Reflection for Action: Designing Tools to Support Teacher Reflection on Everyday Evidence

L. P. Prieto[a] and Paul Magnuson[b] and Pierre Dillenbourg[c] and Merike Saar[a]

[a]Tallinn University, Narva mnt 29, 10120 Tallinn (Estonia); [b]Leysin American School, Chemin de La Source 3, 1854 Leysin (Switzerland); [c]École Polytechnique Fédérale de Lausanne, 1015 Lausanne (Switzerland)

ABSTRACT
Improving educational practice through reflection is one of the most widespread foci of teacher professional development (TPD) approaches. However, such teacher reflection operates under practical classroom constraints that make it happen infrequently, including the reliance on disruptive peer/supervisor observations or recordings. This paper describes three design-based research iterations towards technological support for teacher reflection based on everyday evidence and feedback. We collaborated with 16 teachers from two different secondary schools, using a variety of prototype technologies (from paper prototypes to web applications and wearable sensors). The iterative evaluation of such prototypes lead us from a high-tech focused approach to a more nuanced socio-technical one, based on lightweight technologies and 'envelope routines' that also involve students. After illustrating the potential of this approach to change teacher practice and students' learning experience, we also present a series of guidelines for the design of technology that supports such reflection based on everyday evidence gathering.

KEYWORDS
Teacher professional development; teacher reflection; technology design; design-based research

1. Introduction

1.1. Effective teacher professional development (TPD) and reflection approaches

Teacher professional development (TPD, often just referred to as PD) can be defined as 'the professional growth a teacher achieves as a result of gaining increased experience and examining his or her teaching systematically' (Glatthorn, 1995, p. 41), and is one of the main strategies for the improvement of student learning (Villegas-Reimers, 2003).

Teacher professional development is widely accepted as a way to foster improvements in teaching. However, there is little consensus on how it should work, nor is there a single overarching theory of teaching and teacher learning (Kennedy, 2016). Among the multiple legitimate approaches to TPD out there, Guskey (1994) notes

---

that effective TPD needs to balance general practices that are likely to be associated with the most effective professional development, with other practices that are specifically apt for different situations/contexts, e.g. attending to both the individual and the organization, supporting teams of practitioners, emphasizing feedback and follow-up, and integrating programs. Other important characteristics of effective TPD include the need for 'just-in-time, job-embedded assistance' (Guskey & Yoon, 2009, p. 497) and TPD that is collaborative, collegial, and sustained (Darling-Hammond & Richardson, 2009; Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009).

Collaborative and collegial professional development, sustained over time, both originates from and spawns ongoing reflection. Schon (1989) emphasized the importance of reflection while in the midst of practice (i.e., reflection-in-action). More recently, others have highlighted the importance of reflection as part of long-lasting educational change 'in their classrooms', in order to 'understand, experience, and reflect on innovative methods' (Burke, 2013, p. 248). Teacher reflection of this sort is often linked to the action research cycle in which reflective practitioners pose questions of what went well and what did not, and what useful changes can be made to improve practice (and outcomes) over time.

However, despite the widespread opinion that reflection is something worth promoting and supporting, its implementation in TPD encounters multiple practical challenges (Gelfuso & Dennis, 2014), like the influence of cognitive biases on what we remember and choose to reflect upon (cf. Tversky & Kahneman, 1973) – often leading to reflections which are insufficiently based on evidence (Marcos, Sanchez, & Tillema, 2011). Other practical challenges of supporting teacher reflection are related with the tools and technological scaffolds provided for it, e.g., the fact that keeping journals, recording videos of lessons, etc. is still cumbersome and intrusive in everyday classroom practice (Banville & Rikard, 2001). Below, we briefly review such technological support for reflection, and its main outstanding challenges.

1.2. Technological support for teacher reflection

Aside from pen-and-paper journals and peer observation notes, which have been traditionally used to support teacher reflection, many other technological tools have been proposed to support reflective teaching practice (Romano & Schwartz, 2005). There exist proposals to support written reflection using digital journals (e.g., Lindroth, 2015), as well as e-portfolios (e.g., Winberg & Pallitt, 2016), with a variety of purposes, from teacher assessment (Lambe, McNair, & Smith, 2013) to the creation of professional learning communities (e.g., shared through social media, Kennedy, 2016).

Another set of proposals focuses on the use of online discussions to foster reflection while collaborating with peers. These can be supported, for instance, through weekly e-mails (Cook-Sather, 2007) or blogs (e.g., Hramiak, Boulton, & Irwin, 2009), and even social micro-blogging (e.g., Twitter use in Mieliwocki, 2014).

A large body of work has focused on the recording and analysis of video lessons to support reflection (Es, Stockero, Sherin, Van Zoest, & Dyer, 2015; Santagata, Zannoni, & Stigler, 2007). Some of these proposals just require video recording equipment to be implemented, while in other cases video annotation and coding software is also used (e.g., Clarke, Chen, Bickel, Sherer, & Resnick, 2015; McFadden, Ellis, Anwar, & Roehrig, 2014). Such retrospective video analysis encourages reflection throughout the teaching cycle and helps to zoom in on particular situations to find patterns and relationships between teacher moves and learning (Mosley Wetzel, Maloch, &

Hoffman, 2017), and enables joint discussion/reflection with peers and supervisors (Melville, Bowen, & Passmore, 2011). More recently, such video-based approaches have been combined with other tools like video editing, journal writing or video-clubs (e.g., Bayat, 2010), and the use of wearables (Fleck & Fitzpatrick, 2009) and mobile devices to capture snippets of practice or reflections (Aubusson, Schuck, & Burden, 2009). Video has also been used in conjunction with conferencing tools to analyze teaching practice (Lopez, Ortiz, & Allen, 2015).

Despite this wealth of proposals and technologies to support teacher reflection, such technology use is still far from widespread. Problems and barriers often mentioned in the literature include general usability and computer literacy issues, but also others more specific to the proposed technology: issues of privacy and anxiety (e.g., children afraid to talk on camera, Reid et al., 2016), the need to obtain informed consent (Aubusson et al., 2009) or time constraints (Dreyer, 2015; Hamilton, 2012). Also, such technological proposals share problems common to all reflective approaches, like the need of showing the added value and relevancy of performing the reflection (Friedrich, Ostermeier, Diercks, Krebs, & Stadler, 2012), or the fact that, over time, such reflections feel repetitive (Hramiak et al., 2009). Additionally, this technological support also has other practical shortcomings: they require conspicuous action or breaking the flow of the lesson (e.g., having a human observer present, setting up recording equipment or stopping to record a video, etc.). These practical problems are not trivial to solve, nor should they be dismissed as unimportant (Dillenbourg, 2013), especially given classrooms' high immediacy and simultaneity (Dawson, 2006).

1.3. Research goal

From the research literature reviewed above, we can see that reflection-based teacher professional development is an important path for educational improvement, but also that it operates under very real practical constraints, often relying on too few big reflections too far apart from each other. Its technological support also faces similar hurdles, ranging from the prosaic (e.g., time constraints, usable software) to the fundamental (e.g., the human need for privacy or novelty). This suggests that a purely technological solution may be insufficient to tackle all of them, and that we need to study the socio-technical composite that results from adopting a new tool in a context as socially-charged as classrooms are.

Teacher change is unlikely to come about through one 'aha moment' in a one-shot TPD action, but rather through small, incremental change. Taking a hint from the reported lack of evidence being used as a base for reflection (Marcos et al., 2011), we set out to explore the potential of everyday, small incremental change through frequent evidence gathering and feedback in TPD, hoping to create a washback effect that directly impacts how teachers are thinking about what they are doing. Hence, rather than Schön's (largely time-based) distinction between reflection-in-action and reflection-on-action, we will try to aim at reflection-for-action: when reflection happens is not as important as how often it happens and whether teachers find the right response to their reflections (which seldom happens when reflection is not embedded in everyday practice).

In the rest of this paper, we report on the first three iterations of a design-based research (Wang & Hannafin, 2005) project, in which a total of 16 teachers participated. During these iterations, we collaborated with teachers and professional development specialists, using paper prototypes, wearable sensors and web technologies to gather

data from real practice for several weeks, multiple times a day. In the first iteration, we explored the use of wearable sensors and data visualizations as the evidence and objects of reflection, collaborating with one secondary school teacher. In the second iteration, teachers from another school used paper prototypes mimicking different technology form factors (e.g., a mobile app vs. a desktop application), every day over two weeks. Finally, another set of secondary school teachers from the same school used a digital prototype in the third iteration, leading to a more nuanced socio-technical approach that also involved students.

In the next section, we describe the methodology of our study, followed by a summary of the main evidence and conclusions of each design-based research iteration. Later, we propose a series of design guidelines for the technological support of reflection based on everyday evidence. We close the paper with an outline of the limitations and future implications of our work.

2. Methodology

Against the aforementioned research backdrop, we set out to investigate the following design-oriented research question: 'How can we design tools (and, possibly, practices) that support teacher reflection based on everyday evidence?' There exist inherent methodological challenges in studying reflection (by measuring reflection we normally change reflection, e.g., if we ask any question to a teacher, we actually trigger reflection). Hence, within this broad research question, our central issue was to foster the capture of data that sparks reflection, rather than measuring the reflection itself.

To study this research question we used design-based research (DBR, see Brown, 1992; Wang & Hannafin, 2005) as a methodological framework. We chose this methodology due to its emphasis on daily classroom practice, and the fitting of the designed artifacts within authentic contextual constraints. Furthermore, we also aimed at developing 'humble theories' of how to design socio-technical systems for the purpose of supporting data gathering for reflection (see the guidelines provided in section 6).

Below, we report on the first three iterations of our DBR process. Given the length limitations of this paper, only the main methods, pieces of evidence and conclusions are described in each iteration (rather than a full research report of each study). This is to enable the understanding of where the design guidelines (arguably, the main contribution of the paper) originate.

Throughout these iterations, our focus of inquiry and technological approach evolved, as we explored the design space and incorporated new findings: the first iteration focused on the use of wearable sensors to gather data and advanced visualizations to spark reflection; the second compared different technology form factors as well as data gathering practices of teacher self-observation vs. student-reported experience; finally, the third one evaluated the use of a web application in combination with teacher and student observations.

These iterations took place in collaboration with practitioners in two secondary schools in Switzerland. It is also worth re-iterating that our focus was not only the design of technology alone to support reflection; rather, we wanted to study the co-evolution of such technologies and their classroom usage practices (since tool usage routines are a critical aspect in the integration of new technologies in the classroom, see Prieto, Villagrá-Sobrino, Jorrín-Abellán, Martínez-Monés, & Dimitriadis, 2011).

3. Iteration 1: Exploring reflection on multimodal data with wearable sensors

As a first exploration of the design space, and following previous research work in which we used wearable sensors (mobile eye-trackers) to study teacher cognitive load (Prieto, Sharma, Kidzinski, & Dillenbourg, 2018), we focused our first iteration on studying what aspects of wearable and ubiquitous sensor data (and visualizations) are found useful or interesting by teachers. This approach, like those based on classroom video, have the advantage of being relatively unobtrusive and not breaking the flow of the lesson.

3.1. Context and Method

Given the expense and effort involved in deploying and setting up the sensors in a real classroom, we chose to develop this first iteration as a single-teacher qualitative case study, which took place along four weeks during the school year. The study was set in a private international school in Switzerland, with an experienced secondary-level Mathematics teacher, who had taken part in previous studies and was thus familiar with using the wearable sensors while teaching. During the study, the researchers guided the teacher through a process of teacher inquiry (Dawson, 2006):

(1) An initial teacher interview was conducted to elicit and to agree on what aspects of her practice the teacher found interesting, including her inquiry questions.
(2) Over four weeks, eight lessons were recorded (twice a day, once a week) using a mobile eye-tracker worn by the teacher. The teacher also wore a device that recorded indoor location and physical movement, (a mobile phone app in the teacher's pocket). After each session, additional questions regarding the lesson workload were asked (using the NASA-TLX instrument). The sessions were recorded with two different cohorts of students of the same age, following similar lesson design and content each day, to enable comparison between lesson repetitions (one of the interests voiced by the teacher).
(3) With the data recorded, the researchers manually coded the videos in terms of teaching activity and social plane of interaction with students (to build orchestration graphs, similar to Prieto, Sharma, Dillenbourg, & Rodríguez-Triana, 2016), as well as disciplinary events (both aspects of interest for the teacher). Then, taking a cue from Engeström, Virkkunen, Helle, Pihlaja, & Poikela (1996), a 'classroom mirror' website with different visualizations and access to the eye-tracking videos was built (see Fig.1), as the main artifact to spark reflection about the teacher's practice.
(4) A final teacher interview was conducted, in which the teacher explored the different visualizations and audiovisual materials in the classroom mirror, and expressed her reflections and insights stemming from the classroom evidence through a think-aloud protocol. Further questions about the approach, technology design and potential future usage were also answered by the teacher.

The evidence gathered throughout this process (mainly from the initial and final teacher interviews, but also researcher observations during the process) were analyzed through open coding and conventional content analysis (Hsieh & Shannon, 2005). The main results of this analysis are summarized below.
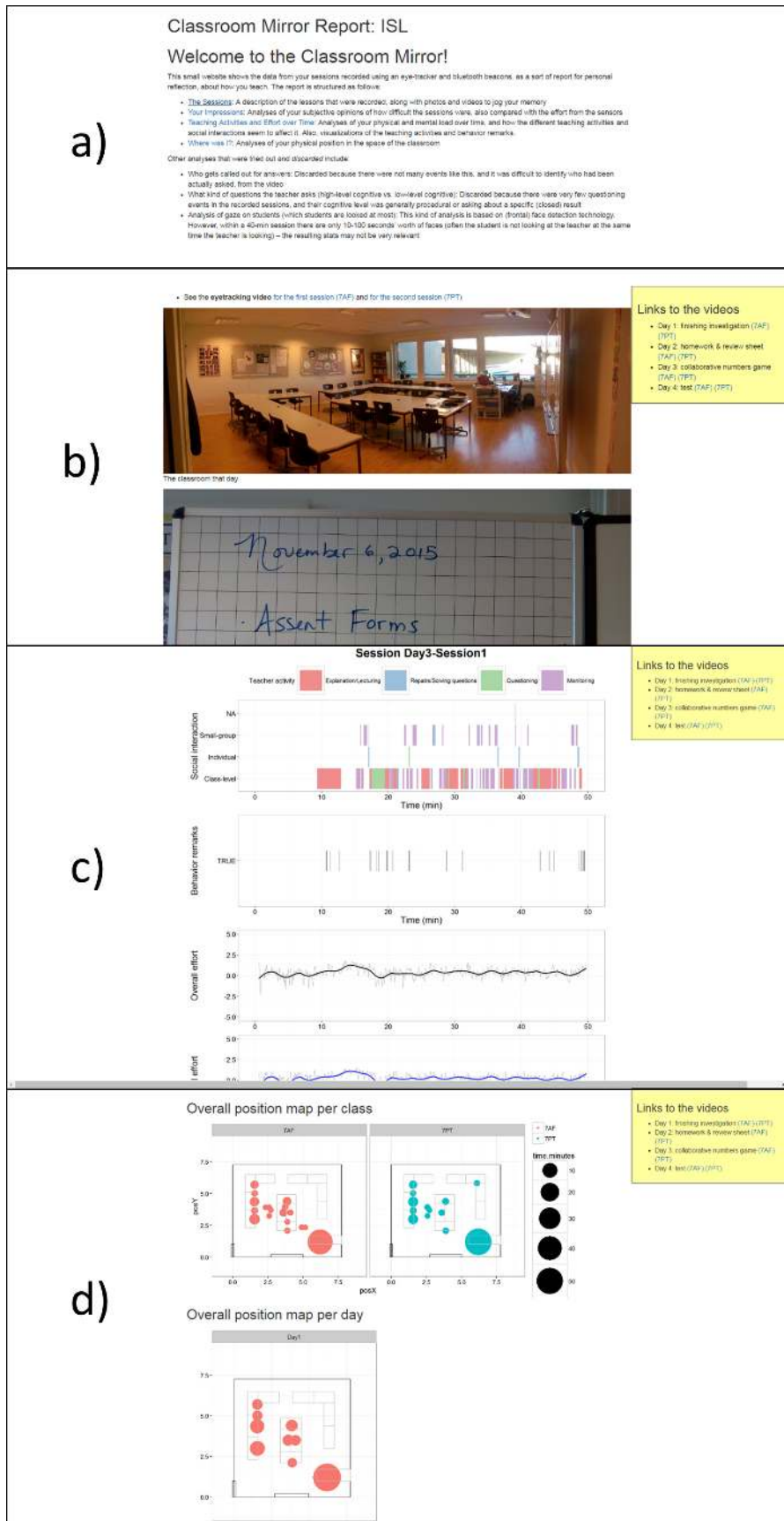
Figure 1. Screenshots from the 'classroom mirror' prototype used in Iteration 1. From top to bottom: a) Welcome screen; b) lesson's contextual description including photos; c) activity, time and effort visualizations; d) teacher spatial location information

3.2. Results

The teacher's reflections upon exploring the data visualizations in the final interview were quite varied. In many cases, the teacher remarked that the data confirmed certain suspicions she had about her own practice (e.g., the fact that she tended to spend most of her time in the front part of the classroom, hinting at a more traditional, explanation-oriented teaching style). Another such confirmed suspicion was that one of the cohorts of students was perceived as more 'difficult to handle' than the other one (hinted by the visualizations of the workload data). There were, however, aspects of the aggregated data visualizations that were surprising for the teacher. For instance, she pointed at a visualization showing that she spent more time walking around the left side of the classroom (see Fig.1, d) – even if she consciously made an effort to spread herself evenly across the room and among her students. A closer examination of the physical layout of the classroom during the interview led the teacher to the discovery that her desk's position inadvertently made it more difficult to access a part of the room, hence producing this lateral bias in the teacher's classroom presence.

Although the initial interests of the teacher regarding her own practice (as elicited from the initial interview) covered a wide range of issues, from physical classroom position to disciplinary events or questioning style, the interest upon actual viewing of the data varied greatly: very detailed temporal series (such as those depicting cognitive load over time, or the orchestration graphs of each lesson, see Fig.1, c) were considered less interesting and actionable than lesson-wide aggregations (e.g., percentage of time spent on each teaching activity, or at a certain physical position).

Finally, it is also worth noting that deciphering all the data visualizations available and reflecting upon them was a lengthy, time-consuming process (as an example, the interview took two hours, while it presented data from eight lessons spanning less than 10 hours of practice). This, and the fact that the reflection took place at the end of the school day, quickly evidenced that such a detailed approach to reflection on teaching practice data may not be feasible and sustainable within the time and energy constraints of the hectic life of a secondary school.

4. Iteration 2: Exploring reflection-in-action through paper prototypes

For the second iteration, now working with a different secondary school, we shifted the focus of our design-based efforts from passively-generated sensor data to a more active data gathering approach (i.e., teachers would have to explicitly take actions to gather classroom evidence about their practice). This was due to the fact that much of the sensor data was not found especially interesting by the teacher in the first iteration. This lack of interest also emerged in our conversations with teachers and the professional development specialist from the school setting of this second iteration. During these conversations, two main approaches for collecting evidence during everyday practice were selected as potentially interesting and feasible:

- a teacher-centered approach, in which the teacher actively makes an observation or recording when a relevant classroom situation emerges (similar to what Könings & Gijselaers, 2015 propose for reflective health professionals).
- a student-centered approach, in which students are asked, at the end of the lesson, about their learning experience and whether they had observed during the lesson some of the selected classroom behaviors of interest.
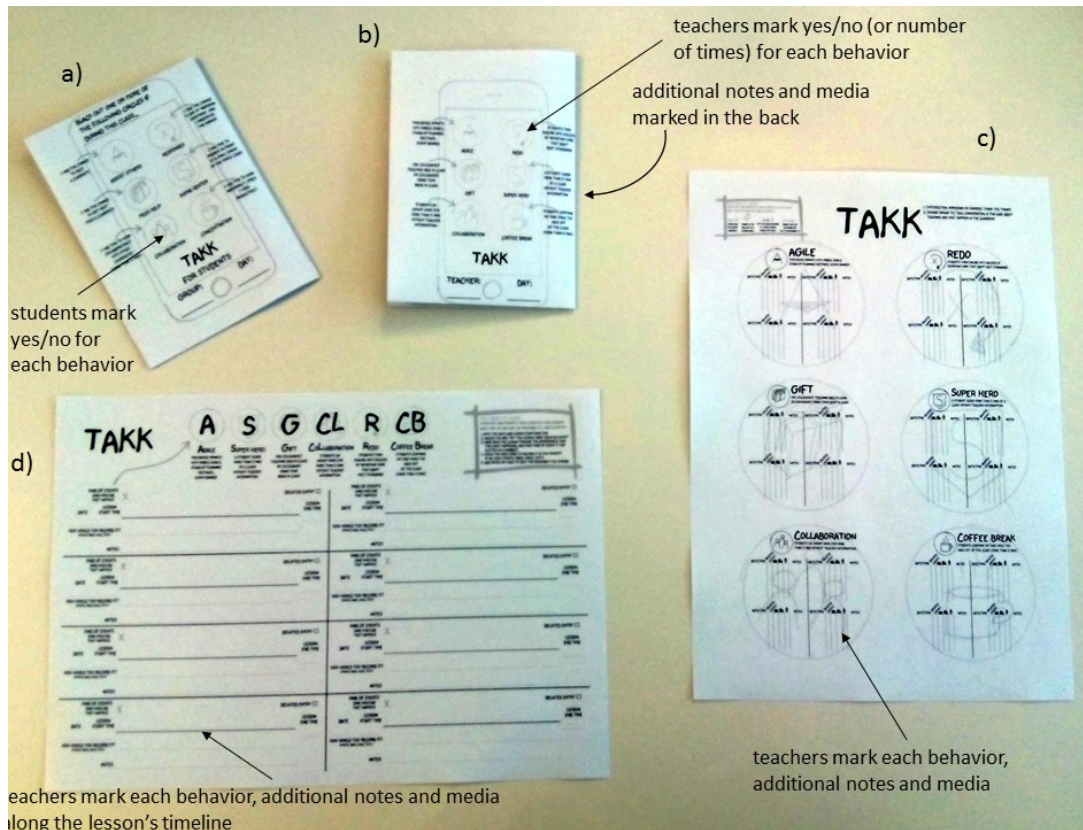
Figure 2. Paper prototypes used for data gathering in Iteration 2. From top-left, clockwise: a) student-oriented data gathering; b) teacher-oriented, mobile form factor prototype; c) teacher-oriented, wall-mounted form factor prototype (scaled down to A4 size); d) teacher-oriented desktop prototype.

In both approaches, the events and behaviors to be reported were agreed upon with the school TPD specialist, on the basis of classroom behaviors that the school wanted to promote (e.g., more stretches of collaborative work, or students presenting in front of the class). Hence, the research focus of this iteration was about what approach for gathering evidence (teacher-centered vs. student-centered) was most effective, and what form factor would be preferrable (e.g., as a mobile app, a desktop software, or a wall display). To compare and explore these different options in the field, paper prototypes were developed for teacher use over several weeks (see Fig.2).

4.1. Context and Method

In this second school, the in-service TPD approach consisted of a mix of training courses and personal teacher inquiry into their own practice (facilitated by the local TPD specialist). In his own words, the specialist described the goal of the technology-enhanced intervention as 'gathering non-threatening evidence to spark conversations about teaching practice'. Nine teachers from the school, with varying levels of teaching experience, volunteered for the study.

Given the small sample size, we chose a within-subjects research design, in which teachers used the prototypes of one approach (i.e., student-oriented or teacher-oriented) for one week, then switched to the other approach for the second week. The procedure followed was:

(1) An initial interview was conducted with each teacher to explain the experiment, obtain informed consent and provide the paper prototypes for the first week.

(2) The teacher used the paper prototype continuously during one week. The teachers were encouraged to use the prototype for every lesson in that week, if possible.

(3) A mid-experiment, semi-structured interview with the teacher, to learn of any problems and gather initial impressions, and to provide and explain the prototypes to be used the second week.

(4) The teacher used the prototypes for the other kind of approach, for another week. Again, use in every lesson was encouraged.

(5) A final semi-structured interview with the teacher, asking about teaching background, impressions and assessment of both kinds of prototypes, as well as opinions regarding the potential future use of a similar, digital prototype.

The events and annotations that teachers made in the paper prototypes were analyzed quantitatively using descriptive statistics. Furthermore, qualitative content analysis was performed on the interviews, to triangulate the quantitative results.

4.2. Results

In total, the nine teachers (n=9) did 132 annotations using the teacher-oriented approach, during a total of 75 different lessons. In turn, there were 551 annotations gathered using the student-oriented approach, from 57 different lessons.

4.2.1. Teacher-centered approach

As we can see in Fig.3, most annotations were made in the prototype that followed a 'desktop form factor' (an A4 paper sitting on the teacher's desk), and most of them were delayed (i.e., at the end of the lesson, not during the lesson itself). In the interviews, teachers reported difficulty in remembering to do annotations during the lesson. It is worth noting that, even if teachers were encouraged to note what kind of media they would have wanted to attach to the annotation (e.g., a video, audio or a text note), most annotations specified no media (i.e., they only specified the type of event/behavior that occurred).

4.2.2. Student-centered approach

The approach in which teachers, at the end of the lesson, asked students to mark what kinds of learning experiences they had had (e.g., 'I had a chance to have my questions answered'), was generally preferred by teachers. They reported that it was easier to remember, and that the routine of asking students to do something is already well-ingrained in any classroom. Several teachers asserted that they found both approaches (the student-oriented and the teacher-oriented one) interesting, providing complementary perspectives. Furthermore, in another interview the local TPD specialist considered that the student-oriented approach gave a clearer idea of the overall practices of the school (see the data shown in Fig.4).

5. Iteration 3: Exploring joint teacher-student data gathering

In the third iteration, we designed a socio-technical composite building upon the results of the previous iteration. We started from the design of the evidence-gathering
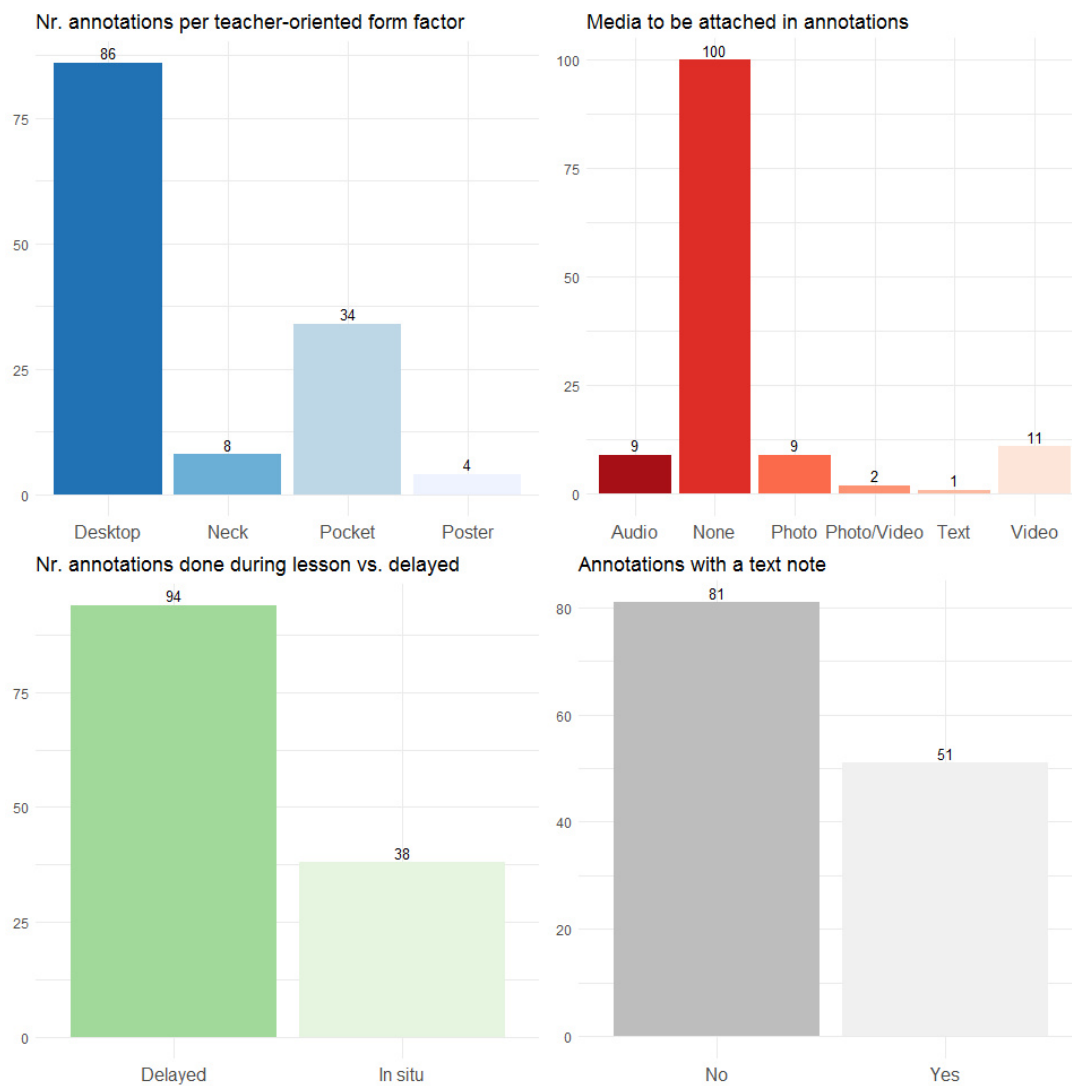
Figure 3. Characteristics of teacher-oriented approach events recorded. From top-left, clockwise: number of annotations per prototype form factor; kinds of media that teachers would attach to events; recorded events with a textual note; number of annotations done during the lesson (vs. delayed).
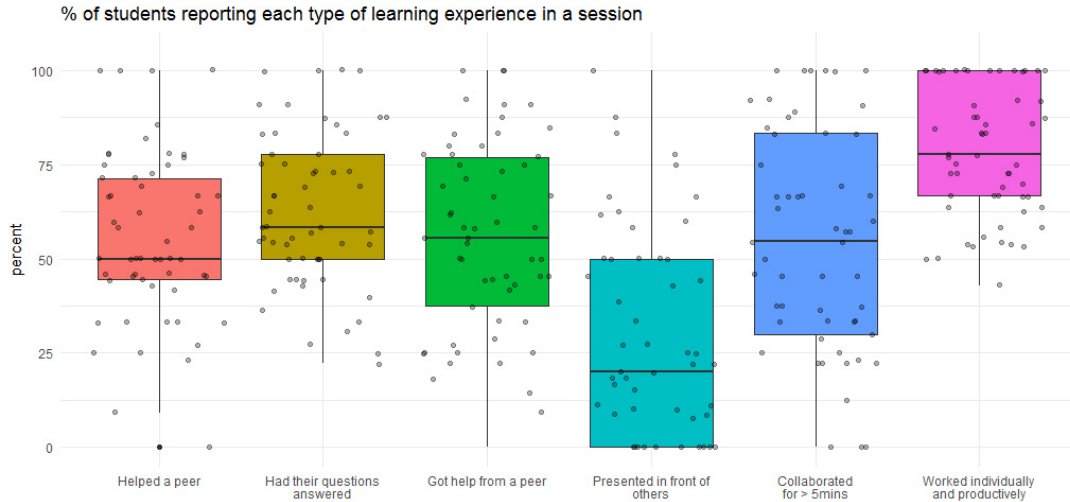
Figure 4. Summary of student-oriented data recorded using the paper prototypes in Iteration 2: each point represents one session, and the y-axis value denotes the proportion of students that reported having that experience, in a binary yes/no question).

practice: a combination of the teacher- and student-centered approaches of the previous iteration. In this data-gathering routine, students are asked at the end of the session about their learning experience, while the teacher reflects and tries to predict what the student responses will be (as a way to focus teachers' attention on the student experience, and the phenomena of interest that they desired to inquire about). To keep the data gathering routine within the strict time constraints of school life, student questions about their experience are limited to yes/no answers (e.g., "I have presented in front of the class today") or linear scales (e.g., "Today I understood this much of the lesson content"). In turn, the teachers are asked to predict what percentage of their students will answer a question positively (if yes/no), or the average score that they will provide as an answer (if linear scale). Teachers also would have the possibility of adding an additional short text reflection after seeing the results of the data gathered in that lesson (e.g., what classroom events that day may have led to such data).

To support this data gathering practice, we designed and implemented a minimalistic data gathering technology: a lightweight web platform that enables this data-gathering routine for both teachers and students (which we called Prolearning, see Fig.5). Hence, we set out to study this new design in the field, to investigate whether teachers are able to perform this routine in every lesson, every day, but also whether there is any evidence of changes in teacher practice (or student experience), and explore what factors an institution should consider when rolling out this kind of innovation.

5.1. Context and Method

We tested the aforementioned technology and practices in the same school as the previous iteration, in collaboration with the same TPD specialist and a new set of six teachers (n=6), with different amounts of teaching experience. The procedure to be followed by teachers was similar to the previous iteration:

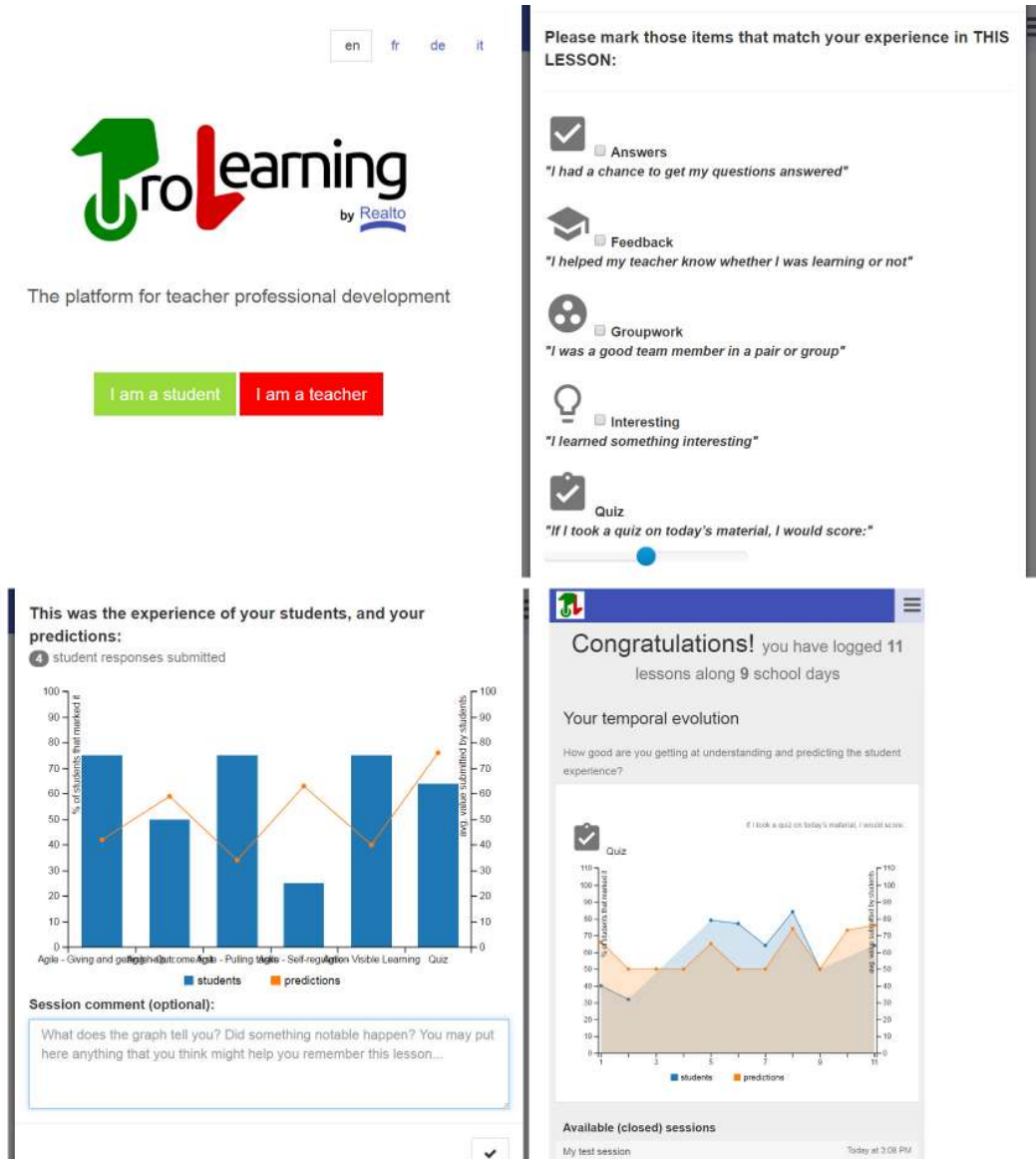(1) Initial interview with the teacher to explain briefly the study, obtain informed

Figure 5. Screenshots of the Prolearning prototype used in Iteration 3. From top-left, clockwise: Welcome screen; student-oriented questionnaire; teacher dashboard (incl. temporal evolution of an item's predictions and student reports); comparison of the student-reported values versus the teacher predictions at the end of each session.

consent and briefly explain the usage of the tool.

(2) The teacher used the tool in everyday practice, for two weeks. The teacher was encouraged to use the tool every day in every lesson, if possible.

(3) A final interview with the teacher, to gather impressions about the tool, perceived changes in teacher practice, and thoughts about intended future adoption and institutional roll-out.

In this case, the qualitative content analysis from the interviews was triangulated with quantitative analysis of the events recorded in the Prolearning tool, including the student responses about their learning experience (again, focused on behaviors that the school wanted to improve, as part of their TPD strategy), teachers' predictions of the student responses, and additional textual reflections.

5.2. Results

In the interviews, teachers reported having used the tool for an average of 70% of their lessons, every day. The main reasons reported for not using it in a lesson were 'being in the flow' of the school day, or particular clashes with a classroom's habits (e.g., a teacher that used to take student computers away at the beginning of the lesson). No teacher reported the lack of time as a reason for the un-recorded lessons. In this sense, teachers reported that the routine of evidence gathering took about 2 minutes (which matches roughly the median duration of a data gathering session from the tool's logs: 151 seconds).

Regarding the effects of using the tool on teaching practice and reflection, several teachers reported thinking more about the classroom behaviors and issues being asked about (e.g., "did I allow enough time today for their questions?"). Teachers reported a variety of student attitudes facing this new tool and classroom practice, often including questions about the potential consequences of this data for the teacher.

Regarding the data recorded by Prolearning (student responses and teacher predictions), teachers themselves showed an improvement in their ability to predict the student responses: a linear regression of the differences between student responses and teacher predictions over time ($n = 625$ predictions) showed that the day within the study was a significant, decreasing predictor of such difference ($\beta = -0.64$, in a 0-100 scale, $p = 0.0065$). It is also worth noting that teachers tended to predict student responses would be lower than they actually turned out to be (average difference: 10.45 in a 0-100 scale).

Reported student experiences in many cases also showed an improvement over time, although time was not found an statistically significant predictor in a linear regression model of the student experience response (see Fig.6 for a visualization of the trends for each of the student experiences inquired).

In general, teachers did not use the short reflection text field available after the teacher introduced the predictions and visualized the differences. Only 32 out of 125 sessions had notes, and most of them were introduced by a single teacher. Example reflections included comments about the prediction of student responses ("I underestimated the responses"), explanations for their scores ("today students had an assessment so no questions, no feedback, no group work"), or classroom events to make the session memorable ("Students worked in group on mindmap and I walked around to help them, guide them, answer their questions.").

Overall, teachers' reported experience of using the tool was good, although the description of the socio-technical composite (tool plus intended practice) varied widely:
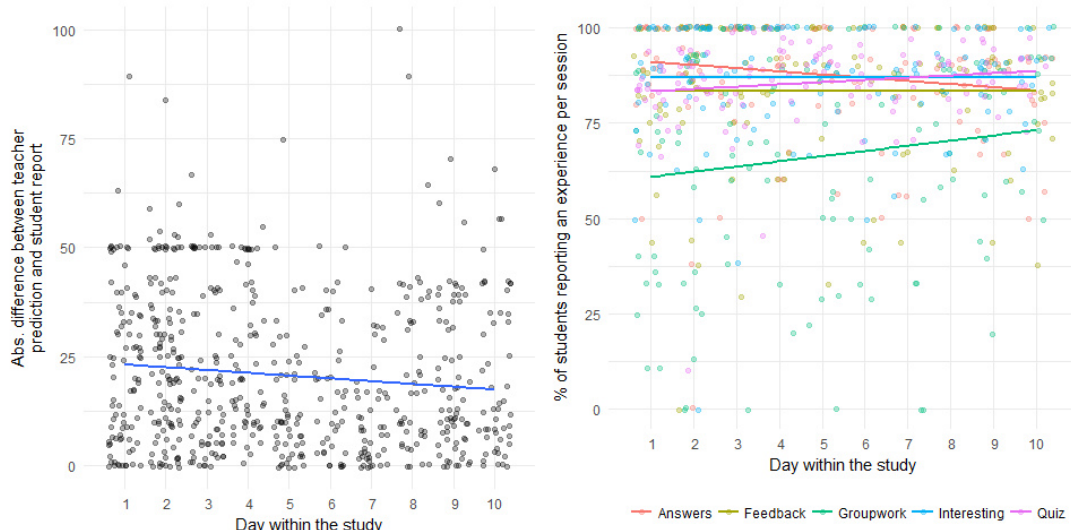
Figure 6. Temporal evolution of data gathered during everyday usage of Iteration 3's prototype. Absolute difference between student-reported values of student experience and teacher predictions (left); student-reported experiences per session (right).

from 'horrible' to 'painless' or 'I liked it'. It is worth pointing out that even the more negative opinions were ameliorated by the statement that, if it were part of a school-wide TPD effort, Prolearning would be preferable to most other TPD options they knew about.

Regarding the intended future usage, teachers stated that they would not use the tool for every lesson, every day, forever. Rather, there were many mentions about what would be the right frequency and duration of use in an eventual school roll-out. There was not consensus on this issue, but many teachers advocated using it as a sort of sampling device to 'check the pulse [of the practice] periodically'.

6. Discussion: Designing technology to support reflection-for-action

Our initial foray into designing a technology to help school teachers in reflecting upon data gathered from everyday practice used very detailed sensor data, and found certain interesting patterns (but was overly expensive to setup and time-consuming to reflect upon). In the second one, we used paper prototypes to look at more active observation by teachers or students, and found that remembering to observe in the middle of the lesson was difficult, but quite easy to do just at the end of the lesson. Finally, we developed a simple web tool (Prolearning) that queries both teachers and students, as a quick reflection wrap-up of the lesson's learning experience.

This concrete evolution, and the form that our final proposal took, are the result of our contextual design-based research inquiry, under a particular local school culture and constraints. Hence, they may not be directly applicable to other contexts. This contextuality of results, along with the limited sample of teachers (a total of 16 teachers) are one of the main limitations of the research presented here. Furthermore, in our iterative study we looked only indirectly at the ultimate goal of TPD: students' learning outcomes (Guskey, 2003). This was only done in the form of one of the Prolearning questions that asked students to self-assess their understanding. While our recorded

14

evidence showed a positive progression in this value during the course of Iteration 3, further research is required to relate the use of our TPD tool with such outcomes. Finally, it is worth noting that our interventions were rather limited in length (a few weeks, in each iteration), and we did not evaluate the overall effectiveness of a TPD program that uses Prolearning (given the large number of other factors that would affect such effectiveness).

Despite the contextuality of our design-based research process, the experience from our iterative study can already be distilled into guidelines and factors that technology designers should consider when building tools to promote everyday data gathering for teacher reflection (in similar contexts). These are the seeds of a 'humble theory' (Cobb, Confrey, DiSessa, Lehrer, & Schauble, 2003) that we will continue refining in the future:

(1) *Design the socio-technical pair (technology + routine) together*: given the crucial role of automaticity and routines in teacher practice (Feldon, 2007; Prieto et al., 2011), we should design technology that is not only easy and fast to use (e.g., the lack of a login screen for students in Prolearning), but also easy to automatize within their everyday practice. The 'envelope routine' of asking students for data at the close of the session is a clear example of that.

(2) *Design for overload*: teachers' lack of spare energy and attention during the school day manifested across all our studies. The technology (and the way it is used) should make it impossible to forget the data gathering and reflection moment (e.g., enlisting students' help to remind the teacher).

(3) *Do not forget the students*: contrary to many other teacher observation and reflection tools (for which students are largely a feature in the background), our final approach required active participation from students. This not only has advantages in ethical terms, as it gives them voice in the TPD process; it also re-centers the whole evidence gathering endeavor around the stakeholder which reflection should most impact: students.

(4) *Data: Space and Time, Activity and Experience*: among the endless variety of data that observers and automated sensors can register, aggregate measures of time spent in activities or places, as well as teaching actions and student experiences, were judged most interesting, easy to interpret and actionable by participant teachers.

(5) *Attention to ownership*: It quickly emerged in our studies that it was necessary to let schools and/or individual teachers personalize the items/behaviors to observe and reflect upon (e.g., certain teachers asked to add subject-specific questions). It also emerged quite clearly that extreme care has to be taken about what data is gathered and for what purpose, whether stakeholders are comfortable with that, or who owns and can see the data. This is especially critical for school leaders trying to roll out this kind of TPD innovation. For instance, the local culture in our context led to anonymous student data and personal ownership of the gathered data by each teacher (not by the TPD coordinator or the school administration).

It is worth noting that these emergent principles and the technologies we designed are aligned with several already-known guidelines for effective classroom technology design, or 'design for orchestration' (e.g., the minimalism of the user interface, see Dillenbourg, 2013). They also match effective TPD guidelines, e.g., the fact that it should be collaborative (Darling-Hammond et al., 2009). However, they may do so

in unexpected ways: instead of teachers collaborating in professional development, we came to have students collaborating with the teacher to gather useful evidence for TPD.

A full comparison of the approach and proposals made here with the many forms of reflection support mentioned in our related work, exceeds the length limitations of this paper. Maybe one of the most related recent applications for professional development and reflection was the mobile app for healthcare interns described by Könings & Gijselaers (2015) (which closely resembles our teacher-oriented observation approach tested in Iteration 2). While they opted for the 'active capture' approach, in our case such capture proved ineffective (possibly due to the contextual differences of classroom and hospital work). In both cases there is a need for further studies to compare the kinds and amounts of reflection that these innovations prompt, versus other approaches like, e.g., journalling.

We can also compare Prolearning with more established classroom observation schemes and tools like iObservation (DuFour & Marzano, 2011). Such tools are also designed to give feedback to be used for reflection, and include multiple well-established and detailed observation protocols. Prolearning, in turn, provides a minimal feature set, and includes students in the data gathering, taking less than 3 minutes of everyone's time (and does not require additional staff to be in the classroom doing the observation). With Prolearning, we do not intend to substitute such complex observation tools; rather, we think both can be adopted complementarily, one for longer stretches of 'taking the pulse' of a classroom (Prolearning), and the others for deeper, punctual dives into a classroom's practices.

7. Conclusion and future work

After three iterations of design-based research, our proposal for a technology to help teachers reflect about their practice in an everyday manner has evolved, from high-tech and data-intensive to a more nuanced socio-technical and student-centered approach. The resulting Prolearning tool is now structured in such a way that teachers know what the students will be reacting to, enabling the teacher to shape their lesson right away. The tool's philosophy of frequent, small practice improvement iterations puts just enough "pressure" on teachers to instruct in a certain way (according to the choices they had made in their TPD meetings). This kind of frequent feedback produced collaboratively with students is distinct from other supports for reflection, and we believe it can be more engaging than traditional observation as a clinical act of data gathering that is done *to* teachers, and only later brought as results to them. It also reflects a different approach to the design of TPD support technologies, that pays special attention to the orchestration constraints and hectic pace of many school classrooms (Dillenbourg, 2013).

Since the time of the studies depicted here, we have continued developing the Prolearning tool, implementing school-level and individual teacher personalizations of the observation items. The tool itself (or other tools based on the same principles) have now been tested, not only by the same Swiss school, but also by other schools in Germany and Estonia.

Our main direction for future research work entails the combination of passive, privacy-friendly sensor data gathering, with the active teacher-plus-student observation data gathering used in Prolearning. These means, along with other assessments of learning outcomes will enable, not only effective reflection for TPD, but also to track

the changes in teaching practice and how it impacts students' learning at different levels (from subjective experience to learning outcomes).

Compliance with ethical standards

The research presented here was approved by EPFL's Human Research Ethics Committee (file no. 001_08/04/2013), including the procedures for participants to provide informed consent, and the reporting of data in anonymized form (indeed, the data in the Prolearning tool itself is anonymous by default).

Disclosure statement

The work reported in this paper did not imply any conflict of interest.

Biographical notes

Luis P. Prieto is a Senior Research Fellow at the School of Educational Sciences in Tallinn University (Estonia). His research interests include learning analytics, especially multimodal learning and teaching analytics, the study of teacher orchestration, and their application for teacher professional development.

Paul Magnuson is Director of Research and Curriculum at the Leysin American School in Leysin, Switzerland. His research interests include student self-regulated learning and the types of professional development for teachers that focus on student learning first, teaching second.

Pierre Dillenbourg is currently Full Professor in learning technologies in the School of Computer and Communication Sciences at the École Polytechnique Fédérale de Lausanne (Switzerland), where he is the head of the CHILI (Computer-Human Interaction for Learning and Instruction) Lab. He is also Academic Director at the Center for Digital Education in EPFL.

Merike Saar is a Ph.D. candidate at the School of Digital Technologies in Tallinn University (Estonia). Her research interests include the technological support to teacher professional development, especially using multimodal teaching and learning analytics.

References

Aubusson, P., Schuck, S., & Burden, K. (2009). Mobile learning for teacher professional learning: Benefits, obstacles and issues. *ALT-J*, *17*(3), 233-247.

Banville, D., & Rikard, L. (2001). Observational tools for teacher reflection. *Journal of Physical Education, Recreation and Dance*, *72*(4), 46-49.

Bayat, M. (2010). Use of dialogue journals and video-recording in early childhood teacher education. *Journal of Early Childhood Teacher Education*, *31*(2), 159-172.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, *2*(2), 141-178.

Burke, B. M. (2013). Experiential professional development: A model for meaningful and long-lasting change in classrooms. *Journal of Experiential Education*, *36*(3), 247-263.

Clarke, S., Chen, G., Bickel, D., Sherer, J., & Resnick, L. (2015). Through the looking glass: Using a classroom discourse visualizer to support teacher reflection on practice. In *International Conference on Computer Supported Collaborative Learning (CSCL 2015)*.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9-13.

Cook-Sather, A. (2007). Direct links: Using e-mail to connect preservice teachers, experienced teachers, and high school students within an undergraduate teacher preparation program. *Journal of Technology and Teacher Education*, *15*(1), 11.

Darling-Hammond, L., & Richardson, N. (2009). Teacher learning: What matters? *Educational Leadership*, *66*(5), 46-53.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the us and abroad* (Tech. Rep.). National Staff Development Council. (Technical report)

Dawson, K. (2006). Teacher inquiry: A vehicle to merge prospective teachers' experience and reflection during curriculum-based, technology-enhanced field experiences. *Journal of Research on Technology in Education*, *38*(3), 265-292.

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers and Education*, *69*, 485-492.

Dreyer, L. M. (2015). Reflective journaling: A tool for teacher professional development. *Africa Education Review*, *12*(2), 331-344.

DuFour, R., & Marzano, R. J. (2011). *Leaders of learning: How district, school, and classroom leaders improve student achievement.* Solution Tree.

Engeström, Y., Virkkunen, J., Helle, M., Pihlaja, J., & Poikela, R. (1996). The change laboratory as a tool for transforming work. *Lifelong Learning in Europe*, *1*(2), 10-17.

Es, E. A. v., Stockero, S. L., Sherin, M. G., Van Zoest, L. R., & Dyer, E. (2015). Making the most of teacher self-captured video. *Mathematics Teacher Educator*, *4*(1), 6-19.

Feldon, D. F. (2007). Cognitive load and classroom teaching: The double-edged sword of automaticity. *Educational Psychologist*, *42*(3), 123-137.

Fleck, R., & Fitzpatrick, G. (2009). Teachers' and tutors' social reflection around sensecam images. *International Journal of Human-Computer Studies*, *67*(12), 1024-1036.

Friedrich, A., Ostermeier, C., Diercks, U., Krebs, I., & Stadler, M. (2012). The team portfolio: A support and evaluation tool? findings from a teacher professional development programme in germany. *Professional Development in Education*, *38*(3), 377-394.

Gelfuso, A., & Dennis, D. V. (2014). Getting reflection off the page: The challenges of developing support structures for pre-service teacher reflection. *Teaching and Teacher Education*, *38*, 1-11.

Glatthorn, A. (1995). Teacher development. In L. Anderson (Ed.), *International encyclopedia of teaching and teacher education (second edition).* London: Pergamon Press.

Guskey, T. R. (1994). Professional development in education: In search of the optimal mix. In *Paper presented at the 1994 annual meeting of the American Educational Research Association.*

Guskey, T. R. (2003). What makes professional development effective? *Phi Delta Kappan*, *84*(10), 748-750.

Guskey, T. R., & Yoon, K. S. (2009). What works in professional development? *Phi Delta Kappan*, *90*(7), 495-500.

Hamilton, E. R. (2012). Video as a metaphorical eye: Images of positionality, pedagogy, and practice. *College Teaching*, *60*(1), 10-16.

Hramiak, A., Boulton, H., & Irwin, B. (2009). Trainee teachers' use of blogs as private reflections for professional development. *Learning, Media and Technology*, *34*(3), 259-269.

Hsieh, H.-f., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288.

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945-980.

Könings, K. D., & Gijselaers, W. H. (2015). Bringing learning to the workplace: A smartphone app for reflection and increased authenticity of learning. In *Transformative perspectives and processes in higher education* (p. 117-135). Springer.

Lambe, J., McNair, V., & Smith, R. (2013). Special educational needs, e-learning and the reflective e-portfolio: Implications for developing and assessing competence in pre-service education. *Journal of Education for Teaching*, *39*(2), 181-196.

Lindroth, J. T. (2015). Reflective journals: A review of the literature. *Update: Applications of Research in Music Education*, *34*(1), 66-72.

Lopez, M. d. C. B., Ortiz, M. E., & Allen, C. (2015). Using mobile devices and the adobe connect web conferencing tool in the assessment of efl student teacher performance. In *Critical call–proceedings of the 2015 eurocall conference* (p. 77).

Marcos, J. M., Sanchez, E., & Tillema, H. H. (2011). Promoting teacher reflection: What is said to be done. *Journal of Education for Teaching*, *37*(1), 21-36.

McFadden, J., Ellis, J., Anwar, T., & Roehrig, G. (2014). Beginning science teachers' use of a digital video annotation tool to promote reflective practices. *Journal of Science Education and Technology*, *23*(3), 458-470.

Melville, W., Bowen, G. M., & Passmore, G. (2011). Pre-service teacher reflections, video-conference and webct: An exploratory case study. *Electronic Journal of Research in Educational Psychology*, *9*(2), 799-822.

Mieliwocki, R. (2014). Morethanedchats: Did you know you can improve as an educator in only a few minutes a day? *Educational Horizons*, *93*(2), 21-23.

Mosley Wetzel, M., Maloch, B., & Hoffman, J. V. (2017). Retrospective video analysis: A reflective tool for teachers and teacher educators. *The Reading Teacher*, *70*(5), 533-542.

Prieto, L. P., Sharma, K., Dillenbourg, P., & Rodríguez-Triana, M. J. (2016). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. In *Proceedings of the sixth international conference on learning analytics and knowledge* (p. 148-157). ACM.

Prieto, L. P., Sharma, K., Kidzinski, Ł., & Dillenbourg, P. (2018). Orchestration load indicators and patterns: In-the-wild studies using mobile eye-tracking. *IEEE Transactions on Learning Technologies*, *11*(2), 216–229.

Prieto, L. P., Villagrá-Sobrino, S., Jorrín-Abellán, I., Martínez-Monés, A., & Dimitriadis, Y. (2011). Recurrent routines: Analyzing and supporting orchestration in technology-enhanced primary classrooms. *Computers and Education*, *57*(1), 1214-1227. Retrieved from https://doi.org/10.1016/j.compedu.2011.01.001

Reid, D. A., Simmt, E., Savard, A., Suurtamm, C., Manuel, D., Lin, T. W. J., & Knipping, C. (2016). Observing observers: Using video to prompt and record reflections on teachers' pedagogies in four regions of canada. *Research in Comparative and International Education*, *10*(3), 367-382.

Romano, M., & Schwartz, J. (2005). Exploring technology as a tool for eliciting and encouraging beginning teacher reflection. *Contemporary Issues in Technology and Teacher Education*, *5*(2), 149-168.

Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service

teacher education: An empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education*, *10*(2), 123-140.

Schon, D. A. (1989). Quotations. a symposium on schon's concept of reflective practice: Critiques, commentaries, illustrations. *Journal of Curriculum and Supervision*, *5*(1), 6-9.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Villegas-Reimers, E. a. (2003). *Teacher professional development: An international review of the literature* (Tech. Rep.). International Institute for Educational Planning Paris. UNESCO.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53*(4), 5-23.

Winberg, C., & Pallitt, N. (2016). "I am trying to practice good teaching": Reconceptualizing eportfolios for professional development in vocational higher education. *British Journal of Educational Technology*, *47*(3), 543-553.