

Assessing the impact of an online inquiry teaching intervention on sixth graders' search performance

Roberto González-Ibáñez^a, Daniel Gacitúa^a, Gonzalo Martínez-Ramírez^a,
Jacqueline Köhler^a, Eero Sormunen^b, Carita Kiili^c, Mirjamaija Mikkilä-Erdmann^d,
Norbert Erdmann^d, Marja Vauras^d and Paavo H.T. Leppänen^e

^a*Departamento de Ingeniería Informática, Universidad de Santiago de Chile*

^b*Faculty of Information Technology and Communication Sciences, Tampere University*

^c*Faculty of Education and Culture, Tampere University*

^d*Department of Teacher Education, University of Turku*

^e*Department of Psychology, University of Jyväskylä*

Abstract

The study of online inquiry competences (OIC) is an important topic on the information literacy (IL) field. Most of the work has been focused on higher education and high-school students, while less has been done regarding primary education. In this work we investigate the effects of an OIC teaching intervention on search performance of a large group of sixth graders from Finland. Our preliminary results show significant improvements in search performance on the intervened group of students compared to a control group when working on a science research task. This work shed light about the potential benefits of a particular approach to develop OIC on elementary school students.

Keywords

online inquiry competences, information literacy, assessment, tests, elementary schools

1. Introduction and related work

The rise of the World Wide Web undoubtedly changed the way people look for information and provided an ever-growing source of information, hence boosting research in the field of Information Literacy (IL) [1]. In this context, special attention has been given to inquiry skills, which can be defined as the skills to gather, interpret, and synthesize different kinds of information and data in order to develop and share answers to questions [2]. From this definition, online inquiry competences (OIC) involve the above skills in addition to the knowledge and the abilities to inquire the Web [3]. OIC are highly valued in the 21st century [4].

[5] established that library and information skills (which are within the scope of OIC) involve a series of cognitive activities: task definition, information seeking strategies, location, access, use, synthesis, and evaluation of information. [6] emphasized the need of de-

veloping OIC in all levels of education. However, to the best of our knowledge, most of the effort on developing and assessing OIC has been focused on tertiary education students [7] and, in a smaller proportion, on high school students [8, 9] and primary education.

Given the above context, the iFuCo project [10] set an ambitious goal to design and carry out an intervention to develop OIC in a group of Finnish and Chilean elementary school students. While the scope of this project targeted students in both countries, interventions were not identical due to cultural differences. Moreover, while the project focused on developing four component skills (i.e., search and locate information, identification of main ideas, critical evaluation, and synthesis) [11], in this article we only study the effects of our intervention on search skills. In particular, we address the following research question (RQ): To what extent, if any, can a teaching intervention targeted to develop OIC on Finnish sixth graders improve their search performance in the context of research tasks in multiple domains?

In the next section we introduce the methodological approach as defined in the iFuCo project. Following, we present preliminary results. Finally, we conclude with a brief discussion of our findings.

Proceedings of the CIKM 2020 Workshops, October 19-20, 2020, Galway, Ireland

EMAIL: roberto.gonzalez.i@usach.cl (R. González-Ibáñez); daniel.gacitua@usach.cl (D. Gacitúa); gonzalo.martinez@usach.cl (G. Martínez-Ramírez); jacqueline.kohler@usach.cl (J. Köhler); eero.sormunen@utu.fi (E. Sormunen); carita.kiili@utu.fi (C. Kiili); mirmik@utu.fi (M. Mikkilä-Erdmann); nwmerd@utu.fi (N. Erdmann); vauras@utu.fi (M. Vauras); paavo.ht.leppanen@jyu.fi (P.H.T. Leppänen)

© 2020 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)



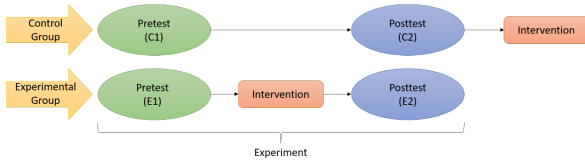


Figure 1: Pretest-Posttest study design.

2. Method

2.1. Study Design

Our approach to address the above RQ involved a quasi-experiment following a pretest-posttest design as shown in Figure 1. To assess students’ OIC in both pretest and posttest, we devised a performance-based test focused on four component skills, namely, (1) searching and selecting relevant sources, (2) identifying main ideas from sources, (3) evaluating the credibility of sources, and (4) synthesizing information across multiple sources [11], which was carried out using NEURONE [12].

As for the intervention, this consisted of a training program involving three modules (i.e., Module 1: Explicit teaching of OIC, which included aspects such as query formulation and analysis of search results; Module 2: Applied skills in a science research task; Module 3: Applied skills in a social science research task) in a span of three to four weeks. The intervention involved conceptual classes, tutorials, and practice without NEURONE. More details of the study design and intervention can be found in [13].

2.2. Sample

We recruited 364 sixth graders from 10 Finnish schools (15 classes) distributed over three cities (i.e., Tampere, Turku, and Jyväskylä). From this group, 344 students were authorized by their parents and 2 of them were absent during the tests. Therefore, our initial sample consisted of 342 students whose mean age was 12.3 (SD=.41) years old. Regarding sex, 165 (48.25%) were girls and 177 (51.75%) were boys.

Classes in which the study was carried out were randomly assigned to control or experimental groups. Note that randomization was school-class-based. As a result, eight classes were assigned to the experimental group (192 students - 46.85% girls, 53.15% boys) and the remaining seven classes to the control group (150 students - 50% girls, 50% boys).

Table 1
Task and domain rotations.

Pretest		Posttest	
Session 1	Session 2	Session 1	Session 2
Soc. Sci. Article	Science Email	Science Article	Soc. Sci. Email
Science Article	Soc. Sci. Email	Soc. Sci. Article	Science Email

2.3. Task

We considered two knowledge domains, “science” and “social science”, and two tasks: “writing an article” and “writing an email response”. The combination of these domains and tasks formed four activities for the students, each one with a multifaceted topic, described as follows: (1) Science article: “Finnish forests”, (2) Science email: “Origins of rain”, (3) Social science article: “Computer games” and (4) Social science email: “Reading on digital screens”. One science and social science task were addressed as part of the pretest and posttest according to the rotations illustrated in Table 1.

For each activity, students were allowed to search in NEURONE within a collection of 20 documents. Three of them, marked as relevant sources, were designed by researchers. The remaining 17 pages were authentic websites.

2.4. Session workflow

The session workflow in both pretest and posttest involved four stages linked to the above mentioned component skills. First, the search and selection phase was conducted in a maximum of 8 minutes. This phase was completed either when students found all three relevant sources or when time was up. Second, students were given 12 minutes to identify main ideas in the relevant sources. Third, the critical evaluation of the sources was done within 7 minutes. Finally, the synthesis phase was completed in a maximum of 15 minutes. Overall, sessions lasted approximately 50 minutes.

2.5. Study setup

Sessions were conducted in schools, using schools’ computers and network connectivity to access the NEURONE server. Sessions were supervised by members of the research team. Once the study data collection was completed, we gathered NEURONE database dumps to perform the analyses.

3. Results

As noted above, in this article we focused on data collected during the search and selection stage. In particular, we performed both within- and between- subjects comparisons based on our study design. We grouped all data from the control and experimental groups regardless the tasks and domains.

After pre-processing the data we were able to consolidate a database of 273 students. Records from the remaining 69 students were discarded due to missing data, incomplete sessions, or corrupted data due to technical issues during sessions (e.g., connection problems, operating system or browser incompatibilities with NEURONE, which mainly affected the search phase). From this group, 448 sessions belong to the control group and 644 to the experimental group.

$$Recall(s) = \frac{\#RelevantDocsBookmarked(s)}{3} \quad (1)$$

$$SearchScore(s) = 5 * \frac{\#RelevantDocsBookmarked(s)}{\#ActiveBookmarks(s)} \quad (2)$$

Data include a wide range of variables linked to search behaviors (e.g., dwell time in pages, mouse movements, keystrokes). From this list of variables, in this article we only analyzed the effects of the intervention on search performance, which was expressed in terms of recall and what we refer to search score [11]. On the one hand, given the goal of the search phase in which the student had to bookmark all relevant pages (3 per task), the recall measure of a student s was appropriate to express the end product of the search process. This measure was computed according to Equation 1. On the other hand, the search score of a student s was defined as a way to measure not only the end product but also some aspects of the search process. In particular, the search score takes into account mistakes as a result of strategies such as trial an error, which was observed in some students who took advantage of system's feedback after bookmarking and submitting random three pages (active bookmarks) to find out if they completed the task. The score is expressed in a 0 – 5 scale and it was operationalized as shown in Equation 2. A summary of descriptive statistics for these two measures is listed in Table 2.

To perform within- and between-subjects comparisons, we relied on the t-test and Wilcoxon test (pairwise and independent samples respectively) depending upon data distribution and variance. A confidence interval of 95% was used for all tests. Normality was tested through the Shapiro-Wilk test, whereas homoscedasticity was tested with the Levene test.

Table 2

Descriptive statistics for recall and search score per group.

	Recall			SearchScore		
	Mean	Median	SD	Mean	Median	SD
C1	0.7262	0.667	0.2935	2.8956	3	1.4379
E1	0.7164	0.667	0.3028	2.782	2.5	1.4326
C2	0.8602	1	0.2304	3.154	3.1665	1.3865
E2	0.8624	1	0.2354	3.4029	3.333	1.3483

Table 3

Search performance results (wr = Wilcoxon Rank Sum, ws = Wilcoxon Signed Rank, * = Significant result at $p < 0.05$).

	Recall			SearchScore		
	Statistic	p-value	Effect Size	Statistic	p-value	Effect Size
C1=E1	wr=36544	0.7771	0.0121	wr=37411	0.4528	0.0321
C1<C2	ws=1435	5.2141e-8*	0.3701	ws=6142.5	0.012*	0.1264
E1<E2	ws=3357	7.5465e-12*	0.3902	ws=10450.5	3.1626e-10*	0.3407
C2<E2	wr=35606	0.3782	0.0133	wr=32319.5	0.0179*	0.0899

To ensure the homogeneity between the distribution of students in the control and experimental groups before the teaching intervention, we first compared C1 and E1 during the pretest in terms of the performance measures listed on Equations 1 and 2. Our results showed no significant differences between the two groups.

Next, we conducted within-subject comparisons to determine the effects of the intervention in the experimental group. Results showed a significant increase in recall and search score in both groups ($p < 0.05$) (Table 3). Yet, between-subject comparison showed that search score achieved by the experimental group (E2) was significantly greater than that of the control group (C2) ($p < 0.05$).

4. Discussion

Our preliminary findings indicate at large that regardless of the intervention, both control and experimental groups exhibited performance improvements in terms of recall and search score (i.e., pretest-posttest within-subjects comparison). While the gain in the control group was not expected, this could be attributed to a learning effect as a result of the exposure to the NEURONE system during the pretest. Indeed, although the topics and domains of the tasks faced by students in the pretest and posttest were different, session structure and the system were the same, thus students were likely to grasp system features, session flow, and strategies to better address the tasks.

In spite of the performance improvement observed in both groups, we found that search score was signif-

icantly better in the experimental group (E2) than in the control group (C2) after the intervention (posttest between-subject comparison). We believe that such difference could be attributed in part to the teaching intervention to which students in the experimental group were exposed to.

We note that no significant differences were found with respect to recall during the posttest. This can be explained by the product-based nature of the measure itself. Recall indicates success (i.e., finding relevant documents), no matter how students achieved the goal (e.g., trial and error). On the contrary, the fact that search scores were higher in the experimental group (E2) illustrates that students in this group were more cautious when bookmarking pages avoiding making mistakes. This positive effect could be the result of the intervention, yet further analyses are needed to test this hypothesis.

Our future work will target analyses to find out whether the intervention had effects on performance at the level of task and domains. Beyond the scope of search performance, which was the main focus of this article, we will also look at the effects of the teaching intervention on search behaviors (e.g., query formulation, mouse actions) as well as other component skills evaluated in the performance test (i.e., identification of main ideas and synthesis).

Acknowledgments

The work described in this article was partially supported by the TUTELAGE project funded by the National Agency for Research and Development (ANID) (FONDECYT Regular, grant no. 1201610); the iFuCo project funded by the Academy of Finland (grant no. 294186) and ANID (grant AKA/EDU-03); ANID Scholarship Program Doctorado Nacional no. 2019-21192162 and Doctorado Becas Chile no. 7608-2020.

References

- [1] P. Stordy, Taxonomy of literacies, *J. Doc.* 71 (2015) 456–476. doi:10.1108/JD-10-2013-0128.
- [2] National Research Council, *Inquiry and the national science education standards: A guide for teaching and learning*, National Academies Press, Washington, DC, USA, 2000. doi:10.17226/9596.
- [3] C. Quintana, M. Zhang, J. Krajcik, A framework for supporting metacognitive aspects of online inquiry through software-based scaffolding, *Educ. Psychol.* 40 (2005) 235–244. doi:10.1207/s15326985ep4004_5.
- [4] American Library Association, *Information literacy competency standards for higher education*, 2000. URL: <https://alair.ala.org/handle/11213/7668>.
- [5] M. B. Eisenberg, R. E. Berkowitz, *Information Problem Solving: The Big Six Skills Approach to Library & Information Skills Instruction*, Big6 Information Literacy Skills Series, Ablex Publishing, Norwood, NJ, USA, 1990.
- [6] J. Delors, *Learning: The treasure within*, UNESCO, Paris, France, 1998.
- [7] B. Lym, H. Grossman, L. Yannotta, M. Talih, Assessing the assessment: how institutions administered, interpreted, and used sails, *Ref. Serv. Rev.* 38 (2010) 168–186. doi:10.1108/00907321011020806.
- [8] S. Majid, Y.-K. Chang, S. Foo, Auditing information literacy skills of secondary school students in singapore, *J. of Information Literacy* 10 (2016) 44–66. doi:10.11645/10.1.2068.
- [9] S. Foo, S. Majid, Y. K. Chang, Assessing information literacy skills among young information age students in singapore, *Aslib J. Inf. Manag.* 69 (2017) 335–353. doi:10.1108/AJIM-08-2016-0138.
- [10] M. Mikkilä-Erdmann, E. Sormunen, T. Mikkonen, N. Erdmann, C. Kiili, M. Quintanilla, R. González-Ibáñez, P. H. Leppänen, M. Vauras, A comparative study on learning and teaching online inquiry skills in Finland and Chile, in: *European Association for Research on Learning and Instruction*, Tampere, Finland, 2017.
- [11] E. Sormunen, R. González-Ibáñez, C. Kiili, P. H. Leppänen, M. Mikkilä-Erdmann, N. Erdmann, M. Escobar-Macaya, A performance-based test for assessing students' online inquiry competences in schools, in: *European Conference on Information Literacy*, Springer, 2017, pp. 673–682. doi:10.1007/978-3-319-74334-9_69.
- [12] R. González-Ibáñez, D. Gacitúa, E. Sormunen, C. Kiili, NEURONE: oNlinE inqUiRy experimentation systEm, *Proc. Assoc. Inf. Sci. Technol.* 54 (2017) 687–689. doi:10.1002/pr2.2017.14505401117.
- [13] E. K. Hämäläinen, C. Kiili, M. Marttunen, E. Räikkönen, R. González-Ibáñez, P. H. Leppänen, Promoting sixth graders' credibility evaluation of web pages: An intervention study, *Comput. Hum. Behav.* (2020) 106372. doi:10.1016/j.chb.2020.106372.