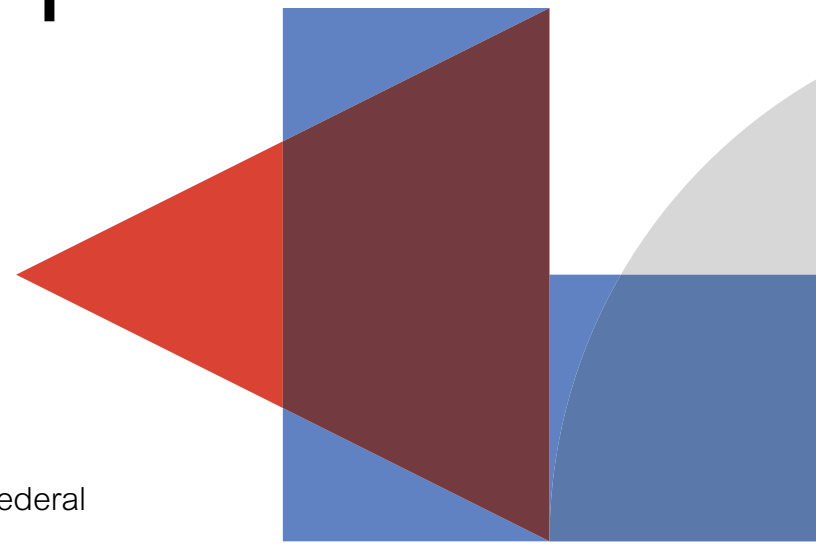


Essay 12

# The Challenges of Scaling- up Findings From Education Research

**Pierre Dillenbourg**

Professor of Learning Technologies, Swiss Federal  
Institute of Technology, Lausanne (EPFL)





The world of education is stratified into layers (class, school, district, state, country) that differ in orders of magnitude or scales (20, 200, ..., 200 million).

One would expect each layer to be simply the aggregation of smaller-scale units, forming a smoothly continuous system. In this essay, I argue that this is not the case and disentangle causes of this discontinuity. Why do the phenomena that appear at the micro-level not propagate to the meso-, macro- or global level? I consider the case of my research on education, which concerns learning technologies.

One could expect that a learning technology, once proven successful in a rigorous classroom experiment, and then in many classrooms, can be brought to larger scale through policy decisions which would then be propagated 'down' to all classrooms. This rarely happens. I present some systemic, though contradictory biases and myths as my explanation for the observed failure to scale up successful learning technology.

A first remark is that this discontinuity simply reflects the segmentation of learning sciences: Educational neurosciences consider brain areas, cognitive psychologists look at learning processes of individuals, social psychologists study teams of learners, education researchers deliberate classrooms, sociologists research school cultures, education economists research national funding schemes, and political scientists research governance principles. Similar fragmentations occur in other fields, but the specific difficulty to connect the dots across education scales hinders the impact our research has on education quality and reality.

A second explanation is that when it comes to learning technologies, there are two systemic opposing biases. On the one hand, we have "gurus" who systematically overstate the positive outcomes of technologies. Seymour Papert has been one of them: A genius in his work but who promised more effects from using LOGO than what teachers could ever achieve. On the other hand, we witness groundless fears that

technology will damage education: “They need a space away from the space of digital technology,” declared Giles Scott about his students in *The Washington Post*.<sup>1</sup> Let’s consider the ‘gurus’ first, specifically in the context of Massive Open Online Courses (MOOCs). The early hype generated over-expectations, disconnected from empirical data. The example of one kid ( $n=1$ ) from Mongolia who ends up at MIT occupied the media<sup>2</sup> more than the fact that most participants ( $n>1$  million) were white males in their 30’s living in urban areas. Conversely, this second fact should not give ground to negative attitudes: when my university provides a high-level MOOC on digital signal processing, it is completely normal that most participants have a bachelor or a master degree, since this topic is neither relevant nor accessible in earlier study paths.

Many critical voices emphasize the attrition rate of MOOCs (which simply reflects that “registration” does not require more than one click); this criticism neglects the millions of students who received education and acquired certificates from top universities. Regarding the use of MOOCs on campus, we found that students who actively participate to a specific MOOC in Physics, in addition to on-campus activities, obtain significantly higher marks on an exam.<sup>3</sup> It is common sense: They engaged more with the learning material. And the relationship may not be causal: It may simply be that better students use the MOOC more since they tend to use any resources we provide them with better. But common sense findings or modest results do not reach headlines; they get lost in translation. On the fear side, our students were strongly against MOOCs, claiming that they would lose contact with teachers. Two years later, the fear was proven unfounded.

In summary, when an empirical finding at scale X is projected to scale Y, it is distorted by unjustified expectations and fears that get more attention than facts. Is this distortion specific to EdTech? No, not even specific to education. As humans, we tend to hear what we are ready to listen to and to see what we expect to see. We also tend to overestimate the significance of exceptions. There are however a few phenomena specific to EdTech research that boost these distortions.

## **THE MYTH THAT LEARNING EFFECTS ARE INTRINSIC TO SOME TECHNOLOGY**

Let’s start with an example from our work on educational robotics. I am often asked about the effect of robots on learning. The answer is simple:



none! In one of our projects, the kids have to teach a small humanoid robot how to write, or to stop it when it makes reading mistakes. They don't learn because they interact with a robot but because they do something demanding with a robot. It is a four dots chain: a technology (dot 1) enables some activities (dot 2) which trigger some cognitive mechanisms (dot 3) that generate learning outcomes (dot 4). *There is no transitivity, no shortcut between dots 1 and 4.*

The same answer holds every time somebody asks what is the effect of MOOCs, or of augmented reality, or any other technology used in learning. I often reply by stating that good MOOCs are better for learning than bad MOOCs. This superfluity effectively illustrates the myth I want to emphasize: it is quite possible to produce a MOOC in which no participant learns anything. What is a good MOOC? Not one with Spielberg-produced superb videos but a MOOC in which learners face rich problem-solving activities (between good videos), at the right, for each learner, level of difficulty. In other words, technologies have affordances, i.e. potential effects; designers are supposed to turn them into actual effects. This myth is behind most failures of large scale educational technology deployment. It explains why the effects observed in small-scale experiments do not scale up: simply stated, **scaling up fails because it is not the device per se that generates effects, it is the activities students perform with this device.**

## THE MYTH OF INNOVATION

EdTech are often presented as pedagogical innovations, even if they have been around for decades. I cannot deny that if some robots are used for the first time in a classroom, this constitutes some kind of novelty. Now, if I teach with an egg on my head, is this innovative? Probably, but irrelevant—unless it is dance that I am teaching! Let's avoid the term "innovation" and instead focus on technology uses that address educational problems. For instance, carpenter bosses asked us if we could teach apprentices notions of statics for roof structures without any mathematical formulae. We responded by developing an intuitive augmented reality environment. Whether it is innovative is irrelevant; it only needs to be educationally effective. This strife for innovation may be justified only by the need to resist an even stronger myth encapsulated in "school used to be better in the olden days." Humans have this amazing selective memory that makes them prone to nostalgia: there were so many bad things about the school of our parents, but what gets attention is that pu-

pils made fewer spelling mistakes at that time—the few ones who did get through school. Between nostalgia and innovation, there is space for mature discourse. Now, if I had to, I would rather err towards the innovation side: I prefer a school culture that favors trying new methods, exploring new paths, keeping all actors engaged in efforts to make schools lead the changing society rather than being dragged into it.

## **THE TEMPTATION TO BLAME TEACHERS**

I often hear that the difficulty in scaling up is due to teachers, sometimes also that it is the fault of learners, administrators and/or parents. Let's consider the case of teachers. First, they are accused of technophobia. In western countries, teachers cannot order an airline ticket, a concert ticket, or save their holiday pictures without using technology. They have no issue with technology in general but, if we suddenly introduce 20 tablets into their classroom, we actually create a problem, a super-competitor to kids' attention. Nobody likes a tool that makes their work more difficult. Are teachers resistant to change? Yes, as we all naturally are. One could even argue that they have chose a routine job. But, I suggest we stop blaming teachers, and instead undertake the responsibility ourselves. Technologists: our technologies are not always suitable for schools—for a mountain road one needs a Land Rover, not a Rolls Royce. Consider the login process. When learners have to login, problems arise. Jennifer forgot her password while George pressed the caps lock key. This can waste 5 minutes of the lesson time (i.e. 10 percent, with absolutely no usefulness). We design technology for learners, but forget that a classroom is more than a set of individual learners; it is a collective body orchestrated by a teacher who has to manage multiple constraints: time, discipline, noise, safety, curriculum. Recent work on classroom orchestration has stressed the notion of classroom usability<sup>4</sup>: How demanding is it for a teacher to use digital tools while handling 25 or more kids and all the daily classroom constraint?

## **TESTING TECHNOLOGIES IN CLASSROOMS IS NOT ENOUGH**

Difficulties of scaling up also stem from very small differences between the method or technology used in classrooms during empirical studies and what can be generalized to larger scales later on. There are well-documented methodological biases in experiments, such as the



Hawthorne effect (subjects over-perform because they know they are being observed) or the Rosenthal effect (subjects perform according to their perception of the experimenter's expectations). But, there are also subtler differences that prevent generalization. Teachers who decide to join an empirical study are not selected randomly, neither are they typical teachers; there is a self-selection bias. For an experiment, they might come 15 minutes before the class to prepare the material, while they usually come only a few minutes before. They might allocate 60 minutes to a curriculum item to which they usually allocate only 30. Even if we do our best as researchers to set up ecologically valid conditions, our presence alone changes any education routine and thus distorts the observed effects. Once the same tool or method is used routinely, innovation wares off and these small differences may lead to abandoning the method or tool. Learning science has to integrate these practical details into the understanding of what is a scalable learning technology.

## **MODELLING THE TEACHER EFFECT**

Education research is condemned to generalization: if the learning outcomes obtained with a set of learners cannot be extended to a population of learners with a similar profile, the study is of little use. To reach the necessary sample size for statistical generalization, we conduct the same experiment in multiple classrooms, hoping that the same effect will occur across classrooms. Consider an experiment in which a technology is tested and the researchers find the same effect in each classroom. The researchers would be happy to see that this effect is invariant despite the difference of learners and teachers. In learning science publications, they would be pleased to state that “there is no teacher effect.” Is that really good? How could we expect any engagement of a teacher if we claim that, with some technology, there is no teacher effect? Education is not like medical research: there is not double-blind approach, a teacher may not be unaware of his teaching method. The same point can be made about the school effect: we expect that the same method would work independently of the school, as if an efficient school management were non-significant. One approach to generalization is to consider these variables, teacher-excellence or school-management, as noise and expect noise to cancel out when measuring many teachers and many schools. Another approach to generalization is to model these effects in greater detail (i.e. describe a teacher in our statistical or computational models, not with one, but with many variables that account for their behavior in class).

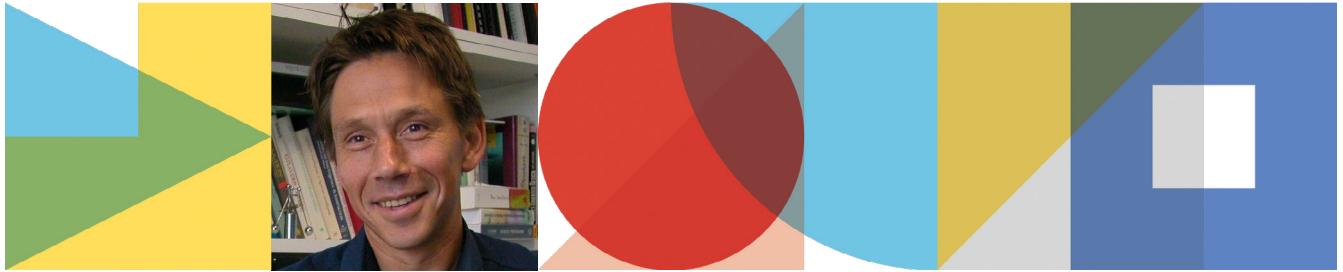
The problems reported in this contribution are well known in learning sciences. I do not have a magic bullet for them, but some hopes based on new developments in data sciences. Which activities produce learning for an individual is a fundamental question, but answering this question is not sufficient to succeed in scaling up. Scalable models have to integrate many more variables such as classroom logistics, teacher behavior, and school management. I expect that learning analytics will allow us to model rich learning situations, in real classrooms, and to integrate parameters specific to each scale of the education system.



## ENDNOTES

- 1 Strauss, Valerie. "High school teacher: I'm banning laptops in class — and not just because they are distracting." *The Washington Post*. August 5, 2016. <https://www.washingtonpost.com/news/answer-sheet/wp/2016/08/05/high-school-teacher-im-banning-laptops-in-class-and-not-just-because-they-are-distracting>.
- 2 Pappano, Laura. "The Boy Genius of Ulan Bator." *The New York Times*. September 13, 2013. <http://www.nytimes.com/2013/09/15/magazine/the-boy-genius-of-ulan-bator.html>.
- 3 Center For Digital Education (CEDE) Ecole Polytechnique Fédérale De Lausanne (EPFL). *MOOCs Annual Report – 2015*. March 2016. [https://documents.epfl.ch/groups/m/mo/moocs/www/MOOCs\\_Report\\_2015.pdf](https://documents.epfl.ch/groups/m/mo/moocs/www/MOOCs_Report_2015.pdf).
- 4 Dillenbourg, Pierre et al. "Classroom orchestration: The third circle of usability." *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings. Volume 1 — Long Papers, vol. 1*. January 2011. [https://www.researchgate.net/profile/Pierre\\_Dillenbourg/publication/228913602\\_Classroom\\_orchestration\\_The\\_third\\_circle\\_of\\_usability/links/004635162f51d79765000000.pdf](https://www.researchgate.net/profile/Pierre_Dillenbourg/publication/228913602_Classroom_orchestration_The_third_circle_of_usability/links/004635162f51d79765000000.pdf).





## Pierre Dillenbourg

Professor of Learning Technologies, Swiss Federal Institute of Technology, Lausanne (EPFL)

A former teacher in elementary school, Pierre Dillenbourg graduated in educational science (University of Mons, Belgium). He started his research on learning technologies in 1984. He obtained a PhD in computer science from the University of Lancaster (UK), in the domain of artificial intelligence applications for education. He has been assistant professor at the University of Geneva. He joined EPFL in 2002. He is currently full professor in learning technologies in the School of Computer & Communication Sciences, where he is the head of the CHILI Lab: “Computer-Human Interaction for Learning & Instruction.” He is also the academic director of Center for Digital Education, which implements the MOOC strategy of EPFL. EPFL recently passed over 1.5 million MOOC registrations. He wrote a book entitled “Orchestration Graphs” that proposes a formal language for instructional design (EPFL Press). With EPFL colleagues, he recently launched the Swiss EdTech Collider, an incubator with 30+ start-ups in learning technologies.