SE path. The first phase of the script required students to submit an individual answer (i.e. individual concept explanation). Upon submission, the second phase began by randomly allocating students to small groups. During this phase, students could see the keywords which were configured by the teacher during the activity design stage for SE and the individual answers submitted by peers in their groups. Students were required to give an "out of 5" star rating to each group member's answer taking into consideration their alignment with the keywords listed. At the end of the pyramid learning activity, students could see the highly rated answers from the different Pyramids formulated and were subsequently prompted to rewrite their explanation of the key concept.

To analyze the written responses from step (3) in both groups, writing analytics were used. Writing analytics involve the measurement and analysis of written texts for the purpose of understanding writing processes and products, in their educational contexts [19]. There are a number of tools that incorporate Natural Language Processing (NLP) techniques for the provision of automated writing analytics [19]. Among the existing options, Coh-Metrix [13] is a computational tool that produces indices of the linguistic and discourse representations of a text. These values can be used to investigate the cohesion of an explicit text and the coherence of the mental representation of the text [13]. Both cohesion and coherence are used to assess the quality of a written text. According to the Coh-Metrix authors, cohesion consists of the characteristics in an explicit text that play some role in helping the reader mentally connect ideas conveyed in the text [13]; while coherence reflects the interaction between linguistic representations and knowledge representations. Coh-Metrix provides indices of cohesion and coherence characteristics [13]. Thus, to understand whether and how individual and CSCSE differed in terms of cohesion, we analyzed the linguistic features of the students' SE responses using Coh-Metrix. Among the measures obtained from the automated evaluation of texts, we analyzed the following indices: descriptive indices such as sentence length, and word and sentence count; easibility indices such as: (a) **Narrativity**: whether a text tells a story with characters, events, places and things that are familiar to the reader; (b) **Syntactic Simplicity**: the use of simple sentence structures that are easy to understand; (c) **Word Concreteness**: the use of words that can be easily imagined; (d) **Referential**

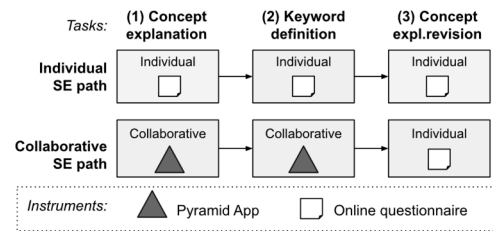


Figure 1. Individual vs. collaborative paths: procedure and instruments.

Cohesion: the presence of overlapping ideas and concepts in a text; and (e) Deep Cohesion: the presence of explicit causal relationships. Moreover, additional Coh-Metrix indices were reviewed to explore whether significant differences existed between conditions.

Statistical analysis was performed using SPSS software (version 23). Significance levels were set at the 5% level with the student t-test or Mann-Whitney U test being used depending on the distribution (normality) of each measure.

III. RESULTS AND DISCUSSION

In total, 37 students participated in the first reflection activity (19 in the individual path and 18 in the collaborative option), whereas 32 students performed the second reflection activity (17 students followed the individual path and 15 the collaborative one). Overall, no significant differences appeared in the descriptive and easibility indices (see Table 1), however, significant differences were found in measures of word overlap and readability suggesting that the CSCSE group demonstrated a higher level of comprehension than the individual SE group.

CSCSE responses had a significantly higher level of word overlap between sentences in terms of nouns and pronouns (Argument overlap: Individual, mean 0.13, Std. Error 0.05; Collaborative: mean 0.33, Std. Error 0.08; $p=0.046$); while the individual SE responses had a significantly higher level of verb overlap (Ind.: mean 0.54, Std. Error 0.04; Collab.: mean 0.39, Std. Error 0.02; $p=0.036$). Verb cohesion has been found to be greater in lower grade level texts when compared to higher grade texts [20]. Yet, CSCSE responses received higher readability scores on the Flesch Reading Ease measure (individual SE condition: mean 31.40, Std. Error 3.08; CSCSE condition: mean 42.97, Std. Error 2.98; $p=0.009$). These results suggest that CSCSE responses were more readable and of a higher-grade level.

A. Descriptive indices

Despite no significant differences being found, the Coh-Metrix descriptive indices revealed that CSCSE responses tended to use more words (36.33 vs. 39.48); contain more sentences (1.53 vs. 1.60); and longer sentences (26.50 vs. 27.22) than individual SE responses.

B. Easibility indices

No significant differences were found in the easibility indices which are presented in Table 1.

(a) **Narrativity**: CSCSE responses tended to use more familiar words which can be interpreted to mean they were more related with everyday conversation.

(b) **Syntactic Simplicity**: CSCSE responses tended to use simpler and more familiar syntactic structures that are less challenging to process.

TABLE I. COH-METRIX TEXT EASIBILITY INDICES

<i>Coh-Metrix indices</i>	<i>Individual SE</i>	<i>Collaborative SE</i>
(a) Narrativity	25.55%	39.10%
(b) Syntactic simplicity	17.11%	19.34%
(c) Word concreteness	72.28%	68.25%
(d) Referential cohesion	60.90%	67.50%
(e) Deep cohesion	35.17%	44.77%

(c) **Word Concreteness:** CSCSE responses tended to contain fewer concrete words and more abstract words. Abstract words reflect concepts that are difficult to represent visually. Texts that contain more abstract words are more challenging to understand. A possible interpretation is that the CSCSE condition contained texts with a higher degree of difficulty as evidenced by the lower degree of concreteness in the words used.

(d) **Referential Cohesion:** CSCSE responses tended to have more words and ideas that overlapped across sentences and throughout the entire text; thereby forming explicit threads that connect the text for the reader (referential cohesion).

(e) **Deep Cohesion:** CSCSE responses tended to contain a higher degree of causal and intentional connectives (deep cohesion). When a text contains many relationships but does not contain such connectives, the reader must infer the relationships between the ideas in the text. If the text is high in deep cohesion, then those relationships are more explicit [13] and easier to understand [5].

The easibility results support the word overlap and readability results in suggesting that CSCSE responses were better (i.e. more coherent and cohesive) and more likely to reflect a higher level of comprehension.

IV. CONCLUSION AND FUTURE WORK

The results of the Coh-Metrix analysis suggest that students in the collaborative SE condition demonstrated a higher level of comprehension than those in the individual SE condition. However, these findings may be affected by the short length of the texts analyzed which can limit the effectiveness of the measures being used and the limited sample size of the study. Overall, the results of our analysis indicate that it may be worthwhile to further investigate the differences between both conditions. For instance, to triangulate the quantitative results obtained from Coh-Metrix with a qualitative analysis provided by content experts (i.e. a manual grading of the texts). Future research in this direction can explore the use of writing analytics in CSCL systems as a way of supporting the real-time monitoring and regulation of learning.

ACKNOWLEDGMENT

This work has been partially funded by the EU Regional Development Fund, Erasmus+, and the National Research Agency of the Spanish Ministry of Science, Innovation and Universities under project grants TIN2017-85179-C3-3-R, 2017-1-ES01-KA201-038220, RED2018-102725-T, and the UPF PlaClik program. D. Hernández-Leo acknowledges the support by ICREA under the ICREA Academia programme.

REFERENCES

- [1] R. E. Mayer and P. A. Alexander (Eds.), *Handbook of research on learning and instruction*, 2nd ed. Routledge: London, 2017.
- [2] C. E. Weinstein, T. W. Acee and J. Jung, "Self-regulation and learning strategies," *New directions for teaching and learning*, vol. 2011(126), June 2011, pp. 45-53.
- [3] K. Bisra, Q. Liu, J. C. Nesbit, F. Salimi and P. H. Winne, "Inducing Self-Explanation: a Meta-Analysis," *Educ. Psychol. Rev.*, vol. 30(3), March 2018, pp. 703-725.
- [4] C. H. Legare and T. Lombrozo, "Selective effects of explanation on learning during early childhood," *Journal of Experimental Child Psychology*, vol. 126, October 2014, pp. 198-212.
- [5] L. K. Allen, D. S. McNamara and M. McCrudden, "Change your Mind: investigating the effects of self-explanation in the resolution of misconceptions," in D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings and P. Maglio (Eds.), *Proc. Cog Sci 2015*, pp. 78-83. Pasadena, CA: Cognitive Science Society.
- [6] D. S. McNamara and J. P. Magliano, "Self-explanation and metacognition: The dynamics of reading," in *Handbook of metacognition in education*. Routledge, 2009, pp.72-94.
- [7] K. Berthold, T. H. S. Eysink and A. Renkl, "Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations," *Instr. Sci.*, vol.37(4), July 2009, pp.345-363.
- [8] S. Rojas-Drummond, N. Mazón, K. Littleton and M. Vélez, "Developing reading comprehension through collaborative learning," *Journal of Research in Reading*, 37(2), 2014, pp.138-158.
- [9] K. Topping, "Peer assessment between students in colleges and universities," *Review of educ. research*, 68(3), 1998, pp. 249-276.
- [10] L. Kobbe, A. Weinberger, P. Dillenbourg, A. Harer, R. Hämäläinen, P. Häkkinen, and F. Fischer, "Specifying computer-supported collaboration scripts," *International Journal of Computer-Supported Collaborative Learning*, vol. 2, 2007, pp. 211-224.
- [11] D. Hernández-Leo, J. I. Asensio-Pérez, Y. Dimitriadis, and E. D. Villasclaras-Fernández, "Generating CSCL scripts: From a conceptual model of pattern languages to the design of real scripts," *Goodyear P, Retalis S, eds. Technology-enhanced learning: design patterns and pattern languages*, 2010, pp. 49-64.
- [12] P. Hermann and P. Dillenbourg, "Elaborating new arguments through a CSCL script," P. Dillenbourg (Ed.), *Learning to argue*, vol. 1, 2003, pp. 205-226, Dordrecht, The Netherlands: Kluwer.
- [13] D. S. McNamara, A. C. Graesser, P. M. McCarthy and Z. Cai, *Automated evaluation of text and discourse with Coh-metrix*. NY: Cambridge University Press, 2014.
- [14] M. Dascalu, D. S. McNamara, S. Trausan-Matu and L. K. Allen, "Cohesion network analysis of CSCL participation," *Behavior Research Methods*, vol. 50 (2), April 2018, pp. 604-619.
- [15] J. W. Creswell, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. New Jersey: Pearson/Merrill Prentice Hall, 2002.
- [16] L. Albó, M. Beardsley, I. Amarasinghe, and D. Hernández-Leo. Individual versus computer-supported collaborative self-explanations: how do their writing analytics differ? Coh-Metrix results and students' excerpts, 2020, [Data set], Zenodo.
- [17] Y. Weinstein, M. Sumeracki and O. Caviglioli. *Understanding how we learn: A visual guide*. Routledge, 2018.
- [18] K. Manathunga and D. Hernández-Leo, "Authoring and enactment of mobile pyramid-based collaborative learning activities," *British Journal of Educational Technology*, 49(2), 2018, pp. 262-275.
- [19] S. Buckingham Shum, S. Knight, D. McNamara, L. K. Allen, D. Bektik and S. Crossley, "Critical perspectives on writing analytics" in: *Proc. of the Sixth International Conference on Learning Analytics & Knowledge*, ACM, New York, 2016, pp. 481-483.
- [20] D. S. McNamara, A. C. Graesser and M. M. Louwerse. "Sources of text difficulty: Across genres and grades" in *Measuring up: Advances in how we assess reading ability*, 2012, pp. 89-116.