Towards evolving collective explanations

Xavier Coiteux-Roy
coitex@usi.ch

CQI blog.

The vision. Here is a rather common dream: You are approaching a scientific paper. You look it up online on a website — the *open-science community platform*: There, associated to the paper, you find a list of user-made contributions ("posts"). The most popular post is a short YouTube video explaining the main idea of the paper, made by some user named Alice321. In the comment section of the post is a discussion started by BobTheCat, maybe to clarify some points about the video or to correct its inaccuracies. Other posts by various users feature multiple summaries (one of which is a cartoon), diagrams, step-by-step explanations, questions (often paired with answers), and even a quiz to test your understanding. Thanks to all of this collectively made content, your own study of the paper is faster, and also *deeper*. And what you find lacking, you can always contribute yourself, or ask in the open — "How did the authors jump from Equation 5 to Equation 6", "Has this open problem been solved in follow-up work?" — people expressing expertise or interest in that paper might answer you quickly.

And even more. This hypothetical platform already sounds a promising learning tool and a supportive seed for original research. But let us dream further, for there is something even deeper than explaining well papers that such a collective platform could achieve. A meta effect. Since contributions are user-made and everyone can participate, there is bound to be a diversity in the media and types of explanation used. If the system is well designed, such that good explanation methods get more attention, the popular explanations associated to a paper are progressively improved and extended, while the bad contributions are mostly ignored and quickly forgotten. On a higher level, the successful methods are selected and replicated

for other papers, and globally, with time emerges from the collective platform an understanding of not only what explanations are good, but also of what *kind* of explanations are good, and which are not. This neo-Darwinian evolution of explanations (and of meta-explanations!) can transform the way we think about scientific ideas, and a good collective open-science platform could well induce a drastic change in the scale of that evolution.

Why it often fails. The open-science vision is at the base of many modern initiatives, like technical wikis and community peer-review platforms. But those initiatives often fail short of achieving their grand goal.

The problem is well summarized in Michael Nielsen's book *Reinventing discovery*: While most scientists morally support collective initiatives; they seldom contribute themselves. The problem is often the so-called *publish or perish* motto of modern academia — in present times, writing papers and writing grant proposals are so crucial to an academic career that spending time on anything else seems unproductive to the active researcher: Why would they "waste" time explaining papers?

How it could succeed. The *publish or perish* problem is infamously difficult to solve. While I believe it can and *should* be addressed, I want to follow a different avenue here: *I suggest to incorporate collective open science into the educational curriculum* (of both undergraduate and graduate students). More precisely, why not turn (some) class activities and assignments into open-science participation and contributions? Instead of sweating over an assignment that only one TA will ever read, the student could for example contribute to expanding the collective knowledge and understanding of the lecture's field (I offer one concrete experiment in the next section).

For the student, such an approach is motivating as they can from very early on participate in the collective science activity through small contributions[1]. These contributions could take various forms

[1] There is a link to be made here with *competency-based education*.

depending on each student's style: addition or improvement to the extensive wiki textbook; step-by-step guides, YouTube tutorials; flashcards — creativity is not to be discouraged[2]. Through that freedom, students are personally invested: a key element to success.

And the obstacles to student participation are lesser than the ones linked to active researchers, for students taking courses can be easily given an extra incentive to participating in the open-science assignments — better grades. Unlike national grant agencies' policies and university-wide practices, the grading protocols are usually decided uniquely by the lecturer and TAs; no concerted global revolution is needed. The flexible format allows for trial and experimentation — in other words, the classroom is a good ecosystem for evolution.

One modest proposal to experiment. One common evaluation format, which I have come across at the three universities I have had the opportunity to TA at, is to ask students to study an important research paper and to present it to their peers (usually through a slideshow presentation). The presentations coming out of this seminar format are often impressively informative; it does not matter that the involved students are fresh undergrads. In fact, I remember Prof. Gilles Brassard, while attending Prof. Stefan Wolf's quantum information seminar for undergraduate students at ETH Zürich, saying after the presentation by two students (Christoph Müller and Fabio Streun) of a paper he had himself co-authored:

> This was the best presentation of this work[3] that I have ever attended, my own included!

[2] In fact, I am sure that there already exist powerful tools online that can be harnessed to expand and communicate better our thoughts. I would love to learn about them and see what they could bring to serious research and higher education.

[3] The paper, *Parallel lives: A local-realistic interpretation of "nonlocal" boxes* (`https://arxiv.org/pdf/1709.10016.pdf`), exposes one key idea of Paul Raymond-Robichaud's PhD thesis. Interestingly, the main argument is also summarized in an appendix by a poster featuring neat spaceships — an artist was commissioned for the art — and many times in conferences, I have heard Paul's idea referred to as "Ah yes, the argument with the spaceships!" It is a good example that art has its place in creating effective medium to develop scientific thinking.

Prof. Brassard then asked permission to incorporate the students' slides into his own presentations and courses, resulting in an even better exposition. But this case is an exception rather than the rule! Usually, while it is obvious that some students spend a lot of time preparing for the class, the expert understanding they produce is somewhat wasted, for the audience is limited and the ephemeral presentation ends up (often sooner rather than later) forgotten. Without permanence, which is necessary to the evolution of ideas, the explanatory improvements that the students might have brought risk slipping into oblivion (see Figure 1).

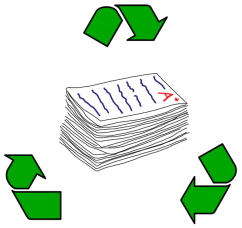

Fig. 1: At the end of the semester, the community of students has usually produced a huge stack of assignments. But while their paper is recycled, their content is often thrashed.

But what if the students were evaluated on their participation to a collective open-science platform such as the one envisioned above. The final contribution to the seminar could then be simply to show and explain what those contributions are (summaries, videos, concept maps, etc.). The platform would then harness the students' expertise and turn it into evolving explanations that could reach a much larger and time-lasting audience, and serve as a nucleus for a future, clearer and deeper, collective understanding.

And even if this particular initiative does not globally catch on: Let us say that, for one given topic, only three classes of 50 students participate. Over 10 years, that is still around 1500 students that will have worked successively on improving the explanations of that

topic's most important papers and concepts. What will be the result? New platforms? New explanations? New vectors of explanations?

I believe it is worth finding out.