Digitally Scaffolding Debate in the Classroom

Adrian Holzer
EPFL / University of Lausanne
1015 Lausanne, Switzerland
adrian.holzer@epfl.ch

Nava Tintarev
Delft University of Technology
2628 XE, Delft, the Netherlands
ntintarev@tudelft.nl

Samuel Bendahan
University of Lausanne
1015 Lausanne, Switzerland
samuel.bendahan@unil.ch

Bruno Kocher
University of Lausanne
1015 Lausanne, Switzerland
bruno.kocher@unil.ch

Shane Greenup
Rbutr
2000 Sydney, Australia
shane@rbutr

Denis Gillet
EPFL
1015 Lausanne, Switzerland
denis.gillet@epfl.ch

CHI'18 Extended Abstracts, April 21-26, 2018, Montreal, QC, Canada

ACM ISBN 978-1-4503-5621-3/18/04.
https://doi.org/10.1145/3170427.3188499

Abstract
Students increasingly have access to information that can be posted by anyone without being vetted, and it becomes vital to support students in evaluating claims through debate and critical thinking. To address this issue, this paper designs and evaluates a light-weight but effective protocol for supporting debate in a classroom activity with university students. It evaluated participants' beliefs on controversial topics (e.g., homeopathy) before and after briefly learning about critical thinking tools, posting arguments, and critically evaluating the arguments of peers. The findings suggest that this intervention led to a statistically significant belief change, and that this change was in the direction of the position best supported by evidence. Consequently, this work in progress presents a constructive approach to scaffold debates in the classroom and beyond.

Author Keywords
Critical thinking; HCI; Interaction systems; Scaffolded online debate

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces - Interaction styles

Claims
Homeopathy cures the flu.
The homicide rate in Switzerland is at its lowest since 1982. (supported)
Measles vaccine is a cause of autism.
Genetically modified foods (GMFS) cause cancer.
Organic products are significantly healthier than others.
A healthy child development requires a father and a mother (as opposed to two fathers or two mothers)
The ``nones" (people with no religious affiliation) are more numerous than any other religious group in Geneva. (supported)
On 9/11, the twin towers collapsed due to a controlled demolition

Figure 1: Claims presented to students in Step 1. The claims are not supported by evidence unless indicated otherwise.

Introduction

Critical thinking is a highly valued skill, but a difficult one to teach or nurture. The ability to cite evidence to justify one's belief is considered a primary form of scientific literacy that should be taught alongside scientific content [**?**]. In part, the difficulty in developing critical thinking may be due to beliefs forming and becoming fixed early in life, i.e., before higher education [**?**]. Unfortunately, critical thinking is a skill that is missing even among people holding a degree in a scientific field of study [**?**]. It is difficult to help adjust unfounded beliefs by simply pointing out alternative explanations. Indeed trying to correct those beliefs might even strengthen people's initial beliefs [**?**, **?**]. In particular when the argument threatens their identity or falls outwith their latitude of acceptance. This issue is called the *backfire effect*, and has been found to be challenging for services that aim to fact check information online [**?**].

A way to improve this outcome is to design systems that give the control back to the user, with sufficient scaffolding (i.e., support and guidance). It is thus crucial to teach skepticism and scientific thinking early as it is most likely to exert an influence at the time of message exposure [**?**].

In the classroom, an effective method to foster science-based reasoning is to ask students to themselves generate counter arguments for unfounded beliefs [**?**]. Teaching critical thinking skills and pointing out flawed argumentation techniques used by providers of misinformation has been shown to be effective to reduce belief in false information [**?**]. Presenting students with automated persuasive arguments has been evaluated in the educational domain in the context of the acceptance of potential learning items [**?**]. The results showed a slightly (unclear if the result was significant) increased item acceptance.

Some authors argue that online debate could reduce beliefs in pseudoscientific claims [**?**, **?**]. Possibly leveraging the fact that arguments from peers can be more persuasive than those coming from more authoritative figures [**?**]. Rbutr[1] is a software solution that scaffolds peer debates on controversial information right where it appears. It does so by allowing users to post and rate rebuttals for web pages through a browser plugin. As such any web page can become a live debate platform. This is in line with a view that there should be a World Wide Argument Web, connecting arguments with each other online (see Schneider et al. [**?**] for a review).

More broadly, Kriplean et al. argue that systems should be designed to reward participants not only to contribute but also to listen well (e.g., restating other people's comments [**?**], providing pros and cons [**?**]). One of the challenges of conducting a debate online, is the fact that they can lead to trolling behaviors, where people with extreme opinions take over the conversation [**?**].

We argue that at this stage, it is unclear whether scaffolded debates lead to more critical thinking and more evidence-based beliefs. To clarify this issue, we design and evaluate an innovative and light-weight debate intervention to foster more evidence-based beliefs.

Debate intervention protocol

We designed a debate intervention to support critical thinking in the educational context. A core idea of this protocol is to enable others to reuse it for educational or research purposes. Therefore, it is designed to fit in a 90-minute lecture with use of a simple classroom interaction tool, which does not require prior expertise to be used in a classroom.

[1] http://rbutr.com, retrieved January 2018

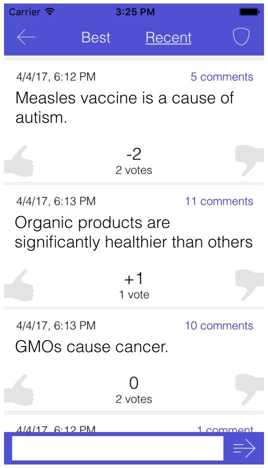

Figure 2: Example of the claims posted on SpeakUp by the instructors for Step 3.

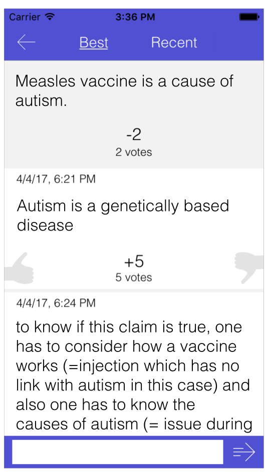

Figure 3: Comments on SpeakUp for a particular claim (i.e., vaccine causes autism)

Step 1: Exposure to controversial claims.
At the beginning of the lesson, students are exposed to a set of claims (see Figure 1). Some are supported by evidence, and some are not. Students are not told which of the claims are or are not supported by evidence. The claims are not just simple facts, but hot button topics in some parts of the world and people tend to feel strongly about them. The idea is to grab student attention that way.

Step 2: Delivering critical thinking lecture.
We designed a one hour lesson providing an overview of critical thinking tools useful to assess new information. This lesson highlighted the importance of verifying information for making decisions for one's private life as well as for public policy. The lecture argued that instead of giving answers on what information is correct, it was more useful to offer a toolkit for students for them to ask the right questions, inspired by Carl Sagan's baloney detection kit [**?**, **?**]. We focused on two main questions when faced with a claim, i.e., *how extraordinary are the claims?*, *how reliable is the source?*

Step 3: Students provide arguments.
At the end of the lesson, students are given time to think about and post arguments supporting or rejecting *two claims* presented previously on the SpeakUp app[2] (see Figures 2-4). SpeakUp is an application that can be seen as a shared group chat app where users can post and rate messages (using the thumb up/down buttons) anonymously. Instructors can create a chat room and students can join by indicating the room number (they thus do not need to register and provide personal information). The tool was designed to make it effortless for students to interact in large classrooms [**?**].Anonymity can be useful to make it easier for students to express their views as they will not be judged for them. Furthermore, it forces students to focus on the content of the posts rather than on the person posting. The fact that the digital artifact is easy to use is an important aspect to make the intervention transferable to others. Figure 2 shows how instructors first post messages containing the claims on the app. Students then write their arguments in the comments of the claim, as shown in Figure 3.

Step 4: Students analyze claims critically.
Students are given time to read the arguments posted by other students and use the use the thumb up/down feature of the SpeakUp app to express their opinions on arguments that they judged strong or weak.

Evaluation in the wild

We used a within-subjects design (with pre- and post-test) in a real classroom setting and analyzed data logs from SpeakUp (the app provides an extract option that provides all activity traces within a room). Participants were recruited in a lesson of the Global Issues course in Communication at EPFL in Switzerland. This course is delivered to 100 first year undergraduate engineering students (26 female). The pre-test took place between Step 1 and Step 2, students were asked to indicate their beliefs about the controversial claims on a nine-point scale from 1 (I have no doubt that the claim is correct) to 9 (I have no doubt that the claim is false) on the SpeakUp classroom interaction app (Figure 4). Note that the results of the votes were not shown to the students until after the lesson. The post-test was conducted after Step 4, students were again asked to indicate their beliefs on the set of claims using the 9-point scale on the SpeakUp app. Only when all votes were cast, did we reveal the results of both votes.

[2] `speakup.info`, retrieved January 2018

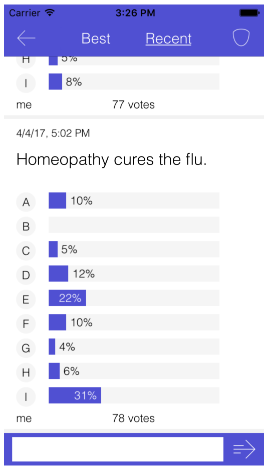

Figure 4: Example of the multiple choice answers in SpeakUp revealed to students after the lesson was over.

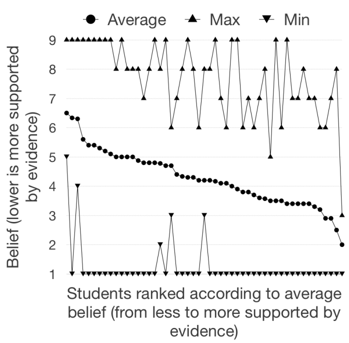

Figure 6: Minimum, average, and maximum initial belief per student (1 = I have no doubt that the claim is true, 9 = I have no doubt that the claim is false).

Results

Analyzing the application logs, a total of 97 students used SpeakUp, 93 of them voted and 55 of them posted messages. A total of 2543 votes were cast, and a total of 165 messages posted during the lesson. The 52 users who completed both votes are included in the analysis below.[3]

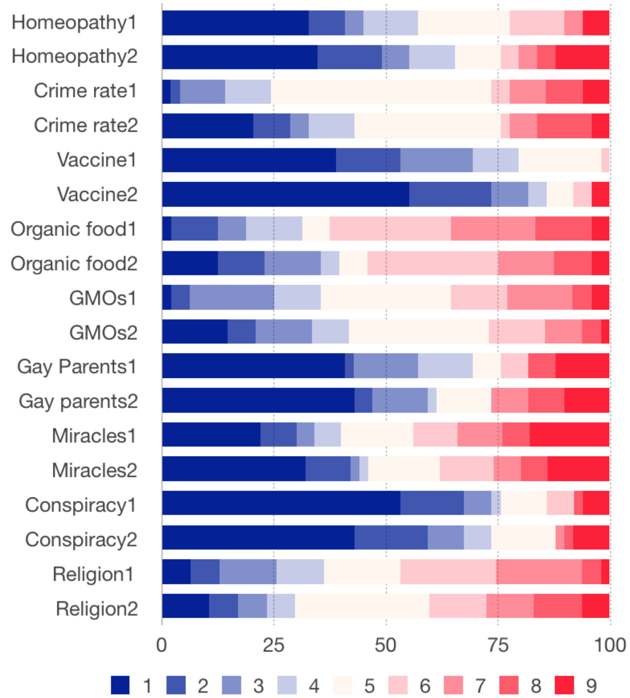

Figure 5: Student opinions.

Figure 5 shows a stacked plot of the opinions of students before and after the intervention. For clarity, we reversed the scores for all claims not supported by evidence in the analysis. This means that for the analysis and in Figure 5, beliefs that the claims are correct (scores from 1 to 4) are supported by evidence. Each bar is divided in segments of different shade representing the beliefs from 1 (I have no doubt that the claim is true) to 9 (I have no doubt that the claim is false). Bars with number 1 (e.g., GMOs1) reflects the first belief (pre-test after Step 1), and with 2 (e.g, GMOs2) reflects the second belief (post test after Step 4). The plot shows that most students (>50%) were on the evidence supported side of four issues (Homeopathy, Vaccines, Gay parents, 9/11 conspiracy). Most students however viewed organic food as significantly healthier than regular food, and more students were on the side not supported by the evidence for the claims about GMOs, crime rate, miracles and religion.

[3] Our data is openly available here: goo.gl/ZjlUE9

Figure 6 shows the minimum, average, and maximal initial belief per student. It shows that most students are on average on the evidence supported side of the issues (value < 5). Nevertheless, it shows relatively large differences between average ratings between students (range=6.5–2.0) and also it shows that most students (56%) held belief ranging from 2-8, and 30% held beliefs ranging from 1-9.

Were extreme students more likely to comment? (yes)
We performed a quadratic regression with cluster robust standard errors. The DV of the regression was the fact that a student had provided a comment (true, false) and the IV was their initial opinion. Students with extreme opinions were more likely to comment ($F(2,51)=6.40$, $p<0.01$).

Did the debate cause more polarized opinions? (no)
One risk of exposing people to both sides of an argument is that they become even more established in their existing views. Figure 7 illustrates the average opinion shift according to initial opinions. A negative shift indicates a shift towards a more evidence supported opinion. We designed

a structural equation model to determine whether extreme opinions tended to become more polarized after commenting but failed to find a significant effect. However, absence of evidence is not evidence of absence, and this lack of evidence could be due to a lack of power as the effect sizes are quite small.

Did students change their minds? (yes)
We performed a linear regression with cluster robust standard errors to determine whether the intervention had an influence on the opinion. The DV of the regression was the opinion and the IV was time (before, after). We find a statistically significant shift towards the position that has more empirical evidence ($r=-0.28$ $F(1,51)=4.29$, $p<0.05$). This opinion shift (toward 1, more negative differences in opinions) can also be seen in Figure 7.

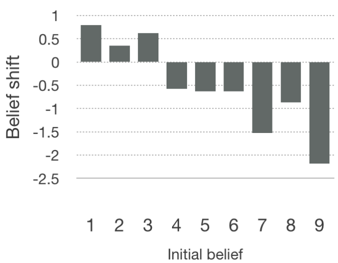

Figure 7: Average belief shift (post-pre). Negative values indicate a shift to a more evidence supported belief.

Discussion, conclusion and outlook

The study has several limitations, which are inherent to evaluation in a live classroom setting. For instance, it meant a large dropout rate (ca 30%). Replications of the study should be conducted to further validate the findings across different educational and cultural backgrounds. Furthermore, the intervention combined several aspects (Steps 1-4), which does not allow us to draw conclusions on the effects of each one individually.

In our study, we asked people to evaluate claims and to give arguments supporting or opposing them. We observed that students commented more on topics that they had strong opinions about. This might be what one expects to see in online communication and raises the question of the importance of investigating ways of nudging users with more moderate opinions to be more active in online debates. In this study we picked topics that are controversial and that threaten students' values by putting them into question. Despite this, we did not observe a shift to more polarized opinions. This could be because students are less threatened by the opinions of peers compared to that of instructors. However, it could also simply be that the lack of power in the analysis prevented us from seeing an effect. Thus, further evidence is required to draw an informed conclusion. We also found a statistically significant change in opinion toward the position that has most support. This can be seen as a surprising result given the simplicity of the intervention. The fact that the intervention is simple make it possibly transferable to other contexts. This would allow to provide support for critical thinking outside of academic institutions for a general public who may not have access to formal education.

In future work we plan to evaluate a number of factors relating to the overarching research question: *does exposure to critical analysis improve critical thinking skills*. Our study design is informed by three principles: training students in critical thinking, active learning, and peer review. We plan to study in a more controlled way how much each design principle contributed to critical thinking. We are also investigating how additional tools can be used to scaffold critical thinking and comparing with more classical offline classroom debate interventions in two courses.

References

[1] Kia Aarnio and Marjaana Lindeman. 2005. Paranormal beliefs, education, and thinking styles. *Personality and Individual Differences* 39, 7 (2005), 1227–1236.

[2] John Cook, Stephan Lewandowsky, and Ullrich KH Ecker. 2017. Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE* 12, 5 (2017), e0175799.

[3] R Kelly Garrett. 2011. Troubling consequences of online political rumoring. *Human Communication Research* 37, 2 (2011), 255–274.

[4] Stella Heras, Paula Rodríguez, Javier Palanca, Néstor Duque, and Vicente Julián. 2017. Using Argumentation to Persuade Students in an Educational Recommender System. In *International Conference on Persuasive Technology*. Springer, 227–239.

[5] Adrian Holzer, Sten Govaerts, Samuel Bendahan, and Denis Gillet. 2015. Towards Mobile Blended Interaction Fostering Critical Thinking. In *MobileHCI'15*. ACM, 735–742.

[6] Adrian Holzer, Sten Govaerts, Andrii Vozniuk, Bruno Kocher, and Denis Gillet. 2014. Speakup in the classroom: anonymous temporary social media for better interactions. In *CHI'14 EA*. ACM, 1171–1176.

[7] Travis Kriplean, Caitlin Bonnar, Alan Borning, Bo Kinney, and Brian Gill. 2014. Integrating on-demand fact-checking with public dialogue. In *CSCW'14*. ACM, 1188–1199.

[8] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *CSCW'12*. ACM, 265–274.

[9] Travis Kriplean, Michael Toomim, JT Morgan, Alan Borning, and AJ Ko. 2011. REFLECT: Supporting active listening and grounding on the Web through restatement. In *CSCW'11*.

[10] Stephan Lewandowsky, Ullrich KH Ecker, Colleen M Seifert, Norbert Schwarz, and John Cook. 2012. Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest* 13, 3 (2012), 106–131.

[11] Richard L Miller and William Wozniak. 2001. Counter-attitudinal advocacy: Effort vs. self-generation of arguments. *Current Research in Social Psychology* 6, 4 (2001), 46–57.

[12] Hien Nguyen, Judith Masthoff, and Peter Edwards. 2007. Modelling a receiverâĂŹs position to persuasive arguments. *Persuasive Technology* (2007), 271–282.

[13] Jarkko Paavola and Harri Jalonen. 2015. An approach to detect and analyze the impact of biased information sources in the social media. In *In ECCWS'15*. Academic Conferences Limited, 213.

[14] Carl Sagan. 2007. The fine art of baloney detection. *Paranormal Claims: A Critical Analysis* (2007), 1.

[15] Carl Sagan. 2011. *Demon-haunted world: science as a candle in the dark*. Ballantine Books.

[16] Jodi Schneider, Tudor Groza, and Alexandre Passant. 2013. A review of argumentation for the social semantic web. *Semantic Web* 4, 2 (2013), 159–218.

[17] Andrew Shtulman. 2013. Epistemic similarities between students' scientific and supernatural beliefs. *Journal of Educational Psychology* 105, 1 (2013), 199.

[18] Chun-Yen Tsai, Chih-Neng Lin, Wen-Ling Shih, and Pai-Lu Wu. 2015. The effect of online argumentation upon students' pseudoscientific beliefs. *Computers & Education* 80 (2015), 187–197.

[19] Daniel T Willingham. 2008. Critical thinking: Why is it so hard to teach? *Arts Education Policy Review* 109, 4 (2008), 21–32.