“Keep Your Eyes on ’em all!”: A Mobile Eye-Tracking Analysis of Teachers’ Sensitivity to Students

Philippe Dessus, Olivier Cosnefroy, Vanda Luengo

Dessus, P., Cosnefroy, O., & Luengo, V. (2016). "Keep your eyes on 'em all!": A mobile eye-tracking analysis of teachers' sensitivity to students. In K. Verbert, M. Sharples, & T. Klobučar (Eds.), Adaptive and adaptable learning. Proc. 11th European Conf. on Technology Enhanced Learning (EC-TEL 2016)(pp. 72–84). New York: Springer, LNCS 9891.

"Keep your Eyes on 'em all!": A Mobile Eye-Tracking Analysis of Teachers' Sensitivity to Students

Philippe Dessus[1], Olivier Cosnefroy[1,2], and Vanda Luengo[3]

[1] Univ. Grenoble Alpes, LSE (EA 602), F-38000 Grenoble, France
[2] DEPP, French Ministry of Education, Paris, France
[3] Sorbonne Universités, UPMC Univ Paris 06, LIP6 (UMR CNRS 7606), Paris, France

philippe.dessus@univ-grenoble-alpes.fr,
olivier.cosnefroy@education.gouv.fr, vanda.luengo@lip6.fr

Abstract. This study aims at investigating which cues teachers detect and process from their students during instruction. This information capturing process depends on teachers' sensitivity, or awareness, to students' needs, which has been recognized as crucial for classroom management. We recorded the gaze behaviors of two pre-service teachers and two experienced teachers during a whole math lesson in primary classrooms. Thanks to a simple Learning Analytics interface, the data analysis reports, firstly, which were the most often tracked students, in relation with their classroom behavior and performance; secondly, which relationships exist between teachers' attentional frequency distribution and lability, and the overall classroom climate they promote, measured by the Classroom Assessment Scoring System. Results show that participants' gaze patterns are mainly related to their experience. Learning Analytics use cases are eventually presented, enabling researchers or teacher trainers to further explore the eye-tracking data.

Key-Words: Mobile eye-tracking, Learning analytics, Classroom supervision, Teacher information taking, Classroom observation system, Visualization techniques

1 Introduction

Maintaining some of the main variables of the classroom in adequate limits is one of the most crucial goals of every teacher, this activity being performed by continuous visual information takes. Teacher's situational awareness [1] is an important skill and is needed for supervising (i.e., taking information from the classroom environment) and controlling (i.e., acting on this environment in turn) the diverse events occurring in the classroom, often at fast pace. This skill has been shown to be directly related to learners' achievement [2].

Teachers' attentional resources are limited, so they cannot equally draw their attention to every event occurring during instruction, or on every learner. Two main concepts from the educational sciences literature have been derived from this assumption. Firstly, the concept of "*withitness*" [3], which refers to the ability of teachers to proactively manage disruptive events, letting their students imagine that "they have eyes in the back of their heads". Secondly, the concept of *steering group* [4], which refers to a group of learners more or less consciously selected, and frequently supervised by the teacher in order to take on-the-fly instructional decisions. It is worth noting that the two concepts are hardly compatible with each other: a "withitness-able" teacher takes the classroom as a whole, whereas a "steering group"-focused teacher selects and a priori targets a small subset of students.

Some concerns have been raised on these two concepts. The operationalization of the "withitness" [5], as well as its empirical support [6], have been subject to difficulties. Whereas the very existence of a steering group is hardly debatable, the literature on this concept does not agree with the main features of this group. For instance, Lundgren [4] argued that the steering group is composed of students between the 10^{th} and the 25^{th} percentile of their cognitive abilities. Wanlin [7] reported two kinds of steering groups, comprehension-centered and behavior-centered, and showed that teachers mostly focus on medium and highly proficient students. Since these scholars did not have the same observation tool, we assume that a finer-grained observation tool may shed some light on the actual features of the steering group.

The main goal of this paper is to bring empirical support about the existence of either of these two concepts. We used a mobile eye-tracker to determine the continuous teacher's eye-fixation behavior during a whole lesson, accounting for their selective visual attention. We related this information to the cues (both behavioral and related to students' achievement) that lead a teacher to focus his or her interest on a given student. Novice and experienced teachers participated to this study in order to seek likely differences of behavior. Eventually, thanks to a Learning Analytics (LA) system, we will argue that we can unveil teacher–students interactional patterns during instruction, which in turn would be useful in some real-life contexts (use cases).

2 Eye-Tracking Devices and Teacher Decision Making

A well-established fact is that every teacher has to keep an overall awareness on the instructional situation [3]. However, the kind of cognitive processes undertaken to maintain this awareness has been studied so far mostly from verbalization procedures (either current or posterior to the activity), which are known to offer an incomplete access to the action and decision cognitive processes [8], because of their partly implicit nature.

Eye-tracking devices have become a reliable way to overcome this problem [9]. They enable the capture of eye fixations and saccades so that two pieces of information can be inferred [10]: which kind of information is extracted from a scene (static or dynamic); how much a scene is complex (the more complex a scene, the longer are the eye fixations). Moreover, the amount of gathered direct information is far larger than with other ways to observe teachers' behaviors, and makes possible LA-based procedures. All in all, they allow the processing of a large amount of "low inference" measures, which can be seen as more objective than measures that rely on the interpretation of a scarcer set of information.

Eye-tracking devices have seldom been used in educational contexts, but they have mainly been used in very constrained environments, like text reading or information seeking on screens. However, a few recent researches used eye-tracking devices for dynamic classroom scenes, either for analyzing student's gaze [11], teacher's cognitive load [12], or the whole classroom [13]. So far, two studies have investigated teachers' selective visual attention through the use of eye-tracking devices.

The study from van den Bogert et al. [14] analyzed teacher's (20 novices and 20 experts) fixations when viewing two videotaped lessons on a TV screen equipped with an eye-tracker. The expert–novice (E–N) paradigm predicted that, firstly, the fixations would be longer and more variable for novices (i.e., more complex) and, secondly, the number of targeted students would be larger for experts than for novices. Three kinds of video segments were identified: "blank segments" (containing no event, as identified by neither novice nor experienced teachers), "low contrast segments" (containing events identified only by experienced teachers), "high contrast segments (with events identified by both). The results showed that novice teachers devoted more time in observing a disruptive student than experienced teachers did, the latter having a wider observation scheme. In low contrast segments, the experienced teachers exhibited shorter fixation times and a wider sampling across students than novices did. No differences were shown for the high contrast segments. There were no significant differences on the observation of blank segments between N and E teachers; no differences on the homogeneity of variance were found either. Since this study captures eye movements on a TV screen displaying a video footage, based on predetermined scenes, its proximity with authentic conditions is weak.

Cortina et al. [15] used a mobile eye-tracker to study the gaze behavior of 24 teachers (12 novices and 12 experienced). They analyzed the relationship between the quality of the classroom climate (using the CLASS, see below for more information), and the level of attention teachers devoted to each student of the classroom, computed with the Gini coefficient (ranging from 0: all students have the same number of fixations, to 1: only one student gets all the fixations). Results showed that the Gini coefficient of experienced teachers was significantly lower than this of novice teachers. Correlations between each CLASS dimension and the Gini coefficient were computed: quality of feedback score correlated positively and significantly with Gini scores ($r = .46$, $p < .05$), showing that the more teachers support learning in delivering feedback, the more their attention is equally drawn towards all the students.

These studies did not attempt to uncover steering groups, nor did they make any assumptions about the actual level of the students. We set up the following study to investigate these questions.

3 Research Questions

The main purpose of this research is to study the strategies of teachers' information gathering through a mobile eye-tracking device and in an ecological context. The use of such a device suits the highly dynamic nature of the classroom environment [16], where the diversity of the potential sources of change are difficult to capture with indirect observation tools. Our research questions are threefold:

- *Classroom awareness*: How can we characterize teachers' attention distribution among students? Is this attention related to some students' characteristics (like performance or behavior)? Does any "steering group" exist? If so, which are its features (number and level of students, number of groups)?
- *Relationship between classroom awareness and teacher–students interaction*: A teacher can be fully aware of what happens in his or her classroom without being reactive to any event. We thus have to check to what extent the teachers' awareness is related to the quality of his or her interactions with students. In other words, we sought to determine the relationship between teachers' visual cues in the classroom environment and their level and quality of the interactions they promote with students.
- *Learning Analytics-based visualization reuse*: Can the large dataset of this study, as well as its LA-related procedures, be spread to every researcher, or even teacher, who wants to investigate gaze teachers' behavior? Can we come up with some use cases of this database for teacher training or educational research purposes?

A novice–experienced comparison was undertaken for the first two research questions, supposing that more experienced teachers would be more aware of students' participation and achievement [17]. A specific LA-based procedure was undertaken to answer the third question.

4 Method

4.1 Participants

Four teachers (100% female) volunteered to participate to this study. Table 1 below shows teachers' main characteristics.

4.2 Measures

First of all, information about the students was gathered: age, gender, quartile level of performance in French and mathematics, special needs, and a 11-item questionnaire assessing the students' behavioral self-regulation abilities [18]. The following abilities were assessed: attention, tiredness, integration into the classroom, work speed, effectiveness, organizational capacity in performing a task, autonomy, and mastery of gestures. Teachers responded on a 4-point Likert scale, ranging from 1 for a behavior never noticed or not learned yet, to 4 for a behavior usually noticed or learned. A maximum likelihood exploratory factor analysis identified one factor as in a previous research [18]. The reliability was satisfactory ($\alpha = .77$). In order to estimate each pupil's behavioural self-regulation perceived by the teacher, each student was given a score taking into account the factor weight of each item.

Table 1. Basic information about teachers

ID	Grade	Nb Students	Experience (nb years)
1	1st	22	High (20)
2	3rd	24	Novice (0.5)
3	2nd	23	Novice (0.5)
4	1st	24	High (25)

Then, we had to represent the occurrence of pedagogical events throughout the teaching sessions. We adapted the Teaching Dimensions Observation Protocol (TDOP) [19], which is a reliable observation tool that captures a large variety of pedagogical practices and events. We used the TDOP to characterize the diversity of pedagogical events that occur in classrooms (e.g., the teacher gives an explanation then the students are doing a guided exercise). This information was coded independently by two researchers from the video footages, and disagreements were resolved by a discussion to reach a consensus.

Eventually, we assessed the level of the teacher–students interactions in the classrooms with the Classroom Assessment Scoring System (CLASS) [2], one of the most used and valid classroom observation systems. The quality of the interactions was assessed upon three main domains: emotional support, classroom organization, and instructional support, derived into ten dimensions (see Table 2 for more information). This judgment of quality of the teacher–students interactions is related to the observation of four 30-minute sessions, hence lasting a whole morning session for each observed teacher.

4.3 Procedure and Data Analysis

After a pre-experiment with a university teacher to rehearse the whole experimental scheme, we undertook a lesson recording in the four teachers' classrooms. The four participants taught a regular lesson of mathematics (numeracy or multiplication) lasting about 45 minutes in the morning, wearing a mobile eye-tracker (ASL Mobile Eye-GX). An additional video camera and an ambient microphone captured the whole classroom activity. Two trained observers gathered CLASS-related information during the whole morning class.

The set of lessons was then transcribed by two trained coders using ELAN [20]. The whole dataset was afterwards exported onto *UnderTracks* [21], a web-based Learning Analytics platform that enables the gathering, analysis, and sharing of a wide range of traces. *UnderTracks* is composed of a web platform to share traces and operators (processing algorithms in Python or R) and a client-side software (*Undertracks-Orange*) to build, share and reuse analysis processes (combination of operators and traces) thanks to the open source *Orange* data mining toolbox (http://orange.biolab.si). A given dataset, as well as its related analysis procedures, can thus be shared, reused, and modified by the *UnderTracks* researchers' community. Once shared, the processes can be applied to other traces. We created a specific data space in *UnderTracks*, called "SuperViseur", for storing raw data of this study, as well as displaying operators and processes used in its analysis. Raw data stored comprises gaze behaviors, students' characteristics, and pedagogical episodes.

The design and processing of the analyses of this study reused processes from within the *UnderTracks-Orange* client application. Fig. 1 shows an *Orange* process that builds several interactive visualizations displaying teachers' gaze behavior under several considerations (pedagogical episode, students characteristics). For example, one of the interactive visualizations shows, for each teacher, his or her gaze behavior by pedagogical episode. Its interactivity lies in the possibility to have more information about a given student when the mouse hovers each gaze target representation.

Any visualization can be saved onto the *UnderTracks* web platform; visualization results can be uploaded in other web sites as well. A dedicated website (http://superviseur.lip6.fr) proposes several visualizations to furthering the exploration of our data, likely following use cases (see Section 6). Moreover, any researcher can conveniently perform, upon registration, some of the analyses described in the paper, as well as new ones.

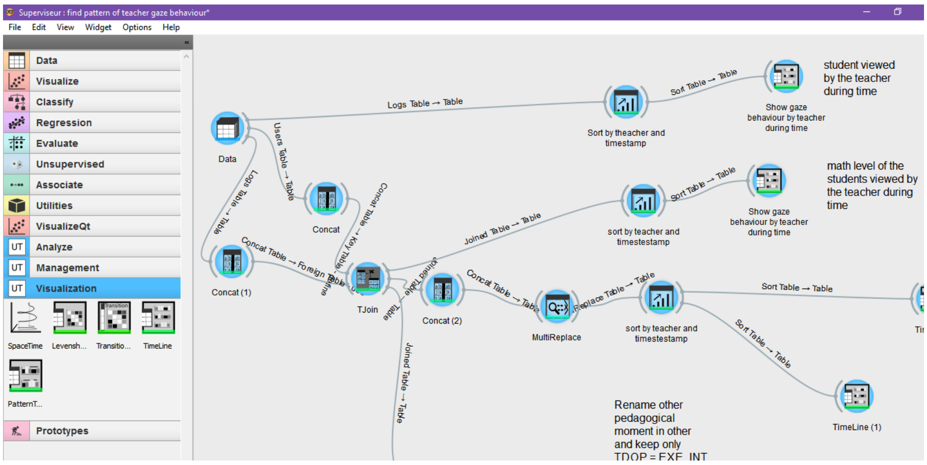


Fig. 1. An *Undertracks-Orange* process producing visualizations from teachers' gaze behaviors

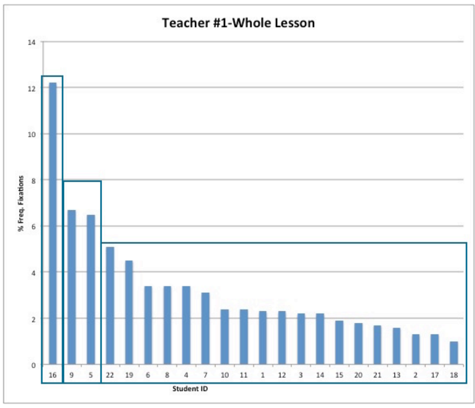
Specific attention was drawn to privacy concerns, and we sought agreements from our university data protection and ethical committees, mostly because the mobile eye-tracker necessarily captures the whole attentional stream of teachers, which makes it difficult to isolate students whose parents decided they would not be videotaped. The parents were given a description of the project and had to confirm their agreement.

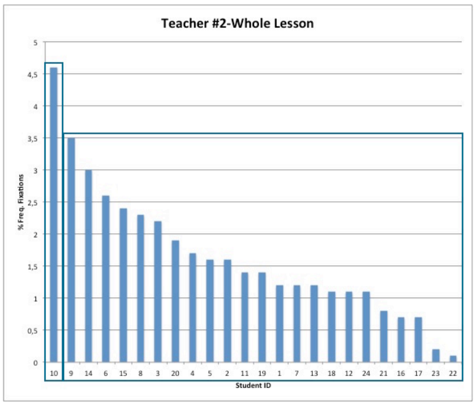
All of them agreed. The dataset available from within *UnderTracks* delivers fully anonymized data only, thus the video shots of the study are not viewable.

5 Results

5.1 Gazing Time: Whole Lesson Analysis

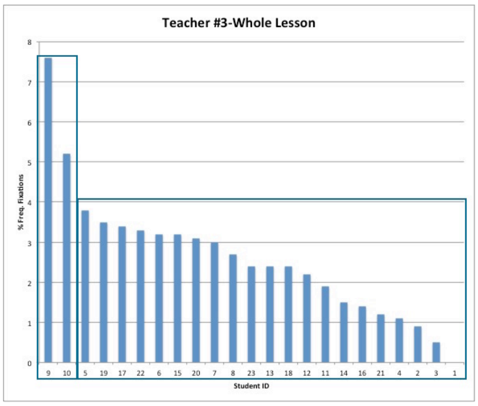
We first selected the same video time range (44 min) to control for time, corresponding to 5,280 eye fixations of 500 ms duration each, per video footage. We then extracted those targeting a student and computed the percentage of time a teacher is focusing a given student (see Fig. 2). Whereas every student was targeted during the whole lesson session, the distribution of the gazing time differed among teachers: the first three devoted most time (about 10%) on a reduced number of student while the fourth distributed her attention between students more evenly.

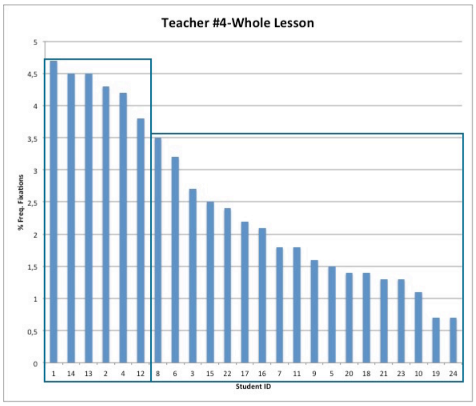




N=3,846, 22 tracked students

N=2,074, 24 tracked students





N=3,176, 23 tracked students

N=3129, 24 tracked students

Fig. 2. Distribution of the percentage of teachers' gazing time per student, sorted by descending rank, whole lesson (time range: 44 min)

We then tried to compose “steering groups” (rectangles in Fig. 2) in function of the attention distribution of the teachers, using the following rule of thumb: We empirically set the cut-off value of eye fixations as 200 fixations (100 s of gazing time) for determining groups. Then, we separated the distribution by tiers every 200 fixations. Results show that Teacher #1 focused her attention towards three distinct “steering groups”, a unique student (#16), a second group composed of two students (#9 and #5), and a third composed of the rest of the students. Teacher #4 exhibited a similar behavior, essentially focusing on a group of six students, the rest of the classroom being almost equally scrutinized. Teacher #2 and #3 focused more often their attention on a more reduced set of students (Student #10 for Teacher #3; Students #9 and #10 for Teacher #3), the others being far less attentionally sampled.

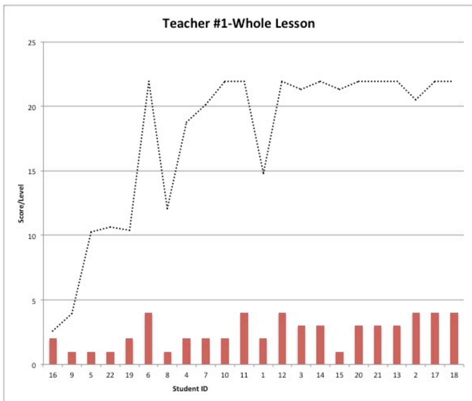
5.2 Teachers Gazing Time in Function of Pupils Behavior and Level

We then wondered if the teachers’ gazing time would depend on some salient characteristics of their students, like their behavior or their level in mathematics. Fig. 3 orders the students as in the previous figure (descending percentage of gazing time per student), in addition with categorical data about their level in mathematics (bars, the higher the better), as well as score data about their behavioral self-regulation (dotted lines, the higher the less dysfunctional). We expected that the more often a student is sampled over the lesson, the lower his or her level is (either in mathematics or related to his or her behavior).

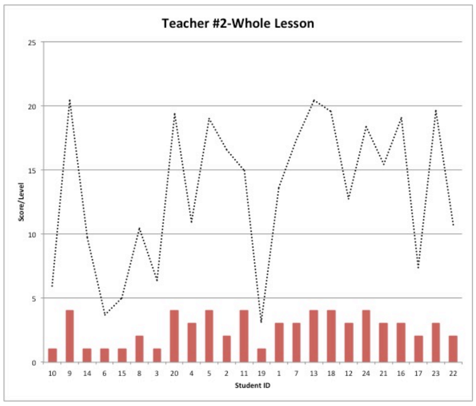
Fig. 3 depicts this relationship, showing that Teachers #1 and #4, again, had similar ocular behaviors: their students’ level curves are globally ascending, even if some irregularities occur (e.g., Students #8 and #1 were less observed by Teacher #1 than their behavioral level would let us think; likewise for Students #22, #5, and #21, Teacher #4). In comparison, the students’ level curves of Teachers #2 and #3 are not ascending and much more erratic, showing no relationship between the percentage of eye fixations and students’ level.

5.3 Analysis of a Specific Episode: Interactive Exercises

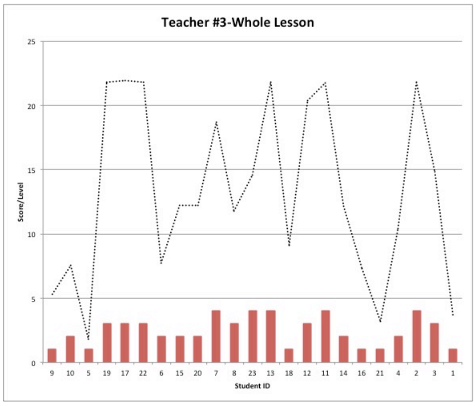
The above analyses made the assumption that teachers behave uniformly during the whole lesson in terms of information takes. In order to control for the kind of pedagogical event, we have now to analyze the participants’ gaze behavior on the same kind of event. We chose to focus on the *interactive exercise* derived from the TDOP taxonomy (a mix of “interactive lecture” and “deskwork”, frequently undertaken in French classrooms, enabling students to do exercises under teacher’s guidance), which lasted sufficiently long in each lesson, and necessitated a larger amount of information from students than others. All in all, the pattern of results related to these episodes was very similar to the overall results (see Section 5.1 and Table 2). For the sake of brevity, the interactive exercise-based results are available at http://superviseur.lip6.fr.



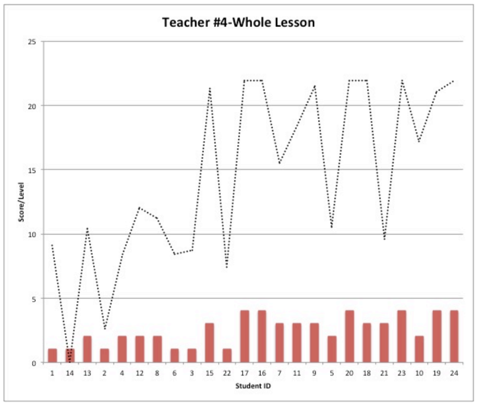
22 tracked students



24 tracked students



23 tracked students



24 tracked students

Fig. 3. Descending rank of the students in function of their gazing time, with information about their levels of behavior (dotted lines) and mathematics (bars) (time range: 44 min)

5.4 Relationship between the Attention Focus and the Classroom Climate

We computed Gini coefficients to measure teachers' attention distribution in interactive exercises [15], appropriate when the variable (in our case, attentional focus, or gazing time) is not independently distributed among students: if a given student is subject to focus, there leaves less chances of attentional focus to the others). The Gini coefficient ranges from 0 (all students get the same number of fixations) to 1 (one student gets all the fixations). Results show that Teacher #4 was the most "egalitarian".

Table 2 also shows the gazing time ratio between the amount of fixations towards less able students (in terms of behavior) and towards more able students, the cut-off between the two groups being the median. For instance, Teacher #1 had an overall behavioral gazing time ratio of about 2, meaning that she gathered two times more information from less able students than from more able students. The pattern of results regarding CLASS shows that, firstly, the smaller their behavior management CLASS-based scores are, the more teachers are "egalitarian", needing to scan a larger sample of students to manage their classroom. We obtained similar results with the gazing time ratio related to performance in mathematics.

5.5 Relationship between Attentional Lability and Classroom Climate

The previous sub-Section considered teachers' gazing time as a whole. However, two teachers may differently distribute their overall – and equivalent – amount of attention over time, one being focused on the same student for many contiguous saccades, the other being constantly changing his or her attention across students. We computed (see Table 2) the percentage of gaze changes, namely "attentional lability," for the whole lesson and for the Interactive Exercises episodes (100% stands for a change every saccade; 50% stands for a change every two saccades).

Table 2. Gini coefficients, behavioral- and performance-related gazing time ratio, CLASS scores, per teacher

	1	2	3	4
Gini Coeff. Overall	0.35	0.33	0.32	0.29
Behav. Gazing Time Ratio Overall	2.07	1.49	1.05	2.14
Overall Attentional Lability	36.9	53.8	49.8	36.8
Gini Coeff. Interactive Exercise	0.33	0.32	0.45	0.29
Behav. Gazing Time Ratio Int. Ex.	2.10	1.14	1.09	2.16
Int. Exercise Attentional Lability	37.6	65.2	48.1	42.2
CLASS Positive Climate	6.0	6.3	5.9	4.5
CLASS Negative Climate	1.0	1.1	1.1	1.7
CLASS Teacher Sensitivity	5.4	6.1	5.6	4.8
CLASS Regard for Student Persp.	4.5	5.6	5.3	4.2
CLASS Behavior Management	5.7	6.3	5.9	5.2
CLASS Productivity	5.7	5.9	5.9	5.2
CLASS Instr. Learning Formats	5.4	5.3	4.9	4.6
CLASS Concept Development	2.7	5.0	3.8	2.9
CLASS Quality of Feedback	4.5	4.9	4.3	3.4
CLASS Language Modeling	3.7	4.8	3.9	3.3

The results on attentional lability are the only ones that are both related to the experience difference between teachers (the more experienced have the lower percentages), and to their CLASS scores. The ranking order of the teachers' overall attentional lability is the same as their respective CLASS scores for Positive Climate, Negative Climate, Teacher Sensitivity, Regard for Student Perspectives, Conceptual Development, and Quality of Feedback. There thus might be a relationship between the teachers' attentional change over the students (his or her sensibility to students needs), and the quality of the teacher–students relationships (briefly put, the classroom climate); at the same time, experienced teachers exhibited lower attentional lability than novices did.

6 *UnderTracks* Use Cases

We can now sketch three uses cases, showing situations where researchers, teacher trainers, and even teachers, would take benefit from the analysis of eye-tracking data with *UnderTracks-Orange*. These use cases enable to foresee advances in the novel research domain of "Teaching Analytics" [22].

Use Case #1: Studying teachers' cognition from classroom management patterns. As argued in the first two Sections of this paper, there are numerous hypotheses on teacher cognition that would take advantage from being more objectively validated through LA-based eye-tracking data analyses. Researchers connected to large datasets of teachers' behaviors would uncover novel fine-grained classroom management patterns.

Use Case #2: Studying teachers' efficacy in relation with students learning. Evidence-based research has recently spread from medicine to educational research [23]. Given that perspective, researchers would use the kind of data we gathered, extended by students' indicators of performance. This would enable the study of the causality between raw behavioral indicators and learning.

Use Case #3: Uncovering behavior patterns for teacher training purposes. Teacher training sessions would also benefit from the device tested in this study. Pre-service teachers would be given access to videotaped lessons and their *UnderTracks*-based data; they would investigate some hypotheses about the teacher's awareness, his or her information takes, and their relationships with the students' behavior and performance, as well as with the classroom climate. Eventually, some instructional strategies would be derived from their conclusions.

7 Discussion

This paper considered the combined use of eye-tracking data together with Learning Analytics procedures leading to open and interactive visualizations of teachers' strategies. Our main results are summarized as follows. Firstly, every student of the four classrooms was looked at by his or her teacher, even a few times. This brings some support to the "withitness" hypothesis. Secondly, steering groups composition differed across teachers: very small groups of students were particularly subject to focus by the teachers, and thus can be considered as more complex in terms of decision-making. The size of the gazed groups seems to be related to the amount of experience of the teachers, as found in [14]. Thirdly, very little variability was observed across different kinds of pedagogical activity. Fourthly, the criterion for choosing a steering group is not clear-cut across teachers: again, teachers' amount of experience better predicted their steering groups-related behavior than the characteristics of their students, in terms of behavior or performance. Eventually, we found a small relationship between teachers' gazing time and the quality of the classroom climate, replicating Cortina et al.'s [15] results, as well as a more obvious relationship between the teachers' attentional lability and many of their CLASS scores.

During their activity, novice teachers engage a larger amount of cognitive load than more experienced do. The way the latter scan a larger "steering group" would make them able to perceive more fine-grained events [24], since they are less overwhelmed than novices are. This "steering group" is action-oriented, so it likely contains students whose behavioral changes may have effects on teachers' strategies (activity change, feedback, etc.) [25]. Novice–expert comparison studies in many fields (aviation, chess, sport, surgery) showed that experts, compared to novices, have fewer fixations of longer duration on nodal points of the situation [26, 27], while novices exhibit more variability. This line of results complies with our paradoxical result, at least at first sight, showing that an experienced teacher might be either egalitarian (i.e., with a smaller gazing time variability across students), and focused on a small set of specific students (i.e., with a restricted "steering group"). Focusing on this group of students allows expert teachers to make sound decisions, grounded on a representative set of students.

Further research will engage a larger set of participants, and consider the actual teachers' location in the classroom to test more ecological hypotheses, as well as finer-grained analyses of more complex episodes, like those involving teacher feedback. The implementation of some use cases in real-life contexts will be considered as well. They are paths to understand how teachers adapt themselves, with sensitivity, to their classroom environment and their students' needs.

Authors' note. This research was partly funded by the Pôle Grenoble Cognition, Univ. Grenoble Alpes, France. This research was approved by the CERNI (University's local ethical committee, n° 2013-09-24-25). We would like to thank the four teachers for having accepted to wear so weird a device and nevertheless doing good teaching; Brigitte Meillon for her invaluable help in calibrating, capturing, and post-producing the video footages; Michèle Arnoux and Mathieu Louvart for CLASS and videos' coding; Luc Sindirian and Pascal Bilau for making this research possible; and Andrea Doos for checking the English of a previous version of this paper.

8 References

1. Rittenbruch, M. and McEwan, G.: An historical reflection of awareness in collaboration. In: Markopoulos, P., Mackay, W. and de Ruyter, B. (eds.) *Awareness systems: Advances in theory, methodology and design*, pp. 3–48. Springer, London (2009)

2. Pianta, R.C., La Paro, K.M. and Hamre, B.K.: *Classroom assessment scoring system: Manual K-3*. Brookes, Baltimore (2008)
3. Kounin, J.S.: *Discipline and group management in classrooms*. Holt, Rinehart & Winston, New York (1970)
4. Lundgren, U.P.: *Frame factors and the teaching process*. Almqvist & Wiksell, Stockholm (1972)
5. Hastie, P.A., Sinelnikov, O.A., Brock, S.J., Sharpe, T.L., Eiler, K. and Mowling, C.: Kounin revisited: Tentative postulates for an expanded examination of classroom ecologies. *Journal of Teaching in Physical Education*, 26, 298–309 (2007)
6. Irving, O. and Martin, J.: Withitness: The confusing variable. *American Educational Research Journal*, 19(2), 313–319 (1982)
7. Wanlin, P.: Elèves forts ou faibles : qui donne le tempo ? Université de Liège, Liège (2011)
8. Ericsson, K.A. and Simon, H.A.: Verbal reports as data. *Psychological Review*, 87, 215–251 (1980)
9. Salvucci, D.D. and Anderson, J.R.: Automated eye-movement protocol analysis. *Human Computer Interaction*, 16, 39–86 (2001)
10. Mele, M.L. and Federici, S.: A psychotechnological review on eye-tracking systems: towards user experience. *Disability and Rehabilitation: Assistive Technology*, 7(4), 261–281 (2012)
11. Yang, F.-Y., Chang, C.-Y., Chien, W.-R., Chien, Y.-T. and Tseng, Y.-H.: Tracking learners' visual attention during a multimedia presentation in a real classroom. *Computers & Education*, 62, 208–220 (2013)
12. Prieto, L.P., Sharma, K., Wen, Y. and Dillenbourg, P.: The burden of facilitating collaboration: Towards estimation of teacher orchestration load using eye-tracking measures. In: *11th Int. Conf. on CSCL 2015*, Vol. 1, pp. 212–219. ISLS, Gothenburg (2015)
13. Wolff, C., Van den Bogert, N., Jarodzka, H. and Boshuizen, H.P.A.: Differences between experienced and student teachers' perceptions and interpretations of classroom management events. In: *Inter-university Center for Educational Sciences Fall School*, Girona, Spain (2012)
14. van den Bogert, N., van Bruggen, J., Kostons, D. and Jochems, W.M.G.: First steps into understanding teachers' visual perception of classroom events. *Teaching and Teacher Education*, 37, 208–216 (2014)
15. Cortina, K.S., Miller, K.F., McKenzie, R. and Epstein, A.: Where low and high inference data converge: Validation of CLASS assessment of mathematics instruction using mobile eye tracking with expert and novice teachers. *International Journal of Science and Mathematics Education*, 13(2), 389–403 (2015)
16. Doyle, W.: Classroom organization and management. In: Wittrock, M.C. (ed.) *Handbook of research on teaching*, pp. 392–431. McMillan, New York (1986)

17. Gettinger, M. and Kohler, K.M.: Process-outcome approaches to classroom management and effective teaching. In: Evertson, C.M. and Weinstein, C.S. (eds.) *Handbook of classroom management. Research, practice, and contemporary issues*, pp. 73–96. Erlbaum, Mahwah (2006)

18. Guimard, P., Cosnefroy, O. and Florin, A.: Evaluation des comportements et des compétences scolaires par les enseignants et prédiction des performances et des parcours à l'école élémentaire et au collège. *Orientation Scolaire et Professionnelle*, 36(2), 179–202 (2007)

19. Hora, M.T., Oleson, A. and Ferrare, J.J.: Teaching Dimensions Observation Protocol (TDOP) user's manual. Wisconsin Center for Education Reserch, University of Wisconsin-Madison, Madison (2013)

20. Sloetjes, H. and Wittenburg, P.: Annotation by category – ELAN and ISO DCR. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*. ELRA, Marrakech (2008)

21. Bouhineau, D., Lalle, S., Luengo, V., Mandran, N., Ortega, M. and Wajeman, C.: Share data treatment and analysis processes in Technology Enhanced Learning. *Data Analysis and Interpretation for Learning Environments Workshop, held in conjunction with the Alpine Rendez-Vous '2013*, Villard-de-Lans (France) (2013)

22. Vatrapu, R., Reimann, P., Halb, W. and Bull, S.: Second Int. Workshop on Teaching Analytics. *Proc. Third Int. Conf. on Learning Analytics and Knowledge (LAK'2013)*, pp. 287–289. ACM, New York (2013)

23. Hattie, J.: *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. Routledge, New York (2009)

24. Hogan, T., Rabinowitz, M. and Craven, J.A.: Representation in teaching: Inferences from research of expert and novice teachers. *Educational Psychologist*, 38(4), 235–247 (2003)

25. Doyle, W.: Making managerial decisions in classrooms. In: Luke, D.L. (ed.) *Classroom Management*, pp. 42–74. University of Chicago Press, Chicago (1979)

26. Ward, P., Williams, M. and Hancock, P.A.: Simulation for performance and training. In: Ericsson, K.A., Charness, N., Feltovich, P.J. and Hoffman, R.R. (eds.) *The Cambridge handbook of expertise and expert performance*, pp. 243–262. Cambridge University Press, Cambridge (2006)

27. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. and van de Weijer, J.: *Eye tracking, a comprehensive guide to methods and measures*. Oxford University Press, Oxford (2011)