

TROLLING IN VIRTUAL LEARNING SPACES: A PROPOSAL FOR ANALYSIS FROM COLLECTIVE LEARNING

Sonia Valero-Tapia¹, Pablo Bautista², & Santos Orejudo¹

¹*Department of Psychology and Sociology, University of Zaragoza (Spain)*

²*Department of Educational Sciences, University of Zaragoza (Spain)*

Abstract

The emergence and evolution of the internet has brought about many changes in society, including in the educational paradigm. With the rise of Learning and Communication Technologies (LCT), the way students learn has changed. One example of this is online collaboration tools, such as the Collective Learning platform, which allows hundreds of students to group together anonymously and synchronously to solve activities on any subject. The aim of the platform is to reach consensus, make decisions and establish learning processes around different topics. However, despite the fact that students interact and communicate indirectly on the platform during the tasks, its purpose is far from that of the social networks in use. Despite this, it has been detected that on Collective Learning, as on other digital leisure platforms on the Internet, trolling behaviour can appear, i.e. behaviour initiated by anonymous users on digital platforms to annoy other users and hinder online communication. In Collective Learning, this behaviour is seen through trolling responses that some students make while doing activities on the platform. These types of responses hinder the development of tasks in Collective Learning and jeopardise the learning processes during them, and should therefore be analysed. In this study, we report on an experiment in which 265 students aged 12-16 collaborated online to solve a case about extreme popularity seeking on social networks. The aim is to detect the troll responses elaborated by the students during the task and to establish common characteristics among them. Our work shows how troll answers can appear in a task on socioemotional development on the Internet with students of Compulsory Secondary Education. These types of responses may reach a certain popularity within the platform, but their general tendency is to disappear as the activity progresses. Finally, with the study presented, it has been possible to establish a set of common characteristics among the troll responses on the platform that has allowed the development of a definition for them and thus open a more favourable path for their detection and mitigation in future activities carried out with Collective Learning.

Keywords: *Collective learning, trolling, collective intelligence, learning.*

1. Introduction

Over the last decade, the emergence of the Internet has transformed educational contexts to a large extent, especially for the very young (Du et al., 2024). In digital spaces, adolescents find numerous opportunities for social interaction (Cebollero-Salinas et al., 2022), both from platforms for leisure and social networks (Volkova and Bachrach, 2015) or even in online collaboration and learning platforms (Faraj et al., 2010; Malone et al., 2010), one of the latter being Collective Learning. Collective Learning is a platform based on the construct of collective intelligence, which assumes that a group of individuals performing a task will perform better than a single individual (Wolley et al., 2010). Collective Learning offers the opportunity to generate spaces for interaction in which hundreds of people can collaborate synchronously and anonymously to solve tasks on any topic through consensus building. With the activities created on the platform, learners work through seven phases. Participants answer certain questions about the activity posed and can see what answers their classmates give, copying from them partially or in full or keeping their initial answers. The phases are developed as follows:

- Phase 1. Students answer the questions in the solo activity.
- Phases 2, 3, 4 and 5. Students can see what their peers have answered and copy from other users' productions if they wish. In these phases the platform considers the most copied answers as the most popular and starts discarding the least copied ones.
- Phase 6. The most popular answers for each question are presented in a top 10, from which students can copy in part or in full if they wish.

- Phase 7. A final top 10 is presented for each question, students can only keep their answers or fully copy answers from the top.

Under this premise, several researchers (Bautista et al., 2022; Orejudo et al., 2022; Orejudo et al., 2024) have used Collective Learning as a platform on which to propose to hundreds of students the resolution of different activities related to socioemotional development and the prevention of digital risks on the Internet through the case methodology (Orejudo et al., 2008). These studies show positive results in terms of socioemotional development; however, it has been detected that while students perform these types of tasks on the platform, trolling behaviours may appear in the answers that students offer to solve the activities (Orejudo et al., 2022; Bautista et al., 2022; Valero-Tapia and Orejudo et al., 2024), these types of answers on the platform have been called troll answers.

Users who engage in trolling behaviour on the internet are called trolls, whose aim is to annoy and provoke other users (Bishop, 2012; Hardaker, 2010; Keita, 2023), hindering online interactions, in this case in Collective Learning. Troll responses in Collective Learning, which not only distance themselves from the objective of the proposed tasks, affecting interaction and cooperation between participants in the activities, but also pose a risk to collective learning (Denoni Buján and Casanova, 2024). In order to protect the learning processes in Collective Learning, studying troll responses in Collective Learning becomes a necessity. Therefore, the following research study is proposed, which aims to detect troll responses and their common characteristics elaborated by a group of 265 students during a task based on the case methodology on socio-emotional development in Collective Learning.

2. Objectives

The aim of this paper is to detect troll responses and their common characteristics elaborated by a group of 265 students during a task based on the case methodology on socioemotional development in Collective Learning. The hypothesis that accompanies this objective focuses on the fact that although troll responses can be detected on the basis of common characteristics, they may contain different content and present differential characteristics.

3. Method

3.1. Participants

The work presented here involved 265 Spanish students from the first to the fourth year of Compulsory Secondary Education (ESO), from four schools located in the Autonomous Community of Aragon. In Spain, ESO students range in age from 12 to 16 years old.

3.2. Instruments: The case

The study participants faced an activity designed from the case methodology (Orejudo et al., 2008) and developed in the Collective Learning platform. The case told a story about Tania, a teenager facing a contentious divorce between her parents and being bullied at school. At a critical moment, she posted images of herself self-harming on Twitter, threatening suicide if her classmates did not stop the bullying. Although initially the bullying diminishes, Tania repeatedly resorts to threats to blackmail her classmates, which eventually leads to a resumption of the bullying. This situation aggravates her emotional state, leading her to self-harm again and hospitalisation. The case was accompanied by five open-ended questions with no predetermined character limit, which alluded to Tania's possible motivations, the importance of the context, the sensitive content uploaded to social networks and the search for solutions to the protagonist's situation.

3.3. Procedure

Prior to the activity, all Secondary Education and Vocational Training schools in the Autonomous Community of Aragon were invited to participate. This session was part of a complete research program, which offered more sessions related to cyber coexistence and digital risk prevention, approved and validated by the Research Ethics Committee of the Autonomous Community of Aragon (CEICA: Comité de Ética de la Investigación de la Comunidad Autónoma de Aragón). The interested centers accepted the conditions of participation, which were to have the necessary time to carry out the activity and that each participating student had an electronic device with Internet access. Prior to the activity, both the participating students and their families were informed of the purpose of the study and the procedure through a letter of informed consent. This letter also explained the voluntary nature of participation and guaranteed the anonymity of the students. To carry out the activity, prior to the session, all the schools from which the participating students came were sent a password and user name for each student with which they could access the activity on the date and time set.

3.4. Data analysis

First, all the answers that the students proposed to solve the activity in the Collective Learning platform were extracted and ordered taking into account their time of appearance and the phase to which they corresponded. Once ordered and taking the conceptualization of trolling by Bishop (2012), Hardaker (2010) and Keita (2023), we proceeded to detect the responses considered as trolling responses by the students, i.e., those that did not respond to the context of the activity and that intentionally deviated from the objective of the activity. After detecting which responses met this identification criterion, an ad hoc empirical categorization system was designed to divide the responses that had initially been categorized as troll responses into more specific categories. The categorization system consisted of the following categories (table 1):

Table 1. System of categorization of troll responses developed.

Category	Name	Definition	Example extracted from the base
1	Absurd troll responses	They do not respond to the question or relate to the theme of the activity	"Salami"
2	Aggressive troll responses	They do not answer the question, express aggressiveness outside the context of the activity	"nerf to the purple dot. Fat girls with blue hair=bad"
3	Troll responses that allude to the face-to-face context	They do not answer the question, they establish a communication channel with the rest of the users alluding to known topics or people	"klk_guillermo is looking for a girlfriend to play with him"
4	Sarcastic troll responses	They do not respond to the objective of the activity, but answer the question sarcastically	"The opinion of the fat women"
5	Troll responses with viral elements	They do not respond to the objective of the activity, their main objective is to allude to songs, video games, memes or viral gossip on social networks	"I want to play fornite"
6	Popularity-seeking troll responses	They do not respond to the objective of the activity, they allude to their own profiles or channels on social networks seeking popularity or followers	"rrrubenn_010 follow me gpas"
7	Incoherent troll responses	They do not respond to the activity, they are meaningless.	"ðđh¶t←↓J€~β¶jt€ ¶đđ"đβ"

4. Results

In the proposed activity, 4950 responses elaborated by the students were recorded. Of the 4950 responses recorded, 404 of them (8.2%) were found to be troll responses, being more frequent in phase 7 and between phases 3 and 5 (table 2).

Table 2. Distribution of troll responses by phase.

Phase	NO TROL			TROL			Total
	N	%	Waste	N	%	Waste	N
1,00	1239	96.9%	7.6	40	3.1%	-7.6	1279
2,00	429	95.5%	3.0	20	4.5%	-3.0	449
3,00	285	86.6%	-3.6	44	13.4%	3.6	329
4,00	511	88.9%	-2.8	64	11.1%	2.8	575
5,00	1193	89.2%	-4.1	144	10.8%	4.1	1337
6,00	756	94.1%	2.6	47	5.9%	-2.6	803
7,00	133	74.7%	-8.5	45	25.3%	8.5	178
Total	4546	91.8%		404		8.2%	4950

Once the category system was applied, it was found that the category of troll responses with the highest frequency of occurrence was category 5, related to troll responses with viral elements (3.4%). Regarding the rest of the categories, the troll responses were evenly distributed among them (Table 3).

Table 3. Distribution of troll responses by category.

Number of category	Name of category	N	%
1	Absurd troll responses	39	0.8%
2	Aggressive troll responses	48	1.0%
3	Troll responses that allude to the face-to-face context	33	0.7%
4	Sarcastic troll responses	46	0.9%
5	Troll responses with viral elements	170	3.4%
6	Popularity-seeking troll responses	19	0.4%
7	Incoherent troll responses	49	1.0%
	Total	404	8.2%

5. Discussion and conclusion

The aim of this work was to detect the troll responses and their common characteristics elaborated by a group of 265 students during a task based on the case methodology on socioemotional development in Collective Learning. The hypothesis was that troll responses could be detected through common characteristics but that there were differential elements among them. In the case developed, 4950 responses were recorded, 404 of them being troll responses, which had a greater representation in the final phase of the activity and between phases 3 and 5. To determine what common characteristics the 404 detected troll responses shared, an empirical categorisation system created ad hoc was designed, which allowed categorising all the detected troll responses, allowing them to be grouped into subgroups of greater concreteness.

In this way, the detection and categorisation of troll responses allows us to establish a definition for them, and it can be pointed out that this type of response would be those that do not respond to the context of the activity, intentionally deviating from its objective and whose content contains absurd, aggressive elements, allusive to the classroom context, sarcastic, allusive to viral elements, aimed at seeking popularity and/or incoherent. Therefore, it has been possible to corroborate the hypothesis put forward at the beginning of this study. Finally, trolling behaviours, or in this case troll responses, are not only likely to appear in online leisure environments such as forums (Phillips, 2015), video games (Thacker & Griffiths, 2012) or social networks (Fichman, P., & Rathi, M., 2023), but can also appear in online learning contexts such as Collective Learning (Denoni Buján & Casanova, 2024). These responses break the dynamics of the activity, however the knowledge generated so far about trolling in learning platforms is limited to understand aspects such as the motivation of users or the profile of users who engage in these behaviours, being Collective Learning an environment with potential to deepen these aspects in future research such as the one presented in this paper.

Acknowledgments

This study was funded by Department of Science, University and Knowledge Society, Government of Aragon (Ref. PROY_S14_24. CONVIVEN-CI-IA. Aprendizaje de competencias socioemocionales para la ciberconvivencia a través de la integración de herramientas de Inteligencia Artificial (IA) + Inteligencia colectiva (IC) en centros educativos). Project financed by: FEDER Aragón Operational Program 2014-2020. Building Europe from Aragon.

References

- Bautista, P. Cano-Escoriaza, J., Sánchez, E., Cebollero-Salinas, A., & Orejudo, S. (2022). Improving adolescent moral reasoning versus cyberbullying: An online big group experiment by means of collective intelligence. *Computers & Education*, 189. <https://doi.org/10.1016/j.compedu.2022.104594>
- Bishop, J. (2012). The psychology of trolling and lurking: The role of defriending and gamification for increasing participation in online communities using seductive narratives. In H. Li (Ed.), *Virtual community participation and motivation: Cross-disciplinary theories* (pp. 160-176). Hershey: IGI Global.

- Cebollero Salinas, A., Bautista, P., Cano, J., & Orejudo, S. (2022) E-competencias socioemocionales e inteligencia colectiva. Propuestas para el desarrollo emocional en las interacciones en línea. *Revista Internacional de Educación Emocional y Bienestar*, 2(1), 13-32. Retrieved from <https://rieeb.iberomx.com/index.php/rieeb/article/view/21>
- Denoni Buján, M., & Casanova, O. (2024) Characteristics of Troll Responses Created by Primary School Students on The Collective Learning Platform. *INTED2024 Proceedings*, 6928-6933.
- Du, W., Fan, Z., Li, D., & Wu, M. (2024). Internet use behavior and adolescent mental health: The mediating effects of self-education expectations and parental support. *Psychology Research and Behavior Management*, 1163-1176. <https://doi.org/10.2147/PRBM.S449353>
- Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge Collaboration in Online Communities. *Organization Science*, 22(5), 1224-1239. Retrieved from <http://www.jstor.org/stable/41303115>
- Fichman, P., & Rathi, M. (2023). Trolling CNN and Fox News on Facebook, Instagram, and Twitter. *Journal of the Association for Information Science and Technology*, 74(5), 493-505.
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research*, 6, 215-242. <https://doi.org/10.1515/jplr.2010.011>
- Keita, M. (2023). Interactional effects of adverse childhood experiences, psychopathy and everyday sadism on Internet trolling. *Personality and Individual Differences*, 214, 3112327. <https://doi.org/10.1016/j.paid.2023.112327>
- Orejudo Hernández, S., Fernández Turrado, T., & Garrido Laparte, M. Á. (2008). Elaboración y trabajo con casos y otras metodologías activas. Cuatro experiencias de un grupo de profesores de la Facultad de Educación de Zaragoza. *Revista Interuniversitaria de Formación del Profesorado*, 22(3), 21-45.
- Orejudo, S., Cano-Escoriaza, J., Cebollero-Salinas, A., Bautista, P., Clemente-Gallardo, J., Rivero, A., Rivero, P., Tarancón, A. (2022). Evolutionary emergence of collective intelligence in large groups of students. *Frontiers in Psychology*, 13, 848048. <https://doi.org/10.3389/fpsyg.2022.848048>
- Orejudo, S., Lozano-Blasco, R., Bautista, P., & Aiger, M. (2024). Interaction among participants in a collective intelligence experiment: an emotional approach. *Frontiers in Psychology*, 15, 1383134. <https://doi.org/10.3389/fpsyg.2024.1383134>
- Phillips, W. (2015). *This is why we can't have nice things: Mapping the relationship between online trolling and mainstream culture*. MIT.
- Malone, T., Laubacher, R., & Dellarocas, C. (2010). The Collective Intelligence Genome. *IEEE Engineering Management Review*, 38, 38-52. <https://doi.org/10.1109/EMR.2010.5559142>
- Thacker, S., & Griffiths, M.D. (2012). An exploratory study of trolling in online video gaming. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL)*, 2, 17-33. <https://doi.org/10.4018/ijcbpl.2012100102>
- Valero-Tapia, S., & Orejudo, S. (2024) Evolution of Digital Stereotypes of Secondary School Students on The Collective Learning Platform. *INTED2024 Proceedings*, 6790-6797. <https://doi.org/10.21125/inted.2024.1783>
- Volkova, S., & Bachrach, Y. (2015). On predicting sociodemographic traits and emotions from communications in social networks and their implications to online self-disclosure. *Cyberpsychology, Behavior, and Social Networking*, 18(12), 726-736.
- Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330, 686-688.