Mohamed Ez-Zaouia

Factors for dashboards design and use to inform teachers' practices in situ

---

EZ-ZAOUIA Mohamed. *Factors for dashboards design and use to inform teachers' practices in situ,* sous la direction de Elise Lavoué, Université Jean Moulin (Lyon 3), 2020.

Disponible sur : http://www.theses.fr/2020LYSE3030

N° d'ordre NNT : 2020LYSE3030

THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée au sein de

L'Université Jean Moulin Lyon 3

Ecole Doctorale N° 512
Informatique et Mathématiques (InfoMaths)

Discipline de doctorat : Informatique

Soutenue publiquement à 16h, le 02/09/2020, par :

Mohamed EZ-ZAOUIA

Factors for dashboards design and use to inform teachers' practices in situ

---

Devant le jury composé de :

Michel, Desmarais, Professor, Université de Montréal, Rapporteur
Davinia, Hernandez Léo, Full professor, Université de Barcelone, Rapporteure
Mar, Perez Sanagustin, Maître de Conférences, Université de Toulouse, Examinatrice
Yannick, Prié, Professeur des Universités, Université de Nantes Examinateur
Chirine, Ghedira Guegan, Professeur des universités, Université Lyon 3, Examinatrice
Elise, Lavoué, Maître de Conférences HDR, Université Lyon 3, Directrice de thèse
Aurélien, Tabard, Maître de Conférences, Université Lyon 1
François, Paret, Directeur général Woonoz

Factors for dashboards design and use to inform teachers' practices in situ

Factors for dashboards design and use to inform teachers' practices in situ

Mohamed Ez-zaouia

Department of Computer Science
University of Lyon, France

ezzaouia.com

For more information about this dissertation, write to :
ezzaouia.mohamed at gmail.com

2020, Edition

Abstract

Information dashboards have great potential in informing teachers' and learners' work, decisions, and practices. In this dissertation, we study the design and use of dashboards to inform teachers' practices when assisting learners in using online learning platforms. We seek to examine not only what teachers are supposed to do or can do with dashboards, but more importantly what they actually *do* with such technologies in their everyday work and in the long run. We argue that this research draws social, pedagogical, and technical ramifications, which are crucial to articulate. On the one hand, to understand the social factors, we need a critical inquiry into the pedagogical practices and the technical challenges underpinned by teachers' dashboards. On the other hand, to properly design such technologies, we need a profound understanding and consideration of the social interactions of users in the context of classroom situations. We collaborated with teachers and designers of two online language learning platforms to design two dashboards, Emodash and Progdash. We then conducted two longitudinal field studies, lasting two and three months, respectively, to examine the impact of our dashboards on 34 teachers in authentic settings.

First, we studied dashboards in the context of video-conferencing language learning. Teachers face one main challenge: lack of emotional awareness in online learning due to distant and technology-mediated interactions. We conducted a case study examining a multimodal approach of learners' emotions awareness, i.e., applying two theoretical models of emotion: discrete and dimensional, in online learning. We demonstrated this approach by using audio and video streams when inferring emotions along with learners' self-reported emotions, as well as teachers' and learners' contextual interaction traces on the platform. We qualitatively and quantitatively analyzed a total of more than 7 hours' recording. We compared emotional cues from the two modalities, audio, and video. We proposed an approach for combining discrete and dimensional emotional cues to capture learners' emotional states. Early results highlighted the pertinence of rendering multimodal and contextual emotional awareness information for teachers in such settings.

Second, we built upon our case study and adopted an iterative design process in which we interviewed five teachers and collaborated with a pedagogical manager, a language researcher, and two teachers. During the design, we conducted two formal formative evaluations leading to the design of Emodash. We conducted a two-month in-the-wild field study with five teacher-learner pairs, to examine how Emodash supports teachers' retrospective awareness of learners' emotions in online learning and how it impacts feedback reports that teachers write to learners after each learning session. The results showed that Emodash led teachers to write significantly more affective and formative feedback as well as less summative feedback, suggesting a stronger awareness of learners' emotions. Also, the dashboard led teachers to reflect on the way they

conduct their lessons, using learners' emotions as a proxy to evaluate their conduct of teaching.

Third, we studied dashboards in the context of remote and blended learning. Teachers face one main challenge: lack of actionable insights to devise and engage in informed interventions with learners. We designed Progdash based on interviews with seven teachers and refined it through collaborative design prototypes. We implemented and integrated Progdash into an online French vocabulary, grammar, spelling learning platform. We conducted a three-month in-the-wild field study with 29 teachers to evaluate whether Progdash provides useful indicators to teachers about learners' progression, and how it informs teachers' practices to engage in informed interventions. The results showed that teachers found Progdash actionable to inform: monitoring, assessments, planning, information sharing, feedback, and lesson provision. Based on our findings, we discussed implications aimed at improving dashboards to bridge online and in-class learning as well as to foster teachers' and learners' dialog and reflection.

Finally, we took a reflexive inquiry building upon the results of the studies of Emodash and Progdash. We articulated the social factors —monitoring, awareness, and reflection, the pedagogical practices —planning, feedback, and coaching, as well as the technical challenges interacting with the design and use of teachers' dashboards. In each dimension we structured and discussed practical assumptions to inform the design and use of teachers' dashboards. Together, these dimensions, serve as a wider conceptual umbrella for the design of information-driven technologies and macro-implications for dashboards fitting teachers' and learners' everyday situations.

Résumé

Les tableaux de bord ont un grand potentiel pour informer les décisions et pratiques des enseignants et des apprenants. Dans cette thèse, nous étudions le design et l'utilisation des tableaux de bord pour informer les pratiques des enseignants quand ils assistent les apprenants dans lutilisation de plateformes d'apprentissage en ligne. Nous cherchons spécifiquement à examiner non seulement ce que les enseignants sont censés faire ou peuvent faire avec les tableaux de bord, mais surtout *ce qu'ils en font* réellement au quotidien et sur le long terme. Nous argumentons que cette recherche impose des ramifications sociales, pédagogiques et techniques, qui sont cruciales à articuler. D'une part, pour éliciter nos réflexions sur les facteurs sociaux, nous avons besoin d'une enquête critique des pratiques pédagogiques et les challenges techniques engendrés par les tableaux de bord. D'autre part, pour bien désigner ces technologies, nous avons besoin d'une compréhension et dune considération profonde des interactions sociales des utilisateurs dans les contextes et situations d'apprentissage et d'enseignement en classe. Nous avons collaboré avec des enseignants et des designers de deux plateformes d'apprentissage de langues en ligne pour concevoir deux tableaux de bord, Emodash et Progdash. Nous avons ensuite mené deux études longitudinales sur le terrain, d'une durée de deux et trois mois respectivement, pour examiner l'impact de nos tableaux de bord sur 34 enseignants en situations réelles.

Dans un premier temps, nous avons étudié les tableaux de bord dans le cadre de l'apprentissage de langues étrangères par visioconférence. Les enseignants sont confrontés à un challenge central à savoir le manque d'*awareness* (conscience) émotionnelle dans les échanges qu'ils ont avec les apprenants en ligne, dû aux interactions distantes et médiées par la technologie. Nous avons mené une étude de cas examinant une approche multimodale, c'est-à-dire en appliquant deux modèles théoriques des émotions à savoir discrets et dimensionnels. Nous avons démontré cette approche en utilisant les flux audio et vidéo pour inférer automatiquement des émotions ainsi que des émotions auto-rapportées ou "*self-reported*" par apprenants. Nous avons contextualisé les émotions moyennant les traces d'interaction d'enseignants et d'apprenants sur la plateforme. Nous avons analysé 10 sessions d'apprentissage de 45 minutes chacune —un total de plus de 7 heures d'enregistrement. Nous avons proposé une approche pour combiner ces deux modèles d'émotion discrets et dimensionnels pour capturer les états émotionnels des apprenants de manière globale. Les premiers résultats soulignent la pertinence de l'approche multimodale pour inférer les émotions dans un contexte éducatif à distance.

Deuxièmement, en s'appuyant sur la première étude, nous avons adopté un processus de design itératif dans lequel nous avons interviewé cinq enseignants et collaboré avec un responsable pédagogique, un chercheur linguiste et deux enseignants durant la conception de Emodash. Au

cours du design, nous avons effectué deux évaluations formelles d'utilisabilité menant au design final d'Emodash. Nous avons ensuite mené une étude longitudinale sur le terrain de deux mois avec cinq paires enseignant-apprenant, pour examiner comment Emodash facilite l'*awareness* d'enseignants par rapport aux émotions des apprenants et comment cela impacte les bilans pédagogiques que les enseignants écrivent aux apprenants après chaque session d'apprentissage. Les résultats ont montré que Emodash a conduit les enseignants à rédiger significativement plus de feedback affectif et formatif et moins de feedback sommatif, suggérant une meilleure *awareness* des émotions d'apprenants. De plus, le tableau de bord a conduit les enseignants à réfléchir sur la façon dont ils mènent leurs séances de cours, en utilisant les émotions d'apprenants comme un proxy pour évaluer l'expérience d'apprentissage.

Troisièmement, nous avons étudié les tableaux de bord dans le contexte de l'apprentissage mixte où les enseignants articulent des activités d'apprentissage à distance et en face-à-face. Les enseignants sont confrontés à un défi principal à savoir le manque d'informations précises pour informer des interventions pédagogiques afin de mieux assister les apprenants. Nous avons conçu et affiné Progdash sur une base d'entretiens et de design collaboratif avec sept enseignants. Nous avons implémenté et intégré Progdash dans une plateforme dédiée à l'apprentissage du vocabulaire, la grammaire et l'orthographe du français en ligne. Nous avons mené une étude longitudinale sur le terrain de trois mois avec 29 enseignants pour évaluer si Progdash fournit aux enseignants des indicateurs utiles sur la progression des apprenants et comment cela impacte les pratiques enseignantes. Les résultats ont montré que les enseignants ont trouvé Progdash utile pour informer plusieurs interventions : le suivi des apprenants, les évaluations de l'apprentissage, la planification des interventions, le partage d'informations, la provision du feedback. Sur la base de nos résultats, nous avons discuté des implications visant à améliorer les tableaux de bord pour faciliter l'apprentissage en ligne et en classe ainsi que pour favoriser le dialogue et la réflexion des enseignants et des apprenants.

Enfin, nous avons mené une étude réflexive sur la base des résultats des études Emodash et de Progdash. Nous avons articulé les facteurs sociaux —mointoring (suivi et régulation), *awareness* et réflexion, les pratiques pédagogiques —planification, feedback et coaching, ainsi que les challenges techniques liés aux tableaux de bord dans un contexte éducatif en ligne. Pour chaque dimension, nous avons structuré et discuté des pratiques d'appropriations pour informer le design et les usages des tableaux de bord. Ces dimensions articulent un cadre conceptuel global pour la technologie d'information en éducation et ainsi des macro-implications pour des tableaux de bord adaptés aux besoins et situations d'apprentissage et d'enseignement.

Dedication

Les mots manquent aux émotions.
Victor Hugo —1802-1885, Le Dernier jour d'un condamné.

Some people too...

To my family

Acknowledgments

I would like to start off by recognizing the people that just make my work on this fascinating and transdisciplinary research possible. I would like to tell my advisors Elise Lavoué and Aurélien Tabard how much I'm thankful and grateful for their timed feedback, healthy critics, continuous mentoring, and informed coaching. Elise first seeded my knowledge of research thinkings and rigor. Aurélien catalyzed my thoughts on designerly practices and deliberation. No words can express how much I appreciate the rich and formative conversations with both of you.

My research would not be possible without collaboration with great people from Speakplus and Woonoz. Célia, Ciro, Fred, Jordan, Kevin, Léa, François, Morgane, Nicolas, Tiphanie, Xavier, without your warm hosting, relentless help and support rendered at different stages, I would not have been able to reach this far in my research.

I would like to thank all the teachers and learners who have contributed to my research. I deeply appreciate their empathy, precious time and energy. Their expertise and thoughts shaped my research from different perspectives. Without their knowledgable collaboration, I would not have been able to delve into teachers' work.

I thank all the members of the Sical research team for their availability, collaboration, and feedback on my papers' drafts and presentations: Audrey, Christine, Benoît, Bertrand, Karim, Jean-Charles, René, Romain, Thibault.

My special thanks to all my teammates in the lab for their help, formal and informal discussions, and for making the workplace such a pleasant and unforgettable experience: D. Alix, G. Alix, Célia, Denis, Joseph, Lili, Liqun, Nicolas, Qinjie, Rubiela, Stuart, Théo, Valentin.

I thank my defense members for investing their time and energy in reviewing and examining my research work. Chirine, Davinia, François, Mar, Michel, Yannick, I will embrace and apply your constructive inputs for life.

I want to acknowledge my family, especially, my mom, sister, and brothers for their constant encouragement and support throughout this journey. Special thanks to all the people who their names do not appear within these few lines, and have supported me socially, emotionally, and many more means, 'you are part of my life', without any of the help that I received from you all I would not have been able to fulfill this degree.

Last but not least, my research would not be possible without funding from the French Agency of Research and Technology Technology [ANRT, grant number 2016/884 and 2018/0093] in collaboration with both companies Speakplus and Woonoz to which I'm very grateful.

Contents

List of Figures

List of Tables

Publications' Notes

Parts of my dissertation have been previously published:

Journal

- **Mohamed Ez-zaouia**, Aurélien Tabard, and Elise Lavoué (2020a). "Emodash: A dashboard supporting retrospective awareness of emotions in online learning." In: *International Journal of Human-Computer Studies* 139, p. 102411. ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2020.102411. URL: http://www.sciencedirect.com/science/article/pii/S1071581918305585

Conference

- **Mohamed Ez-zaouia**, Aurélien Tabard, and Elise Lavoué (2020b). "Progdash: Lessons Learned From a Learning Dashboard in-the-wild." In: *Proceedings of the 12th International Conference on Computer Supported Education - Volume 2: CSEDU*. **Best Industrial Paper Award**. INSTICC. SciTePress, pp. 105–117. ISBN: 978-989-758-417-6. DOI: 10.5220/0009424801050117
- **Mohamed Ez-zaouia** and Elise Lavoué (2017b). "EMODA: A Tutor Oriented Multimodal and Contextual Emotional Dashboard." In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. LAK '17. Vancouver, British Columbia, Canada: ACM, pp. 429–438. ISBN: 978-1-4503-4870-6. DOI: 10.1145/3027385.3027434

Workshop

- **Mohamed Ez-zaouia** (2020). "Teacher-Centered Design Process." In: *Companion Proceedings 10th International Conference on Learning Analytics & Knowledge, in the 2nd International Workshop on Explainable Learning Analytics*. LAK '20. Frankfurt, Germany, pp. 511–528

Some early/short versions of my papers have been presented during national events:

Conference

- **Mohamed Ez-zaouia** and Elise Lavoué (2017a). "EMODA : un tableau de bord de suivi des émotions des apprenants en formation par visio-conférenc." In: *8 Conférence sur les Environnements Informatiques pour lApprentissage Humain*. Strasbourg, France, pp. 385–387. ISBN: 978-2-9552774-6-1

Summer school

- **Mohamed Ez-zaouia** (2017). "Emodash: Faciliter le suivi des apprenants en ligne." In: *7èmes Rencontres des Jeunes Chercheuses et Jeunes Chercheurs en Interaction Homme-Machine*

Preface

Before delving into my thesis research work, I just want to share a bit of my story navigating my way into my current thesis. In 2013, I graduated from the National Advanced Institute for Information Technology and System Analysis —ENSIAS, French abbreviation, Mohammed V University, Rabat, Morocco, as a software engineer. I don't know how exactly, but during the three years that I have spent at grad school, I have developed a big passion for research in computer science. I joined Accenture, Morocco, during my six months-long graduation project. And, by the end of my internship, I joined the company as a software engineer analyst. My journey at Accenture continued for two years. Although it was a very successful experience for me, my passion for research has been growing since then towards three main fields namely, big data, data mining, and machine learning. In mid-2015, I decide on either quit my job to pursue a research career or find a job in the fields that I'm passionate about. In parallel, I applied for IBM, Morocco, knowing that they have more data-oriented projects —the common trait of the three fields of my interest. And, I start emailing professors regarding Ph.D. offers in Europe. Some professors kindly responded to my emails [...]. I understood that for my case, it was more unlikely to get accepted for a Ph.D. directly before pursuing a research master. So, I applied for some research masters, among them was a master by the Advanced Institute of Computer Science, Modeling and their Applications —ISIMA, French abbreviation, Blaise Pascal University, Clermont Ferrand, France. I knew that I have more chances to get accepted in ISIMA's master as this school has a partnership with my grad school. During this process I have conducted several interviews at IBM, Morocco. I get accepted for several research masters in France. I applied for a student visa. Overall, the whole process took me about three to four months. During the last month I get accepted as a software engineer at IBM and I needed to work one last month with Accenture before joining IBM. I joined IBM, for my very first day, I spent the half of the day with the human resources manager (HRM) signing a large pile of papers or as the HRM called: "*the bible of IBM*". One of the main take-homes of all the papers was that every piece of software, including method, algorithm, product, design, and such, that I may contribute to, belongs to IBM. I remember, when I was steeping through those clauses, I said to myself that sounds good, they might be conducting state of the art software engineering as they are so much explicit [...]. My chief engineer was not in the office that day, the HRM said to me that I can take the second half-day off and come back tomorrow for my team onboarding with my chief engineer. Surprisingly, the same day, I get a notification that my student visa was approved and that I can come the next day to get it. I sent an email to IBM explaining the situation and my motives, passion for research [...].

In September 2015, I came to France, for my first time, I conducted a research master 2 in Models and Algorithms for Decision Support which was good as it was related to my fields of interest. For my internship, I applied to my advisor Aurélien about an offer, saying: "*[...] Dans le cadre de mon Master,*

je dois effectuer un stage de fin détudes de recherche. Le sujet de votre stage mintéresse beaucoup car il correspond parfaitement à ma grande passion pour le Big Data, Data Mining et Machine Learning". Aurélien kindly replied to my email saying: *"[...] merci de votre intérêt, votre profil est intéressant, mais venons de confirmer le stage avec un autre candidat. Si vous êtes intéressé par un stage combinant visualisation et données, je vous invite à regarder l'offre de stage suivante : https://speakplus.recruiterbox.com/jobs/fkohywp/ Elle porte sur un projet de visualisation d'émotions dans un environnement d'apprentissage, et peut déboucher sur une thèse CIFRE (collaboration laboratoire - entreprise)*". And that was the very beginning of my journey into this thesis.

Mohamed Ez-zaouia
Lyon, France, 25 May 2020

Chapter 1

Introduction

> *The educational marketplace is filled with products promoting learning dashboards [..]. However, the failure to transition from exciting concept demonstrators to embedded practical tools has long dogged educational technology.*

—Shum et al., 2019, p. 2

1.1 Overview

Over the last two decades, technology has been transforming ways in which teaching and learning are delivered to millions of learners around the globe (Technavio, 2016) for both formal and non-formal education and training. Many stakeholders, policy makers, and researchers advocate education technology to improve experiences and outcomes (Siemens and Long, 2011; Chasteen et al., 2011; Ferguson, 2012). Teachers are, however, the *front line end users*, who strive to orchestrate technology needs and uses for both inside (Vermette et al., 2019) and outside (Hillman et al., 2019) classrooms, as well as to harness data to improve teaching and learning (Campaign, 2018).

In face-to-face learning environments, teachers are, more or less, skilled to grasp the classroom's climate. They frequently observe the status of their living classroom. They can assess learners' participation and involvement with the learning activities. They can build an awareness of contingencies taking place at individual and group levels. They can engage in formative interventions. They (re)scaffold, (re)instruct, (re)explain parts and aspects of the materials gearing lessons to successfully attain the learning objectives (Hattie, 2012, p. 81).

In online learning, the vital cues of co-located human interactions, including facial expressions, direct contact, dialog with and feedback from learners, among others, are to a large extent lacking. It can be hard for teachers to build a strong emotional, social, and instructional awareness of learners' online learning (Jennings and Greenberg, 2009). Additionally, learners develop multifaceted behaviors online, including different learning styles, acquisition, progression, pathways, pacing (Onah et al., 2014; Y. Lee and Choi, 2010). Coupled with the challenges of articulating both remote and co-located learning activities (Vermette et al., 2019; Hillman et al., 2019), it can be hard for teachers to acquire fine-grained knowledge about learners' interactions with the online learning materials.

Education technology can yield large streams of information on a daily basis, which can provide valuable insights into learners' behaviors, processes, progress, and skills (Siemens, 2013; Siemens and Baker, 2012). Analyzing learning traces to inform day-to-day practices is, however, both a complex and a laborious process (Baker and Inventado, 2014). Teachers come up against barriers for using learning data in their schools. They lack data, visual, and analytical literacy training to analyze, make use, and sense of learning data (Barbara Wasson, 2015). They also lack formal time-slots for leveraging data-informed practices (Campaign, 2018; Sandford et al., 2006). There is a growing need for tools that translate learning data into actionable insights for teachers. *Information dashboards* can fill this gap.

Dashboards are now commonplace in society (Few, 2004; Alper Sarikaya et al., 2018). They capitalize on our familiarity and experiences in using different information-driven technologies for everyday use: phones, watches, monitors, to name but a few. They present formative insights in a way that facilitates at-a-glance readings. They can shape presentations to fit a wide range of audiences. For teachers and learners alike, dashboards seem promising tools to support a variety of practices on a regular basis, such as monitoring learners (Holstein et al., 2018), supporting learners with lesser abilities (Molenaar and Campen, 2017), providing personalized support to learners in a timely fashion (Aslan et al., 2019), and providing feedback to learners (Roberto Martinez-Maldonado et al., 2015).

Recent research, however, underlines the lack of longitudinal research examining the design, use, and impact of dashboards on teachers' practices in "real-world" settings (Rodriguez-Triana et al., 2016; Schwendimann et al., 2017; Bodily et al., 2018). From a *design perspective*, Shum et al. (2019) argue for a growing need for design-based research centered around needs of teachers and learners in-situ to best serve pedagogical practices through dashboards. Verbert et al. (2014) argue the need for leveraging authentic and "richer set of relevant data" sources in education dashboards. Schwendimann et al. (2017) underline the need for examining dedicated designs that go beyond simple logs-driven summarization metrics in dashboards. From an *empirical perspective*, several reviews of more than 150 dashboards underline the need for longitudinal field studies to better understand how teachers appropriate, leverage, and shape their pedagogical practices around dashboards in the long run (Verbert et al., 2014; Schwendimann et al., 2017; Bodily et al., 2018). From a *methodological perspective*, much work on teachers' dashboards remains somewhat disjoint. Shum et al. (2019) underline the need for macro-conceptual implications covering the design space and accounting for the interactions between underlying users and stakeholders. Ferguson (2012) highlights the importance of triangulating the social, pedagogical, and technical ramifications in the design and use of learning analytical technologies.

This dissertation further examines teachers' dashboard design and use in-the-wild. Specifically, it provides empirical findings from two design and field studies on:

1. Teachers' routines surrounding dashboards. We examine the benefits for teachers from using dashboards to improve learning experiences and outcomes of learners.

2. Design and use implications of the social, pedagogical, and technical dimensions of dashboards. These implications can guide information-driven classroom technology and pedagogy.

1.2 Thesis Statement

The use of learning technology brings forth both opportunities and challenges. Such technologies often add more features, and tend not to consider teachers' practices as much. From the learners' perspective, they are often designed around "do your best". Teachers' dashboards can bridge the gap between learning technology and pedagogy. We argue that this research draws social, pedagogical, and technical implications. On the one hand, to understand the social factors, we need a critical inquiry into the pedagogical practices and the technical challenges underpinned by teachers' dashboards. On the other hand, to properly design such technologies, we need a profound understanding and consideration of the social interactions of users.

1.3 Research Approach

We conducted this thesis in the context of two commercial learning platforms. Both platforms specialize in online language learning. SPEAKPLUS is a video-conferencing platform dedicated to improving oral communication skills in a foreign language, i.e., English, French, and Spanish. Teachers use SPEAKPLUS to deliver virtual live classes to learners. PROJET-VOLTAIRE is a web-based platform dedicated to learning French vocabulary, grammar, and spelling concepts. Teachers use PROJET-VOLTAIRE to level up the writing skills of learners. Both platforms are used by public/private schools, as well as professional teachers. With SPEAKPLUS we examined teachers' dashboards in a fully remote learning setting. On the other hand, with PROJET-VOLTAIRE we examined teachers' dashboards in a blended setting where teachers articulate both face-to-face and remote online learning. Together, both platforms provided us with interesting contexts for examining the design and use of teachers' dashboards in real-world learning settings.

We took an iterative mixed research approach and built upon research from different fields, including social science, human-computer interaction, human-centered design, information visualization, learning analytics. Figure 1.1 illustrates the methodological, design, and empirical approach of this thesis.

In the context of SPEAKPLUS, teachers face one main challenge: lack of emotional awareness in online learning, which is due to remote and technology-mediated interactions. We built upon emotion models to seamlessly incorporate, in the platform, cues to foster teachers' awareness of learners' emotions. We then iteratively designed EMODASH an interactive dashboard, to support teachers' feedback reports, written to learners after each live learning session.

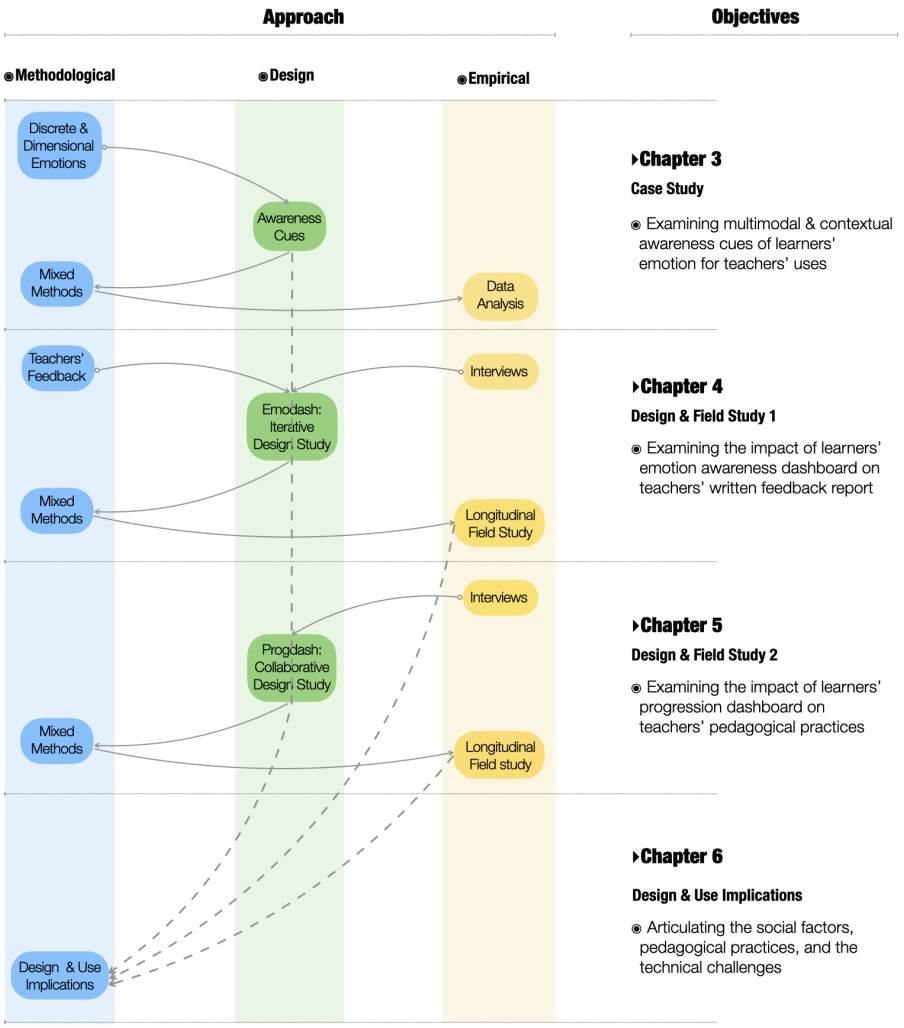


Figure 1.1: The methodological, design, and empirical research approach and objectives of this thesis.

In the context Projet-Voltaire, teachers face one main challenge: lack of fine-grained feedback about learners' interactions with online learning materials to inform pedagogical interventions. We conducted a design study (Sedlmair et al., 2012) with teachers. We collaboratively designed Progdash to assist teachers in bridging between learners' online learning and data-informed pedagogical practices.

We built upon the results of our studies of Emodash and Progdash to drive the social factors, pedagogical practices, and technical challenges intersecting and interacting in the design and use of teachers' dashboards (Shum et al., 2019; Ferguson, 2012).

We hope to contribute from three perspectives:

First, from a **design perspective**, we designed, implemented, and deployed two fully functioning teachers-centered dashboards in real-world settings, to study teachers' dashboards in-situ:

- EMODASH, a dashboard designed based on interviews with five teachers and one learner and refined through an iterative design process, to facilitate teachers' retrospective awareness of learners' emotions, and support teachers' writing of reinforcement feedback reports to learners after video-conferencing language learning sessions.
- PROGDASH, a dashboard designed based on interviews with seven teachers and refined through collaborative design prototypes, to bridge between learners' use of an online grammar learning platform and teachers' data-informed pedagogical practices.

Second, from an **empirical perspective**, we designed and conducted qualitative and quantitative analyses to distill empirical evidence to both inform the design and assess the impact of our built artifacts with respect to teachers' needs and practices:

- We conducted several informal discussions with designers of two online language learning platforms to understand the pedagogical challenges of teachers when using such platforms, as well as their inner working.
- We conducted several formal interviews with teachers to understand their needs, challenges, and pedagogical practices through which they aim to influence and reinforce learners' learning.
- We designed and conducted two of the first design studies and in-the-wild longitudinal field studies, lasting two and three months, respectively, to study how 34 teachers —from both studies, make use and sense of dashboards to inform their pedagogical practices.
- We conducted several formal post-interviews with teachers to evaluate and report on the usefulness, effectiveness, and impact of our artifacts on teachers' everyday practices.
- We conducted a qualitative and quantitative exploratory analysis comparing learners' inferred emotions from two different modalities, namely audio and video streams.
- We conducted a qualitative and quantitative analysis comparing teachers' written feedback reports *with* and *without* EMODASH; and also to articulate teachers' pedagogical practices when using PROGDASH.

And finally, from a **methodological perspective**, we built upon a broad spectrum of research theories, methods, and studies to both inform the design and assess the use of teachers' dashboards in context:

- We built upon emotion recognition theories and studies from psychology (Ekman and Friesen, 1976; Barrett and Russell, 1998) to ground, capture, quantify, and integrate cues (S. K. D'Mello, 2017) to foster teachers' awareness of learners' emotions in online learning environments.
- We proposed a method to unify discrete and dimensional models of emotions to make it easier to combine and use them in a multimodal fashion (Ez-zaouia and Lavoué, 2017).
- We conducted an exploratory analysis of more than 7 hours' recordings to compare learners' inferred emotions from two different modalities: audio and video streams, when using video-conferencing learning sessions.
- We conducted an exploratory analysis of more than 10 hours' screen recordings to examine teachers' experience when writing feedback reports with the use of Emodash.
- We conducted a qualitative thematic analysis (Braun and Clarke, 2006) of interviews to assess the impact of both studies of Emodash and Progdash.
- We conducted a quantitative non-parametric factorial analysis (Wobbrock and Kay, 2016) to compare teachers' written feedback reports *with* and *without* the use of Emodash.
- We proposed design implications that triangulate the social, pedagogical, and technical dimensions underpinned by teachers' dashboards.

1.4 Thesis Outline

In chapter 3, we examine multimodal and contextual awareness cues of learners' emotions for teachers' uses. We present a case study of learners' emotional cues. We apply two theoretical models of emotion: discrete and dimensional. We demonstrate this approach in the context of Speakplus. We propose to infer emotions from the audio and video streams of the live learning session, along with learners' self-reported emotions. We contextualize emotional cues using teachers' and learners' traces on the platform.

In chapter 4, we examine the impact of learners' emotions awareness dashboard on teachers' written feedback reports. We present a design and longitudinal field study of Emodash. We designed Emodash based on interviews with five teachers. We refined the design through an iterative design process. We evaluated its impact through a two-month field study with five professional teachers and five learners.

In chapter 5, we examine the impact of learners' progression dashboard on teachers' pedagogical practices. We present a design and longitudinal field study of PROGDASH. We designed PROGDASH based on interviews with seven teachers. We collaboratively iterated and refined the design through informal discussions with teachers and designers of the platform. We evaluated its impact through a three-month field study with 29 teachers from middle, high, and vocational schools, and university.

In chapter 6, we articulate the social factors, the pedagogical practices, and the technical challenges of the design and use of teachers' dashboards. We engaged with these three dimensions, to examine the opportunities and challenges that we may face in the development of teachers' dashboards. Together, these dimensions serve as a conceptual umbrella for the design of information-driven technologies and macro-implications for dashboards fitting teachers' and learners' everyday situations and contexts.

Chapter 2

Background and Related Work

> *Dashboards can tap into the tremendous power of visual perception to communicate, but only if those who implement them understand visual perception and apply that understanding through design principles and practices that are aligned with the way people see and think. Software won't do this for you. It's up to you.*
>
> —Few, 2006, p. 12

2.1 Overview

In this dissertation, we investigate the design and use of teachers' dashboards to inform pedagogical practices when assisting learners in using online learning platforms. We seek specifically to extend our understanding of not only what teachers are supposed to do or can do with dashboards, but more importantly what they actually *do* with them in everyday practice, along with sociotechnical factors, pedagogical practices, and the design challenges that such tools underpin. We build upon a large body of research from different fields, including social science, human-computer interaction, human-centered design, information visualization, learning analytics, and the wider technology-enhanced learning domains, to name but a few.

In this chapter, we structure and discuss related works in six main themes relevant to this dissertation:

1. **Introducing Information Dashboards:** In this section, we place the work conducted for this dissertation in the context of information-driven dashboards. In particular, we emphasize the most interesting and challenging features of dashboards, such as visual encoding, different contexts of use, different audiences, glanceable readings.

2. **Introducing Learning Analytics:** In this section, we introduce the field of learning analytics. We place particular emphasis on harnessing learning data to improve teaching and learning experiences and outcomes through the use of information-driven tools and visualizations when leveraging human informed judgment.

3. **Monitoring Learners' Activity:** In this section, we focus on research into teacher-facing dashboards aimed at supporting monitoring of learners' activity when using online learning materials. We organize and discuss research in this area in three main themes: proactive monitoring, classroom monitoring, and retrospective monitoring.

4. **Fostering Sociotechnical Factors:** In this section, we focus on research into fostering sociotechnical factors through the use of teacher-facing dashboards, particularly highlighting how dashboards can go beyond log-based measurable metrics and features by fostering important human factors, such as data sense-making, awareness, socio-affective relationships, and reflection, when using learning dashboards.

5. **Evaluating Teachers' Dashboards:** In this section, we focus on methods adopted when evaluating teachers' dashboards. In particular, we organize and discuss empirical studies in two main themes: usability and usefulness oriented evaluations, as well as pedagogical practice-oriented evaluations.

6. **Dashboard Design and Use:** Taking a top-down approach, in this section, we discuss work on teachers' dashboard design and use from a methodological perspective.

2.2 Introducing Information Dashboards

Before describing and discussing some interesting features —which merit in-depth study and research, of *dashboards* in informatics or, as we know and use them nowadays, the etymology of the word "dashboard" also provides some noteworthy facts.

2.2.1 Origin —Enabling Drivers to See Along and Ahead

The main strand of the origin of the term "dash-board" can be rooted in motive (and motion), around 5,000 years ago or, more specifically, to the human-horse relationship (de Barros Damgaard et al., 2018), which was a key turning point in revolutionizing not only transport but the world as we know it today. Although the solid disk wheel was first invented way before domestication of the horse, in the fairly early Bronze age, it was only in the middle Bronze era that Bronze age inventors put together the chariot: a box on top of wheels. This invention opened up immense opportunities for human travel and carriage.

Interestingly, the use of the term "dashboard" originates from the jargon of the horse-drawn carriage. Way back then, as Figure 2.1 portrays, since most routes were not paved, people came up with the idea of placing a *board* ("screen") made of wood or leather —which they called "dashboard", on the forepart of the horse-drawn open carriage to *"stop [prevent] mud from being splashed ["dashed"] into the vehicle by the horse's hooves"* (Etymonline, n.d.). The *tool* was designed to intercept dirt so to protect the vehicle and its passengers, and by doing so the dashboard freed the *riders* ("drivers") from distractions to enable them to focus and see ("monitor") the route along and ahead.

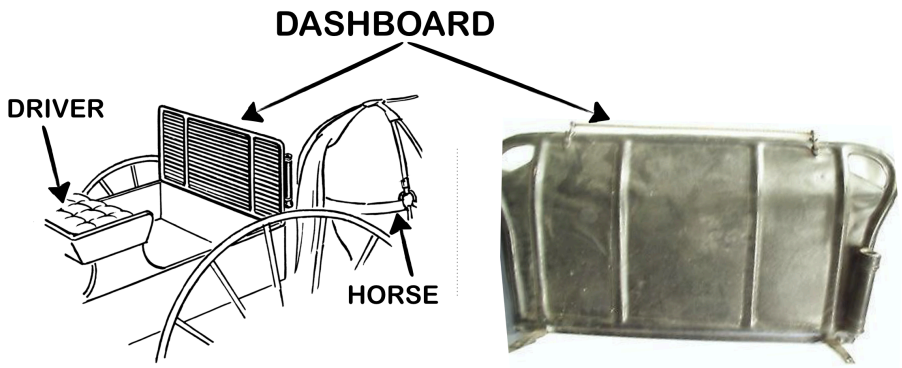


Figure 2.1: Example of a horse-drawn carriage "dashboard". The tool is designed to protect the driver and passengers by catching dirt, dust, and debris thrown up by the horse's hooves. A salient function of the tool is to free the rider from distractions so s/he can see the route along and ahead.

Moving from the horse-drawn carriage to much more modern automotives, see cars, dashboards have retained not only the name but also their focal place and function. As Figure 2.2 portrays, during this evolution, dashboards have evolved from a simple wooden or leather static object in front of the rider to a board *"on which control panels and gauges are mounted"* (Etymonline, n.d.) to enable drivers to best monitor the status of the vehicle and control its behavior. Over time dashboards have gained an essential role in the history of monitoring by providing drivers and passengers with formative, quality, and safety ensuring metrics.

2.2.2 Raise —Supporting Data-driven Business Decisions

It took only 10 years for dashboards —from the first distributed message communicated over ARPAnet in 1969s, to pass across cars to computers. Figure 2.2 portrays a nice joining between a car's dashboard and a digital one: for instance, the Tesla model S dashboard features a full data-driven monitoring digital tool for the whole car. Digital dashboards, as we now use them, first emerged in the 1980s with executive information systems (EISs) summarizing key status measurements, gathered from corporate internal and external resources, and displayed via a simple user interface so that *"even an executive could understand"* (Few, 2006, p. 14), to support data-driven decision-making and communication of information within and across corporate departments (Creps and O'Leary, 1994).

Follow-up work, during the 1990s, focused on the technical aspects of collecting, storing, and processing data, building the infrastructure required for a new field: online analytical processing (OLAP). During this period,

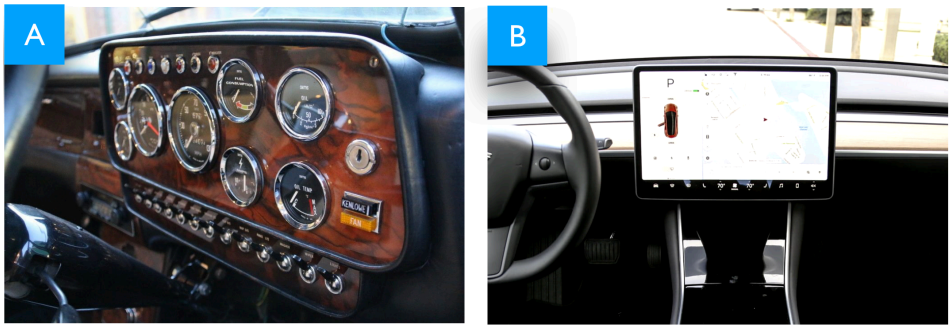


Figure 2.2: On the left (A) is an example of an old Mini Cooper 1275-S-Mk3 car's dashboard. This dashboard presents critical status information, using gauges for, e.g., speed, fuel, oil, water; color-coded flags; buttons; also, the information is structured on the layout to facilitate glanceable readings, for instance, important metrics are scaled up and placed in the center to best match the viewport of the driver —almost all the ingredients of a digital dashboard. On the right (B) is an example of a Tesla model S car's dashboard. This dashboard portrays a nice joining between the car's dashboard and an information-driven digital tool for monitoring and controlling the car.

the focus on business analytics and business intelligence resulted in the raise of key performance indicators (KPIs) dashboards (Kaplan, Norton, et al., 1992) as effective tools to monitor and assess system performance.

Over time, information dashboards have evolved as types of information visualization to more sophisticated information portals, enabling users to make use and sense of data, to inform work, decisions, and practice (Alper Sarikaya et al., 2018). However, dashboards have retained their main role —inherited from car dashboards: in particular, a small appealing board that hides the immense complexity of the underlying system by probing key formative, quality, and safety ensuring metrics, presenting critical information to enable at-a-glance readings, and shaping the presentation to fit a wide range of audiences.

2.2.3 Formation —Key Features of Information Dashboards

In 2006, Stephan Few proposed one of the first working definitions of an information dashboard as: *"a visual display of the most important information needed to achieve one or more objectives consolidated on a single screen so it can be monitored and understood at a glance"* (Few, 2006, p. 26). Few's definition entails what makes dashboard design interesting and challenging at the same time: easy-to-read, understand, and make sense and use of

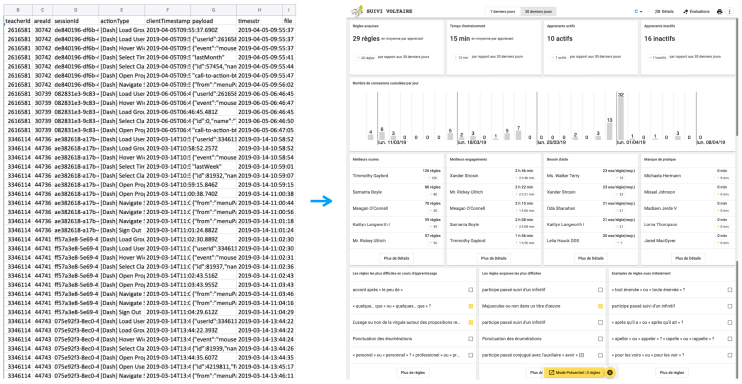


Figure 2.3: Example of a dashboard emphasizing raw learning data through visual encoding —PROGDASH Ez-zaouia et al., 2020.

large, complex, critical amounts of information at-a-glance. To that end, dashboards leverage different key features:

Dashboards **emphasize visual encoding**. As Figure 2.3 portrays, data visualization is at the heart of dashboard design and use. Data in dashboards are abstracted using verbals —text, numbers, and graphics, but with more emphasis on graphics so as to capitalize on human perceptual and cognitive abilities for processing visual information. To that end, along with verbals and numbers, data need to be mapped effectively to graphical elements and properties (Cleveland and McGill, 1984a). Point, line, surface, and volume are the basic graphical elements that can be used and combined to create visual forms. Position, size, color, orientation, texture, and shape are graphical properties that can be used to decorate visual forms.

Dashboards **put forth glanceable readings**. This is important as the essence of dashboards is to emphasize key indicators by compacting the required —all related and relevant, information in a small amount of visual space to inform the audience in a meaningful, efficient, and actionable way at-a-glance. A dashboard serves its purpose well when it fits into a single screen, that is, viewers do not have to scroll or switch between different screens to grasp the needed information. In the former case —*scrolling*, we exceed the viewport of the viewer by expanding the boundaries of the dashboard; and in the latter case —*switching*, we fragment the information and perhaps its context too. In both cases, we hamper the perceptual and communicative intent of the dashboard.

Dashboards **facilitate drilling up and down**. Few's definition can be seen as restrictive as it requires a dashboard to fit a single display (Few, 2006). However, Few's definition aims at overcoming poor dashboard design by delivering the *needed* information in a way to catch viewers' attention at a first glance. As Figure 2.4 portrays, dashboards build fundamentally upon

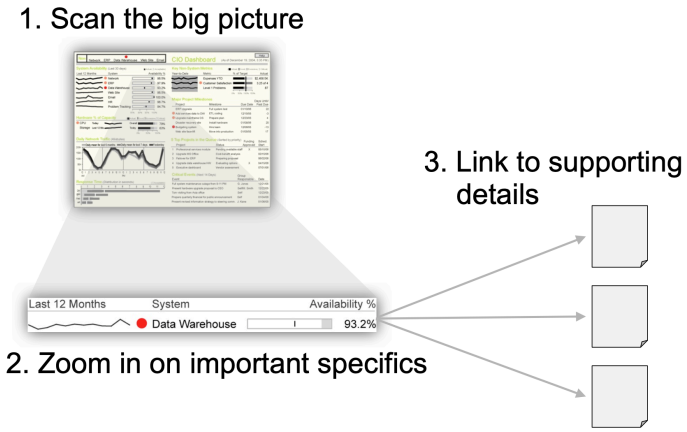


Figure 2.4: Overview, detail on demand —Few, 2013.

the mantra: *"overview first, zoom and filter, then details on demand"* (Shneiderman, 2003). In the overview: viewers get the intended message directly at-a-glance without overwhelming actions or distractions. Interaction methods can then be used to let the viewer examine in more detail specific aspects and different levels of data presented on the dashboard. This can be achieved through navigation and drilling up and down.

Dashboards **fit a wide range of audience**. Dashboard simplicity is intentional to support a wide range of different users with different data, visuals, and analytic literacy. That is, users can interact with dashboards through aggregation, filtering, searching, drilling up and down to make use and sense of the information presented (Alper Sarikaya et al., 2018). However, unlike visual analytics, dashboards do not require end users to engage in active analytical interaction with data, for example, to explore and confirm hypotheses, categorize and identify interesting features of data for further examination (Vieira et al., 2018; Thomas and Cook, 2006).

Dashboards **leverage human informed judgment**. As information-driven tools, design and use of dashboards are, importantly, both: "open" and "slow". Dashboards are open in the sense of their informative design, while they are slow as they are not primarily designed to do things "fast" or to perform productivity. Instead their use is more tailored to foster data sense-making, hence, falling into what is known as the slow category technology (Hallnäs and Redström, 2001). As such, dashboards aim at leveraging human perceptual and cognitive abilities of processing information to help users visually explore data of interest, reason about and make sense of what they see, help guide them towards effective informed decisions, actions, learning, communication, awareness, and reflection (Brath and Peters, 2004; Alper Sarikaya et al., 2018).

2.2.4 Adoption —Facilitating Data-driven of Everyday Practices

Dashboards are now widespread in society. Their interesting features, —highlighted above, make them promising tools for harnessing data to support different day-to-day practices for non-experts (Alper Sarikaya et al., 2018). Academia is not exception when adopting information dashboards. Schools tend most to lower barriers for teachers' adoption of technology for their classrooms (Millar, 2013; Buabeng-Andoh, 2012; Technavio, 2016). Recently, for instance, 71% of surveyed faculties were willing to adopt online materials if they are of high quality and affordable (Consortium, 2016). However, education technology is not without limitations or drawbacks, a problem that has not gone unnoticed by researchers.

In fact, the increased mixing of digital and traditional classrooms has highlighted several challenges for making technology work to its best in everyday learning. For example, understanding learners' state of learning online and devising informed interventions to best engage, coach, and support them when using online learning materials continue to be some of the major issues facing this learning medium (Onah et al., 2014; Y. Lee and Choi, 2010). Fortunately, use of technology yields important streams of traces that can provide valuable insights —if captured and made available to underlying users.

By early 2000, formal data-driven approaches to support informed decisions, assessments, and communication were finding their way in educational contexts, with the goal of harnessing and learning from generated and collected data. The line of reasoning is, in part, to improve teaching and learning outcomes by providing meaningful and actionable data-informed insights to the underlying stakeholders, such as teachers, learners, advisors, administrative staff members, and course designers (Siemens and Long, 2011). Also, in part, a more longitudinal and holistic ambition is actually to understand learners' behaviors, interactions, progress, and process, to provide continuous knowledge about learners' state of learning and skills, and to transfer knowledge from one session to another, from grade to grade, and from one setting to another (Siemens and Baker, 2012).

2.3 Introducing Learning Analytics

"The data flood is here! Traditional methods of data analysis have not kept pace with the amount of data that can be collected and is being collected from educational environments today." (Baker, Duval, et al., 2012, p. 2).

Schools are increasingly deploying learning management systems to deliver course contents online (Paulsen, 2003). Self-paced learning environments and intelligent tutoring systems are also becoming popular in classrooms (Adkins, 2018; Mamun et al., 2016; Kulik and Fletcher, 2015). Moreover, massive open online course platforms, e.g., Edx, Coursera, Khan Academy, as well as paid ones, e.g., Udacity, Udemy, Lynda are being increasingly adopted by leaners for both formal and non-formal education and training (Crues et al., 2018). Learners step through different materials on such platforms, leaving behind important learning traces stored on a daily basis by server logs, which can provide valuable insights into learners' behaviors, process, progress, and skills.

The blending of classrooms with digital learning tools and materials, as well as the resulting challenges of harnessing technology to best fit a complex learning and teaching environment, led to the increase in education data-analytics. Research fields, such as Educational Data Mining and Learning Analytics and Knowledge, emerged. The international society of educational data-mining defines this field as: *"an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand learners, and the settings which they learn in"* (Siemens and Long, 2011). Siemens and Long (2011) defined the field of learning analytics as: *"the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs"*.

The overview of the work of education data-analytics communities is structured around: the *what*, e.g., data, environments; *why*, e.g., monitoring, analysis, prediction, intervention, tutoring, mentoring, assessment, feedback, recommendation; *who*, e.g., learner, teacher, advisor, course designer; *how*, e.g., statistics, data mining, machine learning, visualizations (Chatti et al., 2012; Peña-Ayala, 2018). Although both educational data-mining and learning analytics communities work towards improving learning and teaching by assessing data, it is worth noting that they have different approaches, i.e., in terms of *why* and *how*, for achieving these shared objectives. The key distinction is that the former places *"greater focus on automated discovery"*, i.e., focuses on methods for making and automating discoveries, while the latter places *"greater focus on leveraging human judgment"* through reporting tools and visualizations (Siemens and Baker, 2012), i.e., focuses on the process of collecting, analyzing, measuring, and reporting insights to help underlying stakeholders make use and sense of data, and guide them in making informed decisions.

2.3.1 Harnessing Learning Data

Learning analytics rely on data or, more specifically, learning traces — timestamped click-stream logs, and other learners' related data —bodily signals, self-reported data, grading etc., to improve teaching and learning. For example, learners' use of learning management systems results in collecting large amounts of data about learners' actions, quiz completions, learning artifacts (Paulsen, 2003). Enrollment into massive online courses leads to collecting demographic information, action logs about video and exercise navigation, quiz performance, forum discussions (Crues et al., 2018). Similarly, the increase in self-paced learning environments results in collecting learners' actions, errors, struggles, attempts, moments of idle (Adkins, 2018; Mamun et al., 2016).

Learners step through different content and materials on such learning environments, leaving behind valuable learning traces.

Given the complex nature of learning and teaching, often it is difficult or insufficient to rely on one modality, e.g., logs, to provide insights into learners' states of learning. One of the ambitious goals of learning communities is to collect and use different modalities —also referred to as multimodal analytics, such as using video, audio, gestures, emotions, logs, cognition loads, among others, to characterize, design, and study learning and teaching in a holistic way (Worsley and Martinez-Maldonado, 2018).

In the same way, such modalities are often longitudinal with longer time frames, and are complex and heterogeneous, thus making it difficult to mine for utility purposes using traditional methods and tools, e.g., spreadsheets. Furthermore, teachers, learners, and school staff members face barriers when using data in their workplaces by themselves, including a lack of time (Campaign, 2018; Sandford et al., 2006) and data, visual, and analytic literacy to process, analyze, and make sense and use of data (Barbara Wasson, 2015). There is thus a growing need for methods and tools that facilitate "translation" of learning data into actionable insights for non-experts.

Learning communities work towards building reliable methods, models, techniques, and tools to harness data for critical value (Chatti et al., 2012; Baker and Inventado, 2014); starting by capturing and preparing learning traces into usable forms (Kandel et al., 2011); summarizing and transforming usable traces into insightful indicators (A. Sarikaya et al., 2018); then encoding those indicators into meaningful visual representations to support information-driven practices in teaching and learning settings (Vieira et al., 2018).

2.3.2 Leveraging Teachers and Learners Judgment

One of the key focuses of the learning community is *"leveraging human judgment"* through data-driven approaches and tools (Siemens and Baker, 2012). The globalization of dashboards and their interesting features: emphasizing key formative, quality, and safety ensuring metrics, glanceable readings, wide audience fitting (Alper Sarikaya et al., 2018), caught the interest of learning communities to investigate them when facilitating different kinds of information-driven applications, such as monitoring, regulation, awareness, reflection, learning, collaboration, and more (Chatti et al., 2012; Vieira et al., 2018).

In education settings, information-driven tools are roughly inspired by two approaches, namely "quantified-self" and "data-analytics". In quantified-self, for instance, in health-care, users collect data on their own life's habits, e.g., diet, workout, sleep; and visualizations are used to help users' make use and sense of data and facilitate self-awareness and self-reflection. Most learning dashboards, inspired by quantified-self, are learner-centered (Bodily et al., 2018; Jivet et al., 2017), with the goal of facilitating self-monitoring, awareness, reflection, and metacognition (duval, 2011; Loboda et al., 2014; Ruiz et al., 2016).

The data-analytics approach, on the other hand, considers that teaching and learning are complex internal processes that need to be formally informed using systematic approaches. For instance, when a teacher opts for an existing learning material or activity or updates it, how is the learning really improved, and how can such improvements best be assessed? The analytical approach aims at filling this gap by leveraging formative assessment and data-driven decision support through teacher-centered dashboards (Black and Wiliam, 1998; Wayman, 2005; Sedrakyan et al., 2019). Both movements, learner-centered and teacher-centered, have different approaches for collecting, analyzing, and designing visualizations to assist teachers and learners in emphasizing indicators about "what is doing on" to improve teaching and learning experiences and outcomes.

2.3.3 Teachers' Dashboards

When drawing a parallel with traditional learning, for instance, in face-to-face, teachers frequently track and assess different cues about learners, such as participation, social and behavioral interactions, or engage directly with learners in discussions to assess their understanding and involvement with learning activities. Such cues make the state of learning *visible* for a teacher to devise informed interventions. In online learning, teachers and learners

lack natural cues, such as gestures, facial expressions, direct contact, dialog with and feedback from others. Dashboards can compensate such vital cues of a shared context.

Ever since the increase in data-driven learning communities, dashboards have been at the forefront of digital tools, when harnessing learning traces and other learners' related data, to improve learning and teaching. *Teachers' dashboards* are a type of information dashboards (Few, 2004; Alper Sarikaya et al., 2018; Smith, 2013) aimed at processing, summarizing and presenting meaningful, informative, quality, and safety ensuring indicators about state of learning to teachers. The simplicity of dashboard design and use, visual encoding, easy-to-read, glanceable readings, wide audience fitting, makes them a good fit when instrumenting teachers to allow in-time informed pedagogical interventions.

Early work on learning dashboards focused on instrumenting teachers with data-driven tools to support formative assessments and accountability (Black and Wiliam, 1998; Wayman, 2005). Follow-up research focused on leveraging teachers' judgment through reporting tools and visualizations (Siemens and Baker, 2012; Verbert et al., 2014). Recently, a wide range of teachers' dashboard contributions has been introduced spanning a variety of aspects.

2.3.4 Summary Notes

The research into dashboards could be seen as having an ambivalent audience, roles, and applications in education setting literature (see a review, Schwendimann et al., 2017). Specifically, in terms of "audience", such work can be grouped into two separate, but partially overlapping, strands.

The first strand is driven by theories and studies advocating a shift in the focus from the teacher to learners, namely constructivism in both cases: cognitive or individual informed by Piaget's theory and/or social informed by Vygotsky's theory (Kalina and Powell, 2009). In this model of —learners-centered learning, learners are the active actors in the process of developing, constructing, and acquiring knowledge along the way. Several dashboards have been designed to instrument learners in facilitating self-monitoring, awareness, reflection, and meta-cognition (see reviews, Bodily et al., 2018; Jivet et al., 2017).

The second strand, which is directly relevant to the work of this dissertation, focuses on empowering teachers through the design and use of dashboards to allow them to effectively conduct informed interventions. Effective teachers still play a crucial role as they are experts in the classroom's everyday routines. They coach, inspire, challenge, motivate, and adapt to individuals, thus having an essential role in moderating and facilitating learning (Hattie, 2012).

In this dissertation, we hope to narrow down the scope of teacher-facing dashboards, to identify typical practices surrounding such tools in terms of design goals and purposes, data-informed practices, evaluation, and challenges. We hope to extend our understanding of what teachers actually *do* with dashboards in everyday practice and in the long run, and what *sociotechnical* factors and *pedagogical* practices surround teachers' dashboard design and use.

Consequently, we ground our approach on the assumption that we can augment and support teachers' knowledge through insightful data reporting tools (Baker, 2016; Donald A Norman, 1991; Siemens and Baker, 2012), allowing them to guide their instructions and pedagogical actions to better support learners to learn and improve. In a complex sociotechnical ecosystem of teaching and learning, we believe it is the combination of teachers' expertise, intuitions, and data-driven insights that will most likely improve experiences and produce the best outcomes.

2.4 Monitoring Learners' Activity

Monitoring has been well studied in organizational design (Ouchi, 1979) and computer-supported cooperative work (CSCW), and has been defined as an *"act of observing the actions of others to determine whether they comply with a predefined set of behaviors"* (Robert, 2016). This definition entails two important facts regarding the practice of monitoring: tracking and evaluation (Kirsch, 1996). Tracking involves allocating expectations to members and frequently assessing their associated status: for example, a teacher may expect learners to make an agreed-upon amount of progress on a weekly basis in some topics, so s/he needs to deliberately monitor learners' progression; also, the evaluation involves specifying a set of predefined measures to ascertain expectation status with respect to a desired set of behaviors and/or outcomes: for example, a teacher needs to lay down rules against which s/he will compare learners' behaviors or the outcomes of their behaviors. In this respect, the aim of monitoring is to change (or regulate) individuals' behavior, hence also referred to as behavior control (Eisenhardt, 1985; Kirsch, 1996). The practice of monitoring is consid-

ered as effective when the individuals — monitored behave as expected and comply with the predefined measures (Robert, 2016; Ouchi, 1979). Monitoring can have crucial implications on the actions of users' being monitored: for instance, it is shown that monitoring, i.e., when someone observes users' actions, leads to effective communication and coordination among team members (Dabbish and Kraut, 2008), improves team performance (Marks and Panzer, 2004), and increases trust among users (Robert, 2016).

Similarly, in online learning settings, the practice of monitoring learners' activity can have crucial implications on learners' progress, performance, and experience. Tools that explicitly capture, integrate, and propagate measures about learners' state of learning can facilitate the practice of monitoring. From this viewpoint, monitoring can be performed in two main ways: through self-monitoring, learners monitor themselves with the aim of encouraging self-regulation, reflection and metacognition (Loboda et al., 2014); and through internal (see teachers) or external (see advisors) monitoring, for example, a teacher can set objectives for learners and monitor whether learners are progressing towards achieving those objectives (Florian-Gaviria et al., 2013), a teacher can monitor learners' engagement with the learning material to assist learners in a timely fashion (Aslan et al., 2019). At workspace level, monitoring can be: synchronous, by providing real-time information to underlying stakeholders; asynchronous, by providing post-hoc information; and/or proactive, by providing anticipated information (Rodriguez-Triana et al., 2016).

2.4.1 Proactive Monitoring

Proactive monitoring can be handy at institution level, or in major courses, see University, MOOC, where monitoring of individual (or groups of) learners could be difficult or even unpractical. In this respect, teachers and/or advisors lay down rules for monitoring in order to identify and engage with learners who fail to comply with those rules: not meeting goals, struggling, inactive, or at-risk (Picciano, 2012).

For example, the Northern Arizona University has developed Grade Performance Status (GPS), an academic alert tool designed to enhance teacher-learner monitoring by taking teachers' inputs and turning them into personalized emails sent to learners (Star and Collette, 2010), see Figure 2.5.D. Similarly, Course Signals at Purdue University provides early actionable information, feedback, and interventions to learners, teachers, and advisors about learners' progression by aggregating data from virtual learning environments, grade performance, and outside class help-seeking information, e.g., checkpoint sessions with teachers or advisors (Arnold, 2010). Teach-

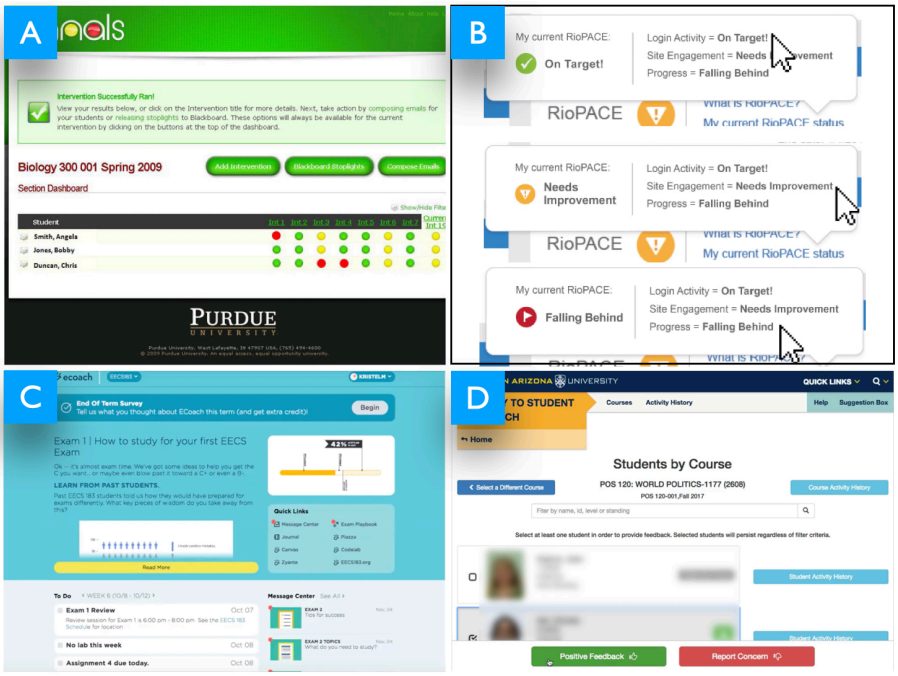


Figure 2.5: Examples of proactive monitoring dashboards. (A) Signals – Arnold (2010); (B) PACE – Grush (2011); (C) ECoach – McKay et al. (2012); (D) GPS – Star and Collette (2010).

ers configure thresholds to categorize learners' progress and implement different strategies to engage with learners, e.g., signaling, traffic alike (green, yellow, and red) flags, emails, text messages, referrals to academic advisors, and face-to-face meetings, see Figure 2.5.A.

Unlike Signals, where teachers set up thresholds for alerts, Rio Salado College developed a Progress and Course Engagement (PACE) system, which predicts learners' online progress using behavioral-data (e.g., course login), online engagement (e.g., course enrollment), and progress (e.g., materials reading, course scores). PACE uses similar encoding to Signals, showing green, yellow, and red flags so that teachers can identify whether a learner is on target, falling behind, or needs help (Grush, 2011), see Figure 2.5.B. To keep track and provide tailored feedback to learners in major courses, E-Coach (McKay et al., 2012) —at the University of Michigan uses data from surveys to provide progress, grade, performance, and feedback visualizations. Teachers can reach learners using todo lists, emails, text messages, prepare and plan exams, and send motivational tips to learners, see Figure 2.5.C.

The main goal of proactive monitoring is to allow teachers, advisors, and staff-members to be proactive about learners' progression by identifying at-risk learners before they get to a critical point. Learners identified as at-risk get follow-up interventions.

2.4.2 Classroom Monitoring

Technology-mediated learning can accommodate teaching and learning for individual learners with different pacing, style of learning, and level of acquisition. Intelligent tutoring systems enable learners to work in a self-paced manner to learn specific concepts, e.g., maths, grammar, while also continuously adapting content and providing feedback to learners while they are learning (S. D'Mello and A. Graesser, 2012; Molenaar and Campen, 2017). In particular, this releases teachers, allowing them to target specific learners need help to provide them with personalized aid.

However, such technologies are not without limitations due to the complex nature of human intelligence and a variety of technical reasons. To be more precise, learners get stuck, try to game the system, abuse hints, avoid hints, get bored, or even get off-task (Baker, 2016). Also, in part, assessing and understanding learners' activity when using such a technology is challenging, i.e., a teacher can, for instance, rely on one-to-one talks with learners, close monitoring, or direct watching of a screen of a learner to assess the state of learning. However, such approaches are cumbersome and time-consuming, as teachers have to target every learner in a classroom (Holstein et al., 2017).

Prior research has shown the benefits of augmenting ITS with fine grain monitoring dashboards, specifically to support teachers when informing timely interventions. For instance, monitoring has been directed at understanding learners' progression while they are learning maths concepts using an ITS (Molenaar and Campen, 2017), see Figure 2.6.C. Monitoring has been tailored to track learners' engagement when using an ITS, namely learners' use of hints, difficulties, facial expressions, and contextual performance to distill learners' emotional state, e.g., satisfied, confused, bored; learning status, e.g., on-task, off-task (Aslan et al., 2019), see Figure 2.6.B. Also, monitoring has been designed to track learners' struggles, performance to identify ill-suited situations: hint abuse, avoidance, off-task, rapid attempt (Holstein et al., 2018), see Figure 2.6.A. Similarly, monitoring has been investigated to enable teachers to track learners' difficulties and differences when learning the C programming language (Fu et al., 2017), see Figure 2.6.D.

2.4.3 Retrospective Monitoring

Technology-mediated learning transforms the way teaching and learning are delivered to learners where learning is no longer bounded with physical classrooms. Since learning occurs remotely, learners' progression,

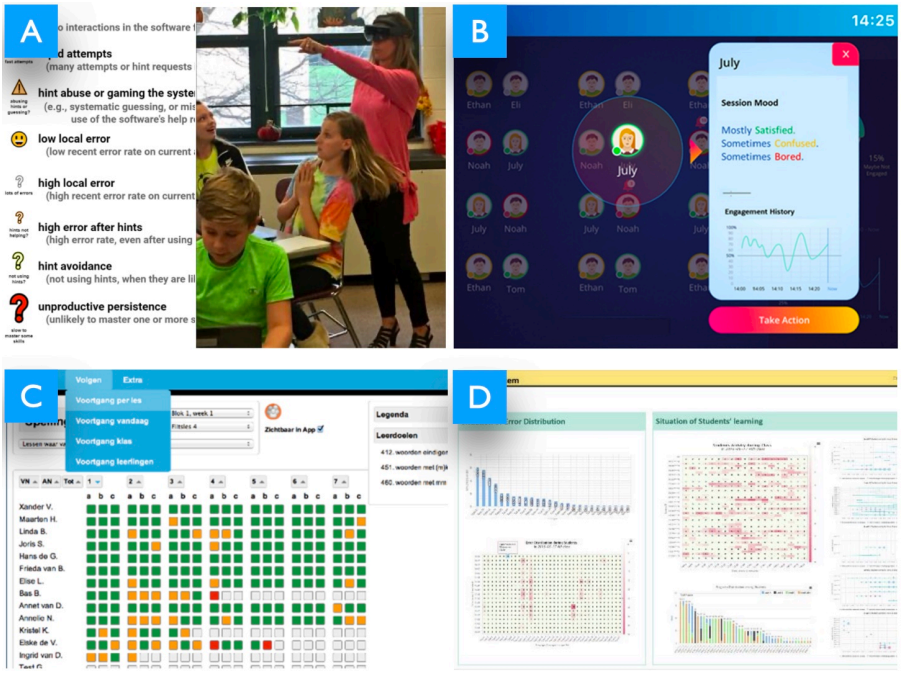


Figure 2.6: Examples of classroom monitoring dashboards. (A) Holstein et al. (2018); (B) Aslan et al. (2019); (C) Molenaar and Campen (2017); (D) Fu et al. (2017).

struggles, and engagement with learning content online become difficult for teachers to follow (Turnbull et al., 2019).

Dashboards have been aimed at facilitating teachers' retrospective monitoring of learners' engagement and acquisition when using online learning materials. For example, monitoring has been designed to track learners' interactions with learning materials to allow teachers to estimate the workload of different materials, such as the most time-consuming exercises, popular materials (Govaerts et al., 2011). Monitoring has been aimed at providing teachers with a set of indicators about learners' skills mastery, progression, time spent, and misconceptions to enable planning and conducting of personalized lessons (Xhakaj et al., 2017). Monitoring has been aimed at assessing where learners get stuck, and at identifying similar problems causing struggles, and common learning strategies in problem-solving environments (M. W. Johnson et al., 2013). Also, monitoring has been investigated to assess learners' performance and competence acquisition with respect to specific levels set by a teacher (Florian-Gaviria et al., 2013).

Furthermore, monitoring has been investigated to facilitate assessment of learners' behavior in online settings. For instance, monitoring has been designed to assist teachers in conducting assessments of learners' pathways

in a self-regulated social network learning environment, e.g., blogs, pages, comments, bookmarks, files (Groba et al., 2014). Monitoring has been directed to allow teachers to understand learners' actions and engagement when engaged in mind mapping learning activities (Carrillo et al., 2017). Similarly, monitoring has been directed to enable teachers to identify tendencies in their classes, and learners who need special attention based on learners' access to online courses, forum discussions, and quizzes (Mazza and Milani, 2005; Mazza and Dimitrova, 2007).

2.4.4 Summary Notes

We build upon this growing body of research, investigating the design of information-driven tools for monitoring, and hope to contribute in two ways.

First, although a large body of research (and reviews) has shown interest in the design of dashboards to support monitoring (see reviews, Schwendimann et al., 2017; Verbert et al., 2014), we found there to be a lack of guidance when designing around the concept of monitoring in dashboards. Indeed, the current conceptual definitions of monitoring are limited to: *"tracking learners' activities and outcomes"* (see a review, Rodriguez-Triana et al., 2016), and *"act of observing the actions of others to determine whether they comply with a predefined set of behaviors"* (Robert, 2016). This could be interpreted as a result of an established consensus among researchers and designers regarding what monitoring is. However, as underlined by different studies, monitoring seems to be a complex construct (Robert, 2016; Kirsch, 1996). There is little insight into what information to provide teachers with to support monitoring e.g., summative, formative, behavioral, outcomes; and how to design around, e.g., exploratory, explanatory.

Second, a great deal of work on dashboards designed for monitoring goes into describing different artifacts, while this body of research demonstrates key facets of what teachers are *supposed* or *can* do with dashboards when monitoring learners. Little research is directed into investigating how monitoring —through use of dashboards, moderates teachers' and learners' practices in real-world settings (Schwendimann et al., 2017; Bodily et al., 2018).

We aim at providing empirical insights into both the design and use of dashboards to support monitoring. Here, we frame monitoring through the use of teachers' dashboards as the practice of visualizing detailed status updates or audit trails about learners' behaviors, process, progress, and/or skills to enable teachers and learners to see what is going on" and gain

insights into engaging in informed and timely actions (Robert, 2016; Few, 2006). In the context of this dissertation, we design and conduct two studies to support teachers in monitoring learners' emotional state. Only a few studies have focused on the impact of learners' emotion on the pedagogical practices in educational settings (e.g., Ruiz et al., 2016). We thus adopted a top-down approach to reflect and relate what we have learned as to the design and use of monitoring dashboards with respect to current literature.

2.5 Fostering Sociotechnical Practices

From early on, research into human-computer interactions has recognized the importance of considering sociotechnical factors, either by integrating human factors into the design of interactive systems, or by capturing such factors in the surrounding of human experience when using technologies; such as affect (Picard, 1999), culture (Clemmensen and Röse, 2010), spirituality (Wyche et al., 2008), awareness (Dourish and Bellotti, 1992), reflection (Schön, 1983). Over time, specifically, awareness, reflection and socio-affect, have gained great attention in different literatures, e.g., computer-supported cooperative work/learning, and, in part, in different domains as such quantified-self, health care, education, and interactive design (Baumer et al., 2014; Desmet, 2002). Dashboards can do more than support the activity of monitoring by fostering the human social factor. In fact, they can give meaning to everyday practice.

2.5.1 Facilitating Awareness of Individual and Group Activity

In physical learning spaces, for instance, in a classroom, a teacher as well as learners develop an understanding about what and how the other is doing, e.g., behavioral actions, emotional state, locations; through natural cues of face-to-face human interactions, e.g., gestures, facial expressions, direct contact, dialog, and feedback. In computer-supported cooperative work/learning, awareness plays a central role in understanding what and how others are doing in context (Dourish and Bellotti, 1992; Gutwin et al., 1995).

From a very early stage, awareness was investigated in cooperative work, specifically, in the context of "media space" tools, where awareness was defined as an *"understanding of the activities of others, which provides a context for your own activity"* (Dourish and Bellotti, 1992, p. 1). This definition entails two main facts for awareness: knowledge and evaluation. Knowl-

edge involves users' realization of the implications of their contributions in a shared context, while evaluation involves users' assertion of their own actions with respect to the objectives of the group as a whole. Tools are considered to be effective in facilitating awareness when they provide users with suitable amounts of cues, contexts, and representations about the setting to be aware of, without overwhelming or distracting information (Schmidt, 2002). Awareness can have a crucial implication on users' knowledge and engagement, for instance, it is shown that awareness, i.e., when cues about users' activity —in a shared context, are captured and reflected back to them, facilitated coordination and communication among users (Dourish and Bellotti, 1992; Buder, 2011), supported collaborative virtual learning among separated learners (Gutwin et al., 1995), and facilitated orchestration of classroom activities (Alavi et al., 2009; Dillenbourg et al., 2011).

In this regard, learning dashboards can facilitate awareness of teachers' and learners' online activities through cues that are captured and made available to them. Users, then, can *seamlessly monitor* such cues, i.e., visualizing, reading, observing, so as to gain an understanding and ascertain the state of the setting to be aware of. For instance, in synchronous settings, awareness cues can improve group dynamics during collaboration through immediate feedback —information about users' interactions made available to the group in real-time (Tausczik and Pennebaker, 2013; Calacci et al., 2016). Similarly, dashboards can also improve teachers' awareness of social interactions among learners online: social participation in forum discussions, group work, chats (Groba et al., 2014; Yasar and Karadeniz, 2019); they can foster teachers' awareness of learners' emotions or the emotional climate of the classroom with respect to different learning activities (Ruiz et al., 2016; Aslan et al., 2019); teachers' awareness of cognitive aspects of learners when using online materials: struggles, workloads, time spent (Dyckhoff et al., 2012; Govaerts et al., 2011); teachers' awareness of ill-suited behavioral interactions of learners' online: gaming materials, rapid attempts, hint avoidance or abuse (Holstein et al., 2018; Mazza and Dimitrova, 2007). These are examples of teachers' awareness of the status and progress of learners when they are engaged with online materials through the use of dashboards.

2.5.2 Fostering Socio-affective Relationships

Over the last decade, researchers in the educational field have showed an interest in emotions. Several studies have shown that emotions have a strong effect on people's experiences and well-being (Pekrun, 2006; Boekaerts, 2010; S. K. D'Mello, 2017), and underscore a key role of emotions in learning, regulation processes, as well as a strong link to motivation and achieve-

ment (Boekaerts, 2010; Pekrun, 2006; Mega et al., 2014; Lavoué et al., 2015). In fact, emotional information can provide cues to enhance learners' learning (Linnenbrink-Garcia et al., 2016). S. K. D'Mello (2017) further argues that "*given the central role of emotions in learning, attempts to analyze (or data mine) learning without considering emotion will be incomplete*". That is, learning is more likely to be successful when teachers, as well as learning platforms, help minimize negative emotions (e.g., confusion, boredom, fear, frustration, stress) and facilitate positive ones (e.g., happiness, enjoyment, pride). Designing tools that foster positive emotional awareness among teachers and learners is a key issue for information-driven learning communities (Bouvier et al., 2014). Emotion awareness refers to the ability to perceive, identify, and understand emotions (Boden and Thompson, 2015), which can be achieved through tools that capture and make available such cues. Much work on tools and environments —in learning settings, which captures and integrates learners' emotions, can be grouped into two separate strands.

The first strand aims at enhancing learning through the design and use of "emotion-aware" learning technologies, which detect and respond to learners' emotional states by providing contextual feedback or intervening in the learning process S. D'Mello (2013). For example, Affective AutoTutor is an intelligent tutoring system that integrates learners' emotional state alongside the cognitive state while they are learning physics, so as to engage them with the materials being learnt through human-like interactions (S. D'Mello and A. Graesser, 2012).

The second strand, which is directly connected to the work in this thesis, aims at capturing and reflecting back emotion cues to users to support emotional awareness and human-informed interventions. Recent research is still emerging on this topic. Dashboards have been directed to capture and present learners' self-reported emotions to support their emotional awareness(Ruiz et al., 2016). Rather than using subjective (or self-reported) emotions, other dashboards have proposed an automated approach to infer learners' emotions online, for instance, using textual cues (Montero and Suhonen, 2014), learners' action logs (Leony et al., 2013), and video cues (GhasemAghaei et al., 2016).

2.5.3 Supporting Teachers' and Learners' Reflection

Reflection in interactive design is informed from different domains: philosophy, cognition, psychology, and education (Baumer, 2015). Schön, for instance, defined reflection as the practice by which we gain evidence from our experiences and become aware of our implicit knowledge (Schön, 1983). In light of Schön's definition, reflection can be seen as a process of

introspection and analysis in which users generate, question, and interpret different aspects of their activities, and project that understanding into the future (Dourish, 2004a). From this perspective, reflection can materialize through both cognitive practices and social interactions (Mols et al., 2016a).

The practice of reflection can happen naturally, for instance, during moments of introspection and/or socialization, such as talks with a friend. However, for most people, reflection is initiated through (external) contextual cues and factors, mainly through use of technology —also referred to as information-driven reflection, or through concrete verbal or written dialogs, e.g., discussion with a mentor, a psychologist —also referred to as dialog-driven reflection, (Mols et al., 2016b). The former case is directly relevant to our work as it deals with users' reflection through use of data-driven tools. In fact, as stated by Mols et al. (2016b), technology plays three roles for reflection: "triggering" provokes a reason to initiate reflection, e.g., on some content or information, historical events; "supporting" facilitates the execution of the reflection process, e.g., allowing users to explore and inspect information on their activities from different perspectives; and "capturing" facilitates externalization of reflections, e.g., through use of diaries to document moments of reflections; together, these roles bridge the process of reflection of everyday practice, or as stated, assist in: *"remembering and analyzing past, present and future experiences in order to reassess our perceptions, beliefs, feelings, and actions regarding our everyday life"* (Mols et al., 2016b). Reflection, then, facilitates inspections of information of interest to produce a deep understanding and knowledge and better actions, for instance, better behavioral change and well-being (Mols et al., 2016a), reinforcement of learning (D. Norman, 1993).

Dashboards by their very nature are information-driven tools. This, in part, leverages human exploration and interpretation of the information presented on dashboards. As a result, dashboards could provide great space for reflection. For instance, through dashboards a teacher might be able to explore and assess learners' states of learning from different perspectives: behaviors, progress, process, and skills, which can then "trigger" reflection and "support" its process (Mols et al., 2016b). Few papers have reported how the dashboard supported teachers' reflection, by assessing outcomes seemingly resulting from reflection, for instance, the ability to identify opportunities for interventions in a classroom (Molenaar and Campen, 2017), to identify parts that need more focus in the classroom (Chetlur et al., 2014) or parts that need to be improved to support learners' comprehension (Xhakaj et al., 2017), to understand the way teachers prepare and run their courses (Mazza and Dimitrova, 2007).

2.5.4 Summary Notes

We seek to build upon this great body of research investigating the design of information-driven tools to foster awareness and reflection, and we hope to contribute in three ways.

First, it is worth noting that the large majority of learning dashboards rely on server logs and that few papers design dedicated probes to capture and present relevant information to support awareness, emotional awareness, and reflection (Schwendimann et al., 2017). Indeed, we found there to be a need for, or as stated: "a much richer set of relevant data" to foster awareness and reflection about learning (Verbert et al., 2014; Rodriguez-Triana et al., 2016).

Second, several papers aim at supporting teachers' awareness and/or reflection of various factors including social (Groba et al., 2014; Yasar and Karadeniz, 2019), emotional (Ruiz et al., 2016; Aslan et al., 2019), and behavioral (Holstein et al., 2018) factors. However, one crucial mismatch persists between supporting awareness or reflection through dashboards, and the focus of the evaluations. Very few papers have reported on awareness and/or reflection *per se*. Indeed, "few studies actually look at how (or how much) [to what extent] awareness and/or reflection are improved or, more importantly, what are the effects of such enhancements" (refer to a review Rodriguez-Triana et al., 2016).

Third, tools are considered effective in supporting awareness and/or reflection when they provide users with the right information, at the right time, and through the right representation. That is, too little information might induce a lack of knowledge about the setting, while too much information might overwhelm, distract, and hamper ease of use of the tool. Indeed, there is little insight into what information should be provided to teachers to support awareness and/or reflection, how it can be provided, and how to design around such concepts (Rodriguez-Triana et al., 2016).

We aim at providing empirical insights into how dashboards might support teachers' awareness and understanding of individual (or a group of) learners' activities, and emotional states, as well as how dashboards might trigger and support teachers' reflection.

2.6 Evaluating Teachers' Dashboards

Approaches for evaluating teachers' dashboards are mostly informed by methods for evaluating interactive systems. New evaluation approaches, such as user experience (UX), aim at capturing both the pragmatic and hedonic aspects involving users' subjective perception and responses that result from the use and interaction with interactive tools (Bargas-Avila and Hornbæk, 2011). However, *usability* and *usefulness* are still widely adopted when evaluating interactive systems (Hornbæk, 2010).

Usability has different attributes focusing on evaluating whether a system is *"easy to use"*. Early work on usability focused on developing accurate and generable methods, for instance, by gathering systematic performance data about users' use of a tool, for example, using questionnaires, such as Questionnaire for User Interface Satisfaction (QUIS) (Chin et al., 1988), Computer Usability Satisfaction Questionnaires (J. R. Lewis, 1995), and System Usability Scale (SUS) (Brooke et al., 1996). Although important, ease of use is not the only factor to evaluate a system, or as stated, *"if ease of use was the only valid criterion, people would stick to tricycles and never try bicycles"* (Beale, 2007).

Follow-up research focused on developing formative usability evaluation using methods such as focus group, user observation, think-aloud, heuristics, cognitive and pluralistic walkthrough. Related research has looked beyond usability, yielding new concepts for evaluation of interactive systems such as, usefulness (Hornbæk, 2010) and aesthetics (J. Hartmann et al., 2007). Although there is little consensus and formal methods on how to evaluate whether a system is useful, researchers assume that usefulness depends on the real-world context in which a system is deployed and used by end-users (Macdonald, 2012).

Recently, a wide range of dashboard contributions has been introduced to instrument teachers using dashboards, spanning a variety of technical and pedagogical aspects. A notable amount of teacher-facing dashboard contributions is still system-oriented, where the paper describes the newly created dashboard artifact and highlights how teachers are supposed to (or can) use it (Mottus et al., 2014; De et al., 2010; Bakharia and Dawson, 2011; Gruzd and Conroy, 2018; Mazza and Milani, 2005; Fu et al., 2017). Other papers, discussed below, looked beyond the system and its underlying technology to evaluate the implications on teachers.

2.6.1 Evaluating Usability and Usefulness

In eLAT (Dyckhoff et al., 2012), the authors conducted controlled usability testing to evaluate the dashboard. For this purpose, four teachers were asked to perform certain tasks using the think-aloud method (C. H. Lewis, 1982), which allowed the authors to pinpoint some usability issues. In a similar vein, the authors of SOLAR (Florian-Gaviria et al., 2013) conducted a task-oriented analysis to evaluate the usability of the system. However, unlike (Dyckhoff et al., 2012) where the formative think-aloud method was used, the authors administered a questionnaire to gather systematic performance results from 20 teachers.

In InVis (M. W. Johnson et al., 2013), the authors developed 15 tasks covering different features of the tool that they combined with a usability questionnaire CSUQ (J. R. Lewis, 1995), to conduct quantitative task-based usability testing with seven participants. In SoftLearn (Groba et al., 2014), the authors used the same questionnaire, i.e., CSUQ, as in InVis (M. W. Johnson et al., 2013). However, this time they conducted an A/B testing approach where one teacher used the learning environment without the dashboard (control) and another teacher used the learning environment alongside the dashboard (treatment). It was observed from this that the dashboard was helpful in reducing the assessment time of learners' learning.

Some of the aforementioned papers looked beyond usability. For instance, the authors of SOLAR (Florian-Gaviria et al., 2013) administered a custom questionnaire to gather insights into the usefulness of their dashboard. It was observed that the tool was useful for around three quarters of 20 teachers, to identify patterns of failure and success in learners' learning, to trigger reflection and awareness, and to understand learners' social behavioral aspects. In InVis (M. W. Johnson et al., 2013), the authors conducted a set of case studies with expert teachers, to show how the tool was useful for developing insights into learners' common strategies and mistakes in problem-solving environments. Similarly, the authors of SAM (Govaerts et al., 2011) conducted two case studies with 20 teachers using questionnaires, where it was found that the dashboard helped teachers to identify outlier learners that needed assistance, and to understand the workload of the course on learners online.

2.6.2 Evaluating Pedagogical Practices

Another perspective in the evaluation of teachers' dashboards concerns deploying the dashboard for teachers, to gather insights into the pedagogical

practices that the tool might enable or limit in real-world settings. Research into teachers' data-informed practices using dashboards focuses mainly on how to support monitoring of learners' performance while they are using a learning environment in-class, such as errors, struggles, attempts, responses, and engagement indicators.

Combining field observations and interviews, Molenaar and Campen (2017) looked into how 38 teachers used a dashboard over one session (50 minutes) to conduct assessments both at class and learner levels (errors, progression, etc.), to provide motivational and formative instruction to learners, as well as to adjust both the learning material and the pace of the lesson. Although the results revealed a diversity of use (low, medium, high), three-quarters of teachers' dashboard consultations were followed by motivational feedback or formative instruction, either directed to the class as a whole or to individual learners with medium or high ability. In a similar study, Holstein et al. (2018) investigated how 8 teachers used a virtual reality glass dashboard over a total of 18 courses (30 minutes each) to orchestrate the classroom. In contrast to (Molenaar and Campen, 2017), the authors found that the dashboard tended to divert teachers' attention and time towards learners of lower prior ability.

Combining observations and interviews, Aslan et al. (2019) looked into how a teacher used indicators of learners' engagement shown on a dashboard over two courses (80 minutes each) to support learners through different teachers' interventions: verbal warning, positive reinforcement, scaffolding, close monitoring. The authors observed a medium-sized effect of the dashboard on teachers' scaffolding interventions and use for close monitoring, mainly in assisting the teacher by allocating time to learners who need most support without compromising the overall experience of the class. In a similar study, Roberto Martinez-Maldonado et al. (2015) investigated how three teachers used real-time notifications on learners' actions and their learning artifacts, shown on a dashboard. The dashboard was found to be helpful for teachers in orchestrating multi-tabletop learning environments and providing feedback to learners.

In a similar vein, Groba et al. (2014) proposed an analytics-based tool to help teachers assess learners' activity, and evaluated it with two teachers during a one-session course. The results showed that this tool reduced assessment times significantly and helped teachers understand the learning process of learners. Combining think-aloud and observations, Xhakaj et al. (2017) investigated how five teachers used a dashboard to plan and conduct lessons over eight courses (60 minutes each), where students used a virtual learning environment. The authors found that the provision of both positive and negative notifications on learners' status facilitated the dialog between a teacher and learners, and supported teachers in adjust-

ing and orchestrating learning materials. The authors also observed that the dashboard helped teachers update (confirm, reject, and/or add new items) their knowledge about both the class as a whole and individual learners. In particular, the authors found that teachers mostly integrated the knowledge acquired as to where learners struggle in their lessons.

2.6.3 Summary Notes

We seek to build upon this body of literature, focusing specifically on how dashboards can inform teachers' pedagogical practices in authentic settings, and we hope to contribute in three main ways. First, although the aforementioned studies shed light on the usefulness of some aspects of the design and intelligibility of dashboards in assisting teachers, they mainly focus on evaluating technical aspects of the underlying systems (Schwendimann et al., 2017). Much work on dashboards focuses mainly on implementation of interoperable architectures (Dyckhoff et al., 2012), development through toolkits (Kitto et al., 2015), data aggregation from multiple learning platforms (Mottus et al., 2014), personalization of views (Dyckhoff et al., 2012; Mottus et al., 2014; Florian-Gaviria et al., 2013), usability of the dashboard (Chetlur et al., 2014; Dyckhoff et al., 2012; Ez-zaouia and Lavoué, 2017), usefulness of the dashboard (Florian-Gaviria et al., 2013; Dyckhoff et al., 2012), or interoperability of the dashboard with multiple platforms (Dyckhoff et al., 2012). Also, in many cases the evaluation was conducted in controlled settings or even in the lab (Mazza and Dimitrova, 2007); there is little evidence about the impact of dashboards on teachers' practices in-situ.

Second, there are few notable studies that show that dashboards —when they are well crafted, can successfully assist teachers in monitoring learners and conducting informed interventions (Aslan et al., 2019; Holstein et al., 2017; Roberto Martinez-Maldonado et al., 2015; Molenaar and Campen, 2017). However, the studies that do exist focus almost entirely on specific intelligent tutoring systems used mostly by teachers synchronously inside the classroom while learners are using the virtual learning environment. We aim at extending our understanding of how a dashboard can inform teachers' practices in the wider perspective of web-based online learning settings.

Third, very few studies have focused on teachers' routines surrounding such tools (Greiffenhagen, 2011), or on how they help teachers inform their pedagogical practices. Also, such studies are often conducted in controlled settings, e.g., simulated classrooms (Aslan et al., 2019; Holstein et al., 2017) and over a short-term deployment. In contrast, we aim to examine how teachers make use of dashboards in their everyday practices,

in other words, not only what teachers are supposed to do with such tools, but more importantly what teachers actually *do* do. To that end, we designed and conducted two of the first longitudinal studies, in-the-wild, to investigate how a dashboard might impact teachers' pedagogical practices in the long run in authentic settings.

2.7 Dashboard Design and Use

Dashboard design can be seen as a *process* that includes requirement gathering, ideation, development, and evaluation. In the design process, we link research questions of interest with design ideas, act on those ideas to generate design approximations, and refine those approximations to uncover a solution that meets end-users' needs.

Learning communities have demonstrated significant interest and adoption of dashboards to support a wide range of aspects: data sense-making, communication, learning, collaboration, awareness, reflection (Alper Sarikaya et al., 2018; Verbert et al., 2014). While these contributions highlight an interesting, challenging, and interdisciplinary area of research, there continues to be a lack of guidance in grounding the design of dashboards to support such complex design goals. In fact, the literature on design-based research and practices of design, analysis, and evaluation of teachers' dashboards is unfortunately very scarce, with only very few exceptions. In two recent systematic reviews of more than 150 learning dashboards, almost half of the surveyed papers do not conduct any evaluations or report on conducting a specific —or using an existing, design process (Schwendimann et al., 2017; Bodily et al., 2018). The first welcome exception is the framework proposed by Verbert et al. (2014) to guide the analysis of learning analytics dashboards. Although the framework is an excellent thinking tool, to evaluate the impact of a dashboard (e.g., see Molenaar and Campen, 2017), it mainly captures the evaluation part, and fails to provide a full model of *how* to design dashboards, guiding the whole process from domain characterization to evaluation. Another welcome exception is the four-stage workflow (problem identification, low-fidelity prototyping, high-fidelity prototyping, pilot studies) by Martinez-Maldonaldo et al. (2015), which aims at guiding the design and deployment of awareness tools for instructors and learners. However, the workflow does not shed light on the principles of dashboard design —as information-driven visualization tools, and nor does it tackle the sociotechnical factors, pedagogical practice, or design tensions surrounding dashboards.

Building upon this body of design-based research and practices of design, analysis, and evaluation of teachers' dashboards, we hope to make one major contribution.

First, research into learning dashboards continues to be somewhat disjointed, demonstrating different types of dashboards and spanning various applications, such as understanding collaboration, forum discussions, facilitating monitoring, fostering awareness and reflection, informing interventions (Rodriguez-Triana et al., 2016). Dashboards are, in fact, special types of information-driven visualization tools involving complex and ill-defined tasks: data sense-making, monitoring, awareness, exploration, explanation (Verbert et al., 2014); they still need to foster learning (Ferguson, 2012; Chatti et al., 2012); they need to fit different users who inevitably have different needs and contexts of use (Schwendimann et al., 2017; Vieira et al., 2018), as well as different visual, analytical, and data literacy (Alper Sarikaya et al., 2018; Barbara Wasson, 2015).

Second, in addition to the design challenges: ill-defined and complex tasks, different contexts of use, users with diverse data, visual, and analytical literacy; learning dashboards involve profound sociotechnical factors: reflection, transparency, control, trust, reliance; and pedagogical practices, such as coaching, mentoring, guiding, intervening, e.g., feedback, debrief, scaffold, lesson. A conceptual grounding of sociotechnical factors, pedagogical practices, and design challenges might prove useful in understanding how best to design around dashboards for teachers and what tensions might exist with particular designs (Schwendimann et al., 2017; Vieira et al., 2018; McKenney and Kali, 2017).

Therefore, articulating a larger conceptual umbrella under which research surrounding learning dashboards operates, in terms of *what* (decisions), *why* (rationales), and *how* (process) to structure and guide dashboard design, may provide a holistic view of the design process, the design implications, and the design space and facilitate a deep consideration and/or engagement with the important sociotechnical factors and pedagogical practices underpinned by teachers' dashboard design and use, in a comprehensive and methodological way.

2.8 Summary

Analysis of prior work on teachers' dashboards suggests there are three main issues, which corroborates the results of recent reviews on learning dashboards:

First, from a **design perspective**, there is a growing need for design-based research centered around needs of teachers in-situ (Shum et al., 2019), to best facilitate pedagogical practices through dashboards that are more than simple log-driven summarization metrics (Verbert et al., 2014; Schwendimann et al., 2017; Bodily et al., 2018).

Second, from an **empirical perspective**, there is a need for longitudinal studies to understand how teachers appropriate, leverage, and shape their pedagogical practices around dashboards in the long run: in other words, not only what teachers are supposed to do or can do with dashboards, but more importantly what they actually *do* do (Verbert et al., 2014; Schwendimann et al., 2017; Bodily et al., 2018).

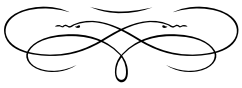
Finally, from a **methodological perspective**, much work on teachers' dashboards continues to be somewhat disjointed. There is still a need for conceptual studies to facilitate consideration and engagement with important, while complex, design challenges, sociotechnical factors, and pedagogical practices underpinned by teachers' dashboard design and use (Schwendimann et al., 2017; McKenney and Kali, 2017; Vieira et al., 2018).

We first took a bottom-up approach by working with underlying stakeholders to prototype and study teachers' dashboards in context. We present, in chapter 4 and chapter 5, two longitudinal studies of dashboards conducted in authentic settings to assist teachers in the context of online language learning.

We then took a top-down approach to structure, discuss, and reflect on our studies into teachers' dashboards to relate what we have learned to the design and evaluation of teachers' dashboards with respect to current research. We then discuss and present in chapter 6 a conceptual umbrella under which research surrounding learning dashboard design and use may operate.

Chapter 3

Multimodal Emotion Awareness in Video-conferencing Learning Settings



> *Not all computers need to pay attention to emotions, or to have emotional abilities. Some machines are useful as rigid tools, and it is fine to keep them that way. However, there are situations where the human-machine interaction could be improved by having machines naturally adapt to their users, and where communication about when, where, how, and how important it is to adapt involves emotional information, possibly including expressions of frustration, confusion, disliking, interest, and more.*
>
> —Picard, 1999, p. 829

3.1 Overview

Emotion plays an important role in learning experiences, processes, and outcomes. Capturing emotional *awareness cues* about learners and making them available to teachers may prove useful in fostering sustainable socio-affective relationships. Such an enterprise, is, however, challenging given the complex nature of human emotions; it would be more complete to rely on different modalities to capture, quantify, and integrate cues about learners' emotional states. We present a case study investigating a multimodal approach or applying two theoretical models of emotion: discrete and dimensional and for using low-cost commercial cloud APIs for emotion recognition when inferring learners' emotions in online learning environments. We demonstrated this approach in the context of video-conferencing language learning, through the use of audio and video streams when inferring automated emotions along with learners' self-reported emotions, as well as contextual teachers' and learners' interaction traces on the platform. In addition to proving the feasibility of using cloud APIs for emotion recognition in online learning environments, we analyze 10 learning sessions of 45 minutes each —more than 7 hours' recording in total, to compare emotions from two modalities: audio and video. We propose an approach for combining discrete and dimensional emotional cues to capture learners' emotional states. Our findings highlight the pertinence of the multimodal approach for inferring emotion in distance learning, where, in particular, different emotions are detected from different sources, i.e., audio and video, during such settings.

3.2 Motivation

Designing online learning environments that are enjoyable, motivating, and inspiring for teachers as well as learners is a key issue for learning communities (Zembylas et al., 2008; O'regan, 2003). Emotion plays a vital role in our social interactions. Learning is more likely to be successful when teachers as well as learning environments help foster positive emotional experiences (Picard, 1999). Several studies have shown a significant effect of emotions on learning experiences, processes and outcomes (Pekrun, 2014). Positive emotions, and particularly task-related ones, such as happiness, curiosity, and enjoyment, can positively impact learning as they help learners focus more on the task, facilitate self-regulation, and foster learners' motivation to improve and learn (Wolters, 2003); negative emotions, such as frustration, stress, or boredom, can have detrimental effects as they divert learners' attention from problem-solving, consume cognitive resources, and inhibit performance particularly as concerns complex learning tasks (Pekrun, Goetz, et al., 2011; S. K. D'Mello, 2017).

Emotion awareness refers to the ability to perceive, identify, and understand emotions (Boden and Thompson, 2015). Information-driven tools help foster awareness of emotion, in online learning settings, by providing *awareness cues* —information that is captured and made available to underlying users. The design of such cues depends on methods for capturing, quantifying, and integrating learners' emotional states. Methods for measuring emotions can be: *objective* —inferred automatically, *subjective* —self-reported by users themselves, *snapshot* —gathered at specific moments, e.g., before and after a learning session, or *continuous* —gathered during a learning session (Afzal and Robinson, 2010).

However, one main strand of almost all proposed information-driven tools is, in part, that they rely mainly on learners' online traces stored by log servers (Schwendimann et al., 2017). Also, in part, and more specifically in the context of emotions in learning, most of the proposed tools are based on one modality, for instance, logs (Leony et al., 2013), texts (Montero and Suhonen, 2014), self-reports (Ruiz et al., 2016), videos (GhasemAghaei et al., 2016; Happy et al., 2016). Given the complex nature of emotions (Izard, 2010), it may be more effective to rely on more than one modality to provide insights into learners' emotional states. One of the ambitious goals of learning communities is indeed to collect and use different modalities, such as using video, audio, logs, among others, to effectively triangulate, characterize, design, and study learning and teaching experiences and outcomes (Worsley and Martinez-Maldonado, 2018). There is a growing need for tools that explore richer modalities of learners' data, such as bodily signals, gestures, emotions, and more, (see reviews, Verbert et al., 2014; Schwendimann et al., 2017).

We follow a case study design (Yin, 2018) to investigate a multimodal and contextual approach of awareness cues of learners' emotion in the settings of online learning. We conduct and demonstrate the case study using a video-conferencing learning platform called SPEAKPLUS, which we describe in Section 3.4. We applied both dimensional (Barrett and Russell, 1998) and discrete (Ekman and Friesen, 1976) models for inferring automated learners' emotions along with contextual information (Dourish, 2004b). We propose to use cloud APIs for emotion recognition in a non-intrusive way. We provide an exploratory analysis comparing learners' inferred emotions, followed by an approach for unifying these two different emotion models. In addition to proving the feasibility of using heterogeneous APIs for emotion recognition in online learning environments (Schwendimann et al., 2017; Verbert et al., 2014), our findings highlight the pertinence of our multimodal approach for inferring emotion in distance learning. The results show that different emotions are detected from different sources, i.e., audio and video, in the settings of live learning sessions.

The contributions of this work are:

1. A case study of using heterogeneous APIs when inferring automated emotional awareness cues of learners' emotions from audio and video alongside self-reported emotions and contextual users' interaction traces in online/distance learning.

2. A method for unifying discrete and dimensional models of emotions.

3. A qualitative and quantitative exploratory analysis comparing learners' inferred emotions from audio and video.

3.3 Background

Incorporating human factors in the design and use of interactive systems has been researched thoroughly for decades (e.g., Picard, 1999; Clemmensen and Röse, 2010; Dourish and Bellotti, 1992; Wyche et al., 2008; Schön, 1983). Affective computing (Picard, 1999), in particular, which is the focus of this chapter, aims at designing technologies that seamlessly leverage our everyday social interactions by incorporating emotional experiences in the design and use of interactive systems. However, the nature of human emotions is complex. The enterprise of designing around such a phenomenology in interactive systems is a challenging endeavor. Different intertwined approaches —from disciplines have been adopted to devise conceptual assumptions about human emotion.

3.3.1 Emotion in Education Settings

Cornelius (1996) traces assumptions about human emotions with respect to four perspectives, and claims the need for theories and studies to unpack the nature emotion. From a perceptional point of view, the **Darwinian** perspective relates emotion to adaptive responses of human *expressions*, while the **Jamesian** perspective relates emotion to bodily changes of human *experiences*; both perspectives describe emotion as a construct that is more or less common and similar to all humans and that evolves to help humans survive through such appropriate emergency reactions in everyday environmental human interactions. The **Cognitive** perspective relates emotion to human thoughts and particularly human appraisal of things in context, namely the process of judging the perceived events as bad or good, and so on. The **Socio-constructivist** perspective relates emotion to human social-interactions and cultural rules.

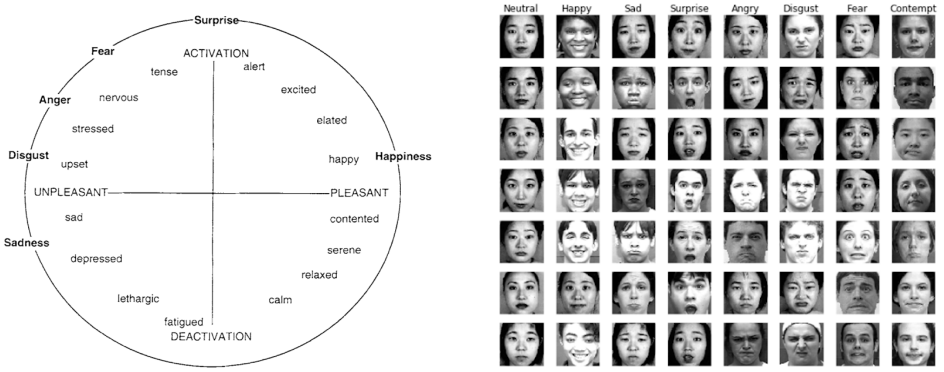
Empirical studies have shown a strong interplay between emotions and learning. Multiple parts of learning involve emotional contingencies and experiences. The taxonomy of educational objectives proposed by Bloom's and his colleagues was one of the first works in education to articulate, alongside cognition and psychomotor, the "affective" construct; or as described: "changes in interest, attitudes, and values, and the development of appreciations and adequate adjustment." (Bloom et al., 1956, p. 17). Plutchik provides a psychological stance describing one's emotion from three perspectives, subjective, such as, fear, joy, rage, trust, acceptance; and behavioral, such as, crying, escaping, biting; and functional, such as, protection, rejection (Plutchik, 1980). Apart from such early investigations of emotions, the interest in such concepts in the learning settings is relatively recent. This can be partly attributed to the complex nature of emotions. Researchers examining the development of human emotional inner workings have resulted in extremely diverse opinions (Izard, 2010), with a categorization of more than 100 multifaceted definitions of emotion (P. R. Kleinginna and A. M. Kleinginna, 1981). The phenomenon of "emotion" is intertwined in neuroscience, psychology, sociology, and education, where the construct goes by many different names, among others, *feeling*, *affect*, *mood*, *stress* (Ketai, 1975); discrete emotions: *anger*, *disgust*, *fear*, *happiness*, *sadness*, *surprise* (Ekman and Friesen, 1976); valence, arousal (Barrett and Russell, 1998), control (Scherer et al., 2013); pride, shame (Pekrun, Goetz, et al., 2011), confusion (Hess, 2003). Transience, prolongation, intensity are also latent attributes of emotion. Academic settings have also investigated more domain-specific taxonomies of "emotion".

Davitz (1969) work, for example, was one of the first that tried to understand emotion from a social phenomenology point of view by providing students with stimulus and asking them to describe the meaning of each stimulus using emotional labels; over 556 *emotional units* were found to be related to one's attitudes, values, and psychological changes, such as "I cry', "it's a state of release'. Pekrun, Goetz, et al. (2011) distinguish between activating (positive) and deactivating (negative) emotions in learning settings, specifically, activating emotions, such as enjoyment, pride, hope, are found positively correlated to different learners' learning strategies (e.g., elaboration) and motivation (e.g., effort, study interest); and in contrast, deactivating emotions, such as anger, anxiety, shame, boredom, are found negatively correlated to learners' learning. To help teachers react or diagnose learners' emotions, Barry Kort et al. (2001) articulated a simplified dimensional space of emotions of four quadrants; using one axis for the construction of knowledge, i.e., from unlearning to learning; and another axis for the affect, from unpleasant to pleasant. Namely, the Kart et al.'s model highlights that learners, when they start learning a new topic, are "fascinated" —quadrant I, or "motivated" —quadrant II, either; as a learner proceeds in learning, motivation might drop, however, to "frustration" — quadrant III, due to moments of struggle, confusion, and more; as a learner

persists in mastering the concept, s/he moves to a state of "hopefulness" —quadrant IV, before moving back to a positive emotional state. Barry Kort et al. (2001) four-state model highlights the cyclic nature of emotional contingencies of the learning experience.

Although such studies shed light on the role of emotions, the classification of emotions in learning settings is, however, a challenging endeavor. Current studies are almost entirely self-reported by learners themselves or subjectively labeled by researchers and teachers. Bloom et al.faced the same challenge when they tried to classify educational objectives under the umbrella of learners' *affective* states, where, or as stated: "teachers do not appear to be very clear about the learning experiences which are appropriate to these [affective] objectives. It is difficult to describe the behaviors appropriate to these objectives since the internal or covert feelings and emotions are as significant for this domain as are the overt behavioral manifestations." (Bloom et al., 1956, p. 17). Some of the challenges in classifying specific emotion for learning settings can be rooted to the psychological, sociological, and neuroscience ambivalent attributes of the phenomenology of emotion. An ongoing debate persists as to whether some learning-specific emotions, for instance, confusion, might in fact be considered as emotions or instead cognitive disequilibrium manifestations (S. K. D'Mello and A. C. Graesser, 2014; Hess, 2003).

Although no precise answer can be stated, we can instead attempt —with these materials as background, to highlight some unique facets of *what emotion is about* in education settings. First, such views of emotion underline at least two partially separate conceptual strands of emotion. One first strand is the informational aspect of emotion, which deals with the phenomenon of emotion as information that can be captured, quantified, encoded, and represented. Another strand is the interactional aspect of emotion, which deals with emotion as a socio-phenomenology that occurs and develops, within and depends on situated, contextual, social-interactions in the surrounding of human experiences (Boehner et al., 2007). Second, emotion, in both the informational and interactional models, is an adaptive inter-(re)action, that is, the emotional reaction manifests in human brain and body, i.e., felt, as a response to an "emitted" signal in situated interactions (Ketai, 1975); for instance, learners' happiness is more likely to be displayed during moments of recreation with fellow learners. Third, not only as an inter-(re)action, emotions are also catalyzed during the flow of our impressions, attachments, appreciations, breakdowns, judgments towards specific things, events, situations, moments, in the environment (Csikszentmihalyi, 2009, p. 116:117); for instance, if a learner is impressed by a specific topic, emotion of enjoyment is more likely to appear, while if a teacher is unsatisfied with learners' engagements, emotion of anger is more likely to occur. Fourth, the covert and overt intensity emitted by an "experienced" (symptoms) emotion is inevitably "expressed" (signs) physiologically and



(a) Emotion in dimensional space: activation vs pleasure (Barrett and Russell, 1998)

(b) Facial expression emotions (Ekman and Friesen, 1976). Images from CK+ & Jaffe datasets

Figure 3.1: Dimensional and discrete models of emotions.

perceptionally and is considerably encodable and recognizable (Cernea and Kerren, 2015; S. K. D'Mello, 2017). Fifth, emotion has not only physiological and perceptual emitters, such as heartbeat, sweat, (e.g., fearful) face, (e.g., frozen) voice, but also cognitive and motivational impulses (Hascher, 2010; Mega et al., 2014). Finally, though not proven to be the most important, while emotion has both intrinsic and extrinsic triggers, it is mostly directed towards the external objectified word (Hascher, 2010).

3.3.2 Encoding Emotional Cues

Building upon the aforementioned assumptions, two perspectives are, however, widely adopted in the affective computing (Picard, 1999). The **discrete perspective** of emotion, as portrayed in Figure 3.1b, claims a small set of distinct emotions (e.g., anger, disgust, fear, happiness, sadness, surprise), called universal or basic emotions, which constitute the core of all humans' emotions regardless of the socio-cultural factors of the person (Ekman and Friesen, 1976). On the other hand, as portrayed in Figure 3.1a, the **dimensional perspective** of emotion claims rather a continuous space in which emotions vary along certain dimensions, such as the degree of *valence* representing the positiveness of the emotion felt —varying between positive and negative, also called *pleasure*; and the *arousal* that represents the physical response and the intensity of that emotional manifestation —varying between low and high, also called *activation*, (Barrett and Russell, 1998). It is worth noting that while other dimensions have been proposed (e.g., "control", Scherer et al., 2013), valence and arousal are still the most widely adopted dimensions. At the same time, some studies have been conducted to bridge between such models (e.g., Scherer et al., 2013).

In addition to the subjective emotions of a person that can be *self-reported*, informational approaches rely on the intensity "emitted" internally or externally by the emotional manifestation to *infer* the emotion felt by the person. To that end, two approaches are used to collect automated measurements: perception-based measurements that include all the bodily manifestations expressed by the person, such as facial, vocal, gesture, textual; and physiological measurements that include all human body responses, such as heart-beat, blood pressure, brain activity, among others (Cernea and Kerren, 2015; Montero and Suhonen, 2014). The process of encoding informational emotion is broadly structured around three steps: *feature extraction* — selecting, normalizing, filtering interesting features from raw data, *feature reduction* — reducing or sampling high dimensionality of the selected features, and *classification* —training a classifier to hopefully be able to predict emotion from raw data.

3.3.3 Multi-modal Emotion Recognition

Rienties and Rivers (2014) study distinguishes seven methods for collecting learners' emotions, namely through the use of data from virtual learning environments (i.e., , content analysis, natural language analysis, behavioral indicators, e.g., click-streams), quantitative (i.e., systematic questionnaire), qualitative (i.e., interviews), self-reported, intelligent tutoring systems (i.e., tracking students' emotions and adapting responses). S. K. D'Mello (2017) highlights the use of learning analytics and educational data mining methods to analyze learners' emotions, based on the analysis of click-stream data (e.g. sequence of clicks labelled by students to build a classifier for emotions), interaction patterns (e.g. using/labelling observations of students' interactions with computers, peers, etc. to build a classifier) and bodily signals (e.g. videos from students).

One main key factor of the collection of emotional awareness cues, which, in part, distinguish between perception and physiological-based approaches, is the degree of intrusiveness. That is, three main features are used for automatic emotion recognition in *non-intrusive* settings. Facial expression recognition can be based on two techniques: a geometric model-based technique, which relies on distinctive facial features such as the position of nose, eyes and mouth; and an appearance model-based technique that considers the face to be an array of intensity values (pixels). Thus, facial expressions can be mapped to emotions using theories such as discrete emotions (Ekman and Friesen, 1976). A variety of tools can be used for this purpose, including FaceReader (Loijens and Krips, 2019), Microsoft Emotion API (Microsoft.com, n.d.), Affdex SDK (McDuff et al., 2016) and Google Emotion API (Google.com, n.d.), among others, to recognize emotions by providing associated numerical values.

Voice or sound based features can be used too. The most commonly used features are prosodic features —pitch, intensity and first formant frequency; and spectral features —Mel-frequency spectral coefficients (MFCC) (C. Lee and Narayanan, 2005). A variety of tools can help infer voice-based emotions, including, OpenEAR (Eyben et al., 2009), and Beyond Verbal (Beyond-Verbal, n.d.), among others. Textual cues are also considered (Montero and Suhonen, 2014). We do not consider textual cues, however, because the learning activity we observe is not heavily textual-based.

Importantly, emotional cues from different modalities, such as visual and vocal, can be combined to provide a more holistic estimation of the experienced emotions. In fact, many studies (e.g. Caridakis et al., 2007; Pantic et al., 2005) show that the multimodal approach for measuring one's emotions is more accurate than the unimodal approach. Different strategies exist for such a fusion, including *feature level* fusion, where features from different modalities are combined and used to infer emotions; and *decision level* fusion, where each modality is classified to emotions separately, and the results from different modalities are combined in a final result. We particularly opted for this second strategy because the result is more explainable and comprehensible.

The primary means of communication of the teacher-learner pair in the learning environment of our demonstration is via audio-visual cues — video-conferencing. We capture both streams and apply both dimensional and discrete emotion through APIs to quantify emotions. We use Beyond Verbal (Beyond-Verbal, n.d.) to infer emotions from audio stream, and Microsoft Emotion (Microsoft.com, n.d.) to infer emotions from video steam. It is worth noting that some tools show high accuracy in detecting basic emotions, but only on specific datasets (e.g., 88% on average, see Lewinski et al., 2014). This may not be the case when classifying emotions in-the-wild. This is based on our investigations of multimodal emotional analysis —discussed in chapter 3. Also, we found that the Microsoft API, which was started as a research project (Microsoft, 2016), was reliable enough for basic emotion recognition. A paper from Microsoft research (Barsoum et al., 2016) highlights the accuracy of emotion recognition for some emotions, such as happiness (94%), surprise (86%), anger (82%), and neutral (90%), but also lays down limitations in classifying other emotions, such as sadness (67%), fear (52%), disgust (26%), or contempt (4%). Such a method should be used very carefully when dealing with learners' emotions. Our approach mainly relies on the aggregation of discrete emotions in two categories (negative and positive) as an insight into learners' emotions experienced during synchronous online interactions. Beyond Verbal API is one of the few services available on the cloud and that is clinically proven (Udave, 2018).

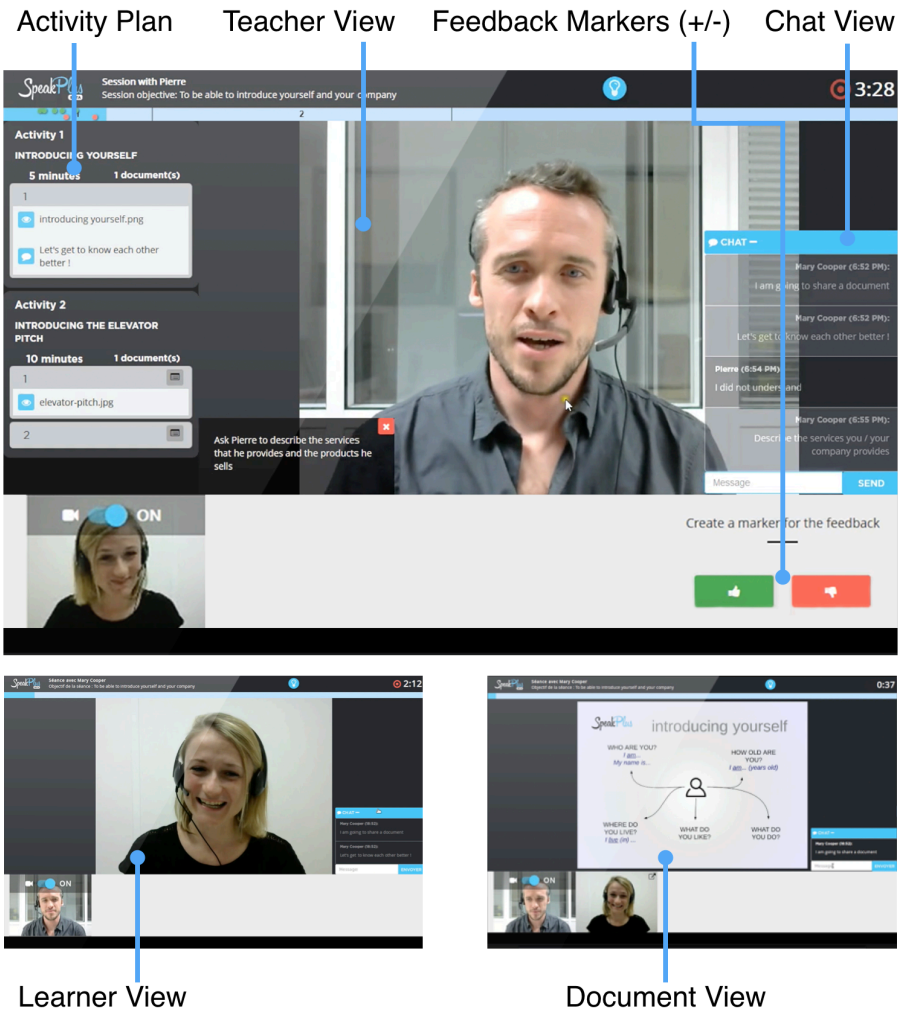


Figure 3.2: Example of a live learning session on SPEAKPLUS.

3.4 SPEAKPLUS a One-to-one Learning Platform

SPEAKPLUS is a web platform dedicated to improving oral communication skills in a foreign language (English, French, and Spanish). As Figure 3.2 portrays, the platform connects professional language teachers with learners in one-to-one remote training sessions. While the study was conducted, there were approximately 50 active teachers (teaching at least one course per month) on the platform. Teachers can join the platform if they have previous teaching experience and a degree in teaching foreign languages, and successfully pass an evaluation interview, after which they receive training to use the platform. About 2,000 users had bought at least 5 learning sessions on the platform at the time of the study. These learners had a very diverse background and skills, ranging from novices to more advanced foreign language speakers.

The learning activities on Speakplus focus on goals set by learners, for instance improving oral presentation skills or preparing a trip abroad. The learning process can be divided into three main stages:

1. Pre-session: before live learning sessions, teachers can create and customize learning materials (activities) for each learning session. Each learning activity has a duration and usually contains a set of documents, instructions for the learner or a set of personal notes for the teacher.

2. In-session: the teacher conducts a synchronous learning session with his/her learner. The pair (teacher and learner) communicates in real-time through video-conferencing. The teacher has the ability to share learning materials/activities with the learner (e.g. pdf, word, image, audio, and video), or to communicate via chat.

3. Post-session: the teacher writes a feedback report for the learner, including a summary of the session as well as a set of instructions for the next session.

During the session, teachers can add annotations (called markers), to pinpoint learners' positive and negative moments. At the end of each session, teachers use these markers to write a feedback report using a dedicated editor. Once the feedback report is finished, the teacher shares it with the learner, who is notified by email. Learners can read the feedback on the platform, alongside an audio recording of the session.

3.5 Research Approach

Our analysis is conducted in authentic and non-intrusive settings:

Data collection. We captured data using sensors already used on the platform for the learning activity, namely a webcam and a microphone for video and audio recording, respectively, as well as the interaction traces of the teacher-learner pair. As we describe in Section 3.6 below, we slightly augmented the platform to collect and process data.

Participants. One male French teacher in his/her thirties and two graduate learners in their twenties participated in this study. They were contacted by the pedagogical manager of the platform. Both learners were enrolled in the platform to train for a job interview in French language speaking. One learner was from Latin-America, and the other was from Asia.

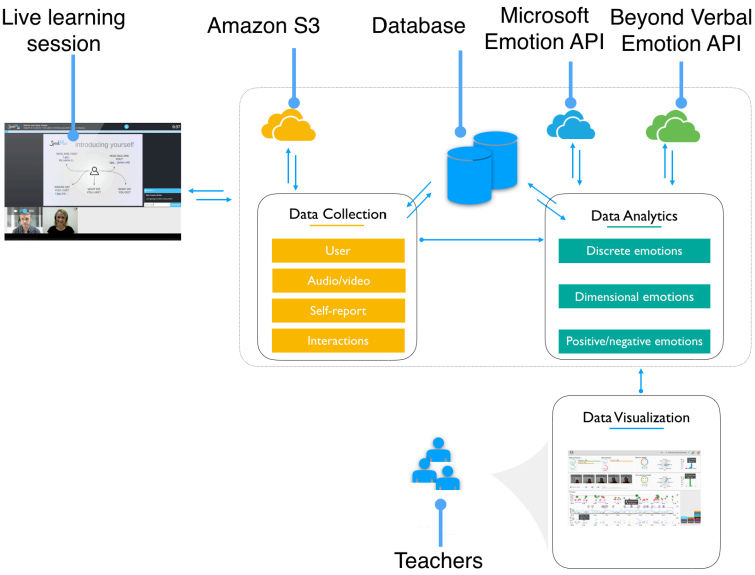


Figure 3.3: Data collection architecture.

Procedure. Each learner had five learning sessions on the learning platform with the same teacher. The learners had one learning session every week. Each session lasted 45 minutes. Data collection lasted for about two months EMODAas the students did not start their learning sessions at the same time. We collected a total of 7 hours and 30 minutes of audio/video recordings along with users' interaction traces on the platform.

3.6 Non-intrusive Emotional Cues Collection

We identified different "trackable" cues to investigate emotional awareness in the setting of SPEAKPLUS, namely four data sources: audio, video, self-report, and interaction traces. Audio and video are the primary communication channels on the platform as the learning activity aims at improving oral skills through video-conferencing sessions. Self-reported data inform on learners' subjective emotions before and after the learning session. Finally, interaction traces help contextualize the learners' emotions and may provide clues as to potential triggers behind learners' experienced emotions. Figure 3.3 portrays an overview of the architecture of the system built upon the SPEAKPLUS platform to collect these four cues during live learning sessions.

Automated emotions are inferred using two cloud APIs, which were chosen to apply both models of emotions: dimensional and discrete. The Microsoft emotion recognition Service API classifies emotions based on *facial expressions*, i.e., video, according to the discrete model, namely re-

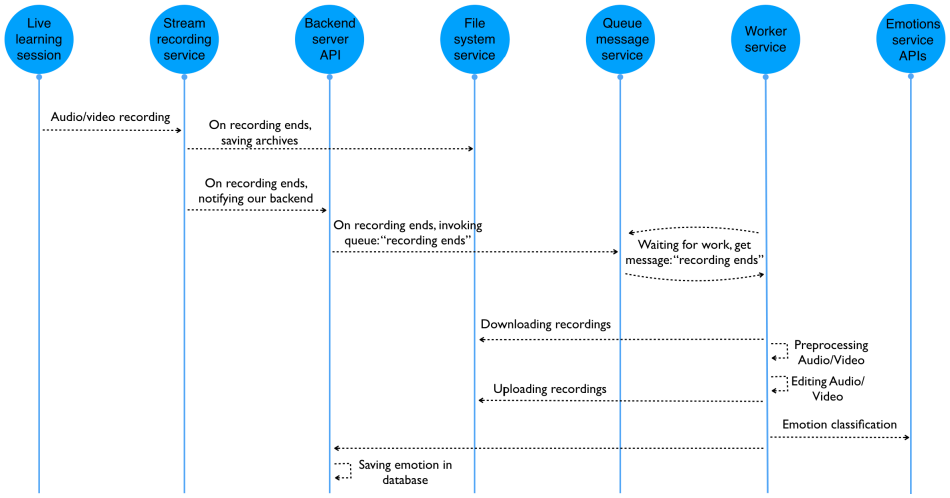


Figure 3.4: Overview of the automated process of inferring emotions.

turning levels of neutral, happiness, surprise, sadness, disgust, contempt, fear, anger. The Beyond Verbal service API classifies emotions based on *vocal expressions*, i.e., audio, according to the dimensional model, namely returning levels of arousal and valence.

Figure 3.4 portrays an overview of the automated process of inferring emotion from audio and video modalities. We store the recording of the learning sessions between the teacher and the learner (audio-video) as we go along in the cloud (Amazon S3). At the end of the session, we download the recordings and process them using the Beyond Verbal API and the Microsoft API. We store the results of classified emotions in a database, via a dedicated service for data collection, together with the contextual interaction traces and self-reported emotion data. We then pull, process, and aggregate the data from the database to build emotional awareness cues. The following subsections describe for each data source how the data are collected and stored, and how emotions are classified using the mentioned APIs.

3.6.1 Self-Reported Emotions

We built an interactive interface that enables learners to express their emotion(s) before and after the learning session —snapshot, according to both dimensional and discrete emotions. The dimensional emotion interface is inspired by the MoodMap tool proposed in (Mora et al., 2011). As Figure 3.5 portrays, we used a squared area for dimensional emotions, namely, for valence and arousal. Learners can interact with the interface to rate their levels of valence and arousal on a scale from -100 to 100. We combined four colors in a conic gradient form to represent the four lev-

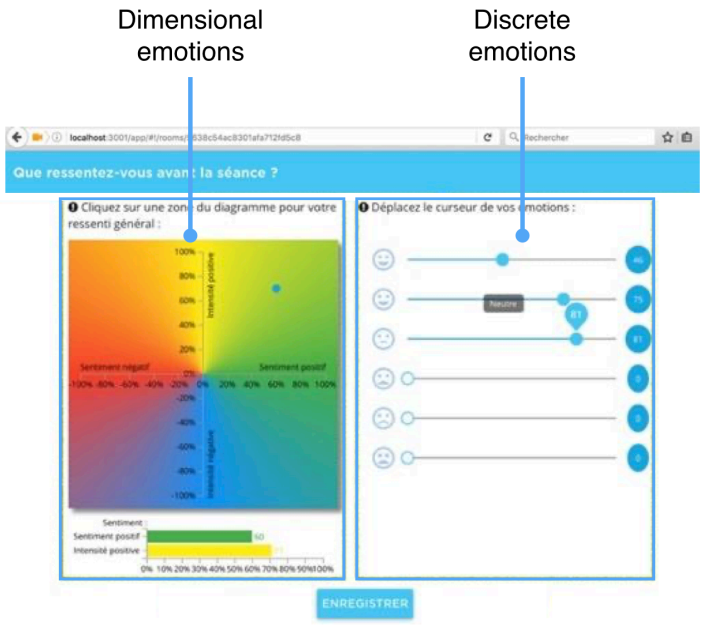


Figure 3.5: Visualization for self-reported emotions.

els: positive and negative valence by green and red colors respectively, and positive and negative arousal by blue and yellow colors respectively. We added a mouse hover event to highlight the position of user's entry. We also reflected user's entry using a mirror bar chart at the bottom of the squared area to guide the learners while they are reporting their two-dimensional emotions. Similarly, we used sliders to allow learners to rate their discrete emotions —at the moment, i.e., neutral, happiness, surprise, sadness, disgust, contempt, anger, on a continuous scale from 1 to 100. We used emoticons and a tooltip to distinguish discrete emotions. We added a mouse event to highlight the value of the rating using a marker associated with the slider and inside a bubble next to it.

We then asked learners to express their emotions both at the start and at the end of the learning session. This emotional information is complementary to the data automatically collected during the learning session. The collected self-reported data are stored in a database as illustrated in Figure 3.3.

3.6.2 Face-Based Emotions —from Video Stream

We recorded the learning sessions on the platform. Each session may have one or more archive folder(s) depending on whether or not there were any interruptions during the session. Each archive folder contains two files: one for the learner recorded stream and the other for the tutor recorded

Table 3.1: Description of an example of a Microsoft API response.

Field	Description
faceRectangle	Rectangle location of the face in the image.
scores	Rectangle location of the face in the image. Emotion scores for each face in the image; (neutral, happiness, surprise, sadness, disgust, contempt, fear, anger).
API Output Example	


```
[
 {
   "faceRectangle": {
     "left":68,
     "top":97,
     "width":64,
     "height":97
   },
   "scores":{
     "anger":0.00300731952,
     "contempt":5.14648448E-08,
     "disgust":9.180124E-06,
     "fear":0.0001912825,
     "happiness":0.9875571,
     "neutral":0.0009861537,
     "sadness":1.889955E-05,
     "surprise":0.008229999
   }
 }
]
```


stream. We used a simple algorithm based on the FFmpeg (FFmpeg, n.d.) library to concatenate the video streams from the archive folders when needed. Once the video sequence had been constructed, we then classified and stored emotions using a simple algorithm, which first uses FFmpeg to extract sample frames, i.e., one frame per second, from the video stream, and then sends HTTP requests to the Microsoft Emotion service API to classify the potential emotions of each frame. We then store the results of emotions in the database. The Microsoft Emotion service API returns specifically the face entries —found in each frame, along with the associated emotion scores. An empty response indicates that no face was detected by the API. Table 3.1 describes a typical response of the Microsoft Emotion service API.

3.6.3 Voice-Based Emotions —from Audio Stream

We used the same algorithm as for the video stream to construct the audio, i.e., voice, sequence from recordings. In the current study, we used only the learner audio channel. An HTTP request to the Beyond Verbal Emotion API returns an array of time interval entries along with their associated scores of emotions. An empty response indicates that no emotions were detected

Table 3.2: Description of an example of a Beyond Verbal API response.

Field	Description
analysisSegments	The array containing analysis segments. The offset of the segment from the beginning of the sample being analyzed (in ms).
duration	Duration of the analysis segment in the sample being analyzed.
analysis	Analysis object. Contains analysis values for the segment. The content of the object is provided as an example.
valence	Valence object score (has value and group): • Value: a value between 0 and 100. • Group: positive, neutral or negative
arousal	Arousal object score (has value and group): • Value: between 0 and 100. • Group: low, neutral or high

API Output Example	{
	"analysisSegments":[
	{
	"analysis":{
	"Arousal":{
	"Group":"low",
	"Value":"4.35"
	},
	"Valence":{
	"Group":"positive",
	"Value":"82.28"
	},
	"duration":37410,
	"offset":4576
	}
	}
	]
	}

in the audio stream by the API. Table 3.2 describes a typical response of the Beyond Verbal Emotion API. We then store the results of emotions of the API in our database as shown in the data collection architecture Figure 3.3.

3.6.4 Contextual Interaction Traces

We decided to track all pertinent actions performed by the tutor and the learner on the SPEAKPLUS environment during the learning session. This information is used to contextualize emotional awareness cues. We list in Table 3.3 the interaction traces collected and stored in real-time.

Table 3.3: Description of list of traces collected on the SPEAKPLUS platform.

Field	Description
action_name	The name of the action (event) triggered between the actors. It can be: • SHARING_DOC: the tutor starts sharing a document • STOP_SHARING_DOC: the tutor stops sharing a document • ACTIVITY_TASK: an action sent to the learner by the tutor • FREE_TEXT: message chatting • SHOW_DOC: the learner enables display of the shared document • HIDE_DOC: the learner disables display of the shared document
action_content_type	The type of event: • PDF: Pdf document sharing • AUDIO: Audio file sharing • VIDEO: Video file sharing • IMAGE: Image file sharing • TEXT: Text action sent to the learner by the tutor • SHOW_HIDE_DOC: for the SHOW_DOC or HIDE_DOC action
activity	Activity Id
item	Item Id in the activity (an activity could have one or more items)
document	Document Id (pdf, audio, video, etc.)
chat_message	Message Id

3.6.5 Implementation

The architecture is built using Javascript and Angular (Google, n.d.). The audio and video are processed using FFmpeg (FFmpeg, n.d.). The contextual interaction are collected using WebSocket (Socket.io, n.d.). The architecture works in a client/server fashion over rest web services. Data are stored in database and pulled for analyses.

3.7 Emotional Data Analysis

This part presents the data exploration and analysis stage of our study. It addresses mainly research question 2), i.e. investigating whether or not audio and video data are correlated. We do not include self-reported information in the analysis because it is not collected at the same time (before and after the session). Throughout this section we answer more specifically the following sub-questions: a) Is there a correlation between audio and video data? and b) What valuable emotional information should be visualized?

3.7.1 Unifying Emotion from Audio and Video Modalities

The use of two different models for emotions (dimensional based on audio, and discrete based on video) was the most challenging part of this comparison. Thus, to be able to compare these cues, both models have to be unified. Several studies have been conducted in this context to convert discrete emotions into a dimensional model (valence and arousal). We used the valence coordinates of the discrete emotions (neutral, happiness, surprise, sadness, disgust, contempt, fear, anger) as proposed in Scherer et al., 2013. The coordinates provided are on a scale ranging from -100 to 100, from negative valence to positive valence (neutral emotion has a valence equal to zero). As described in section 3.1.3, audio valence is on a scale ranging from 0 to 100 (as returned by the Beyond Verbal API). For instance, neutral valence is equal to 50. Therefore, scale unification was needed between the coordinates from (Scherer et al., 2013) and the returned result from the VB API. To this end, we used the formula (f_1), which maps a domain [a, b] interval to a range interval [c, d], to scale the Beyond Verbal interval from [0, 100] to [-100, 100].

$$f(n) = \frac{(d-c)(n-a)}{(b-a)} + c \qquad (f_1)$$

As described in Table 3.1, a successful call to the MS API returns the emotion scores of the frames (images). By combining the coordinates from (Scherer et al., 2013) and the returned scores, we were able to compute for each frame the corresponding value of the valence. The weighted mean was used for this purpose, as illustrated in the formula (f_2), given the set of emotions $S = \{neutral, happiness, surprise, sadness, disgust, contempt, fear, anger\}$:

$$valence_{avg}^{frame^n} = \sum_{i \in S} score^{frame^{n^i}} * valence^i \qquad (f_2)$$

- $valence_{avg}^{frame^n}$ is the weighted valence average —weighted mean, of the $frame\ n$.
- $score^{frame^{n^i}}$ is the score of the emotion $i \in S$ in the result returned by the MS API for the $frame\ n$.
- $valence^i$ is the valence corresponding to the emotion i.

As described in Table 3.2, a call to the BV API returns an array of time segment entries with their associated emotion scores (valence and arousal).

Table 3.4: Correlation between audio and video valence.

	MIC	PEARSON-R	SPEARMAN-R
Correlation coefficient	0.787	0.052	0.172

Thus, a simple algorithm was implemented to map the images extracted from the video, every second, to its corresponding audio time segments. Once the images were grouped, valence is computed first for each image with (f_2). The average over images belonging to each time segment is then computed.

3.7.2 Quantitative Comparison of Emotional Cues

Once the two data models had been unified, the aim was to investigate whether there is any correlation between both variables: audio and video valence.

There are several coefficients for measuring correlations between variables. Spearman's and Pearson's coefficients are those most commonly used. However, they can detect only linear or non-linear monotonic correlations as stated in (Posnett et al., 2012). Conversely, the Maximal Information Coefficient (MIC) may be more powerful as it can detect several associations between variables, such as linearity, non-linearity, asymmetry and even non-functionality (Posnett et al., 2012). Table 3.4 presents the correlation coefficients (MIC, Spearman, and Pearson coefficients) of audio and video valence.

Summary Notes: Regarding the MIC coefficient, the variables (audio and video valence) are correlated (contrary to Pearson and Spearman). We deduce that there is no clear linear correlation between the video and audio variables, but there must be a correlation between the peaks (either positive or negative) of both variables, which is a non-functional correlation that the MIC can detect. At this stage, an exploratory analysis using audio-video records was necessary to further investigate and understand this non-functional dependency between both cues.

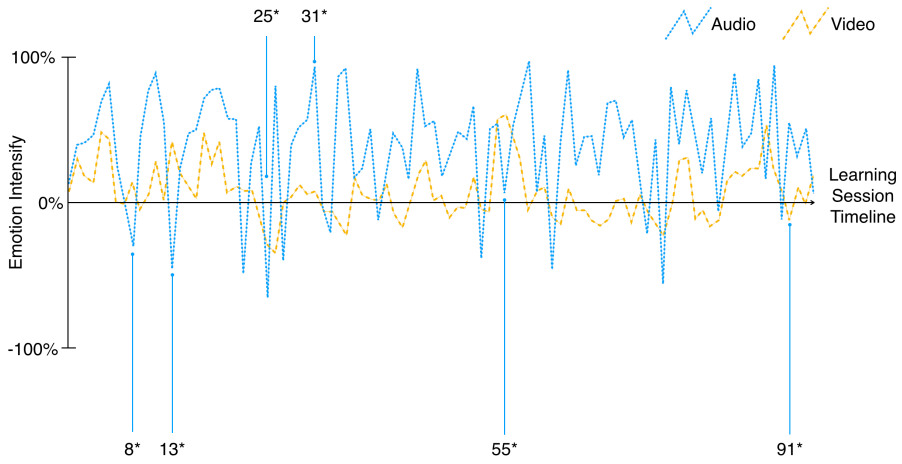


Figure 3.6: Examples of audio and video valence dissimilarities, e.g., 8*, 13*, 55*, 91* and similarities, e.g., 25*, 31*.

3.7.3 Qualitative Comparison of Emotional Cues

We analyzed audio and video valences. The result reveals an explicit dependency between both valences. Figure 3.6 presents some examples of one learning session valence timeline. The x-axis represents the time-segment indices, while the y-axis represents the corresponding valence score between -100 and 100. As annotated with blue color lines, many apparent similarities and dissimilarities exist between the peaks in audio and video valences.

Figure 3.6 presents both examples of **dissimilarities** and **similarities** comparing raw audio and video data for one learning session. Table 3.5 and Table 3.6 present the associated interpretations and comments for the first four segments and first two segments of the dissimilarities and similarities, respectively. The remainder of the segments is more or less similar to the commented ones.

Summary Notes: The comparative analysis conducted in this section reveals some interesting points. First, by adopting different models of emotions and using heterogeneous APIs for emotion recognition, we were able to identify many similarities and dissimilarities between the measured emotions from both audio and video data. This result underscores a dependency between these cues. Second, as described in Table 3.5 and Table 3.6, audio data can reveal some particular levels of valence that video data fail to detect and vice versa. We deduce that both cues need to be considered to best infer learners' emotions.

Table 3.5: Interpretation of valence dissimilarities (audio/video).

Index	Audio	Video	Comment
8	The correction of a mistake by the tutor triggers a regret reaction in the learner. Thus valence decreased.	Almost half of the images have been classified as happiness'. Thus valence increased.	Even if the learner's voice showed regret, there was no sign of this on his/her face.
13	A kind of *"aaaah"* expressed by the learner. Thus valence decreased.	Neutral emotion.	Audio might be a better indicator for some specific emotional vocal expressions.
55	The learner's voice was calm (neutral).	Many images have been classified as *"happiness"*. Thus valence increased.	The learner was smiling, which was a sign of *"happiness"*, but a calm voice was a sign of *"neutral"* for audio.
91	The learner was speaking loudly and confidently. Thus valence increased.	The majority of images have been classified as *"neutral"*. Thus valence decreased.	Neutral was the dominant emotion for the video, while the audio indicated a high level of valence which is not reflected on the face.

3.7.4 Integration of Emotion Awareness Cues

The analysis of emotions from the audio/video recording of the learning session, using a sampling of one frame per second for the video and no sampling for the audio, results in the extraction of a large number of potential emotional expressions (more than 80 vocal expressions and 2700 facial expressions for a 45-minute learning session). Subsets of these data points are, however, likely to constitute an effective emotional state. We thus envisioned to filter and synthetize both dimensional and discrete emotions so as to keep only effective emotional reactions.

As we described in the background (Section 3.3) a variety of approaches are used to combine emotional cues. We discuss here a "greedy" approach to aggregate emotion cues from both audio and video. However, this choice was intentional so as to acquire explainable awareness cues resulting from this merge (fusion).

Table 3.6: Interpretation of valence similarities (audio/video).

Index	Audio	Video	Comment
25	The learner was obliged to interrupt the session in order to talk to a member of his/her family.	*"Neutral"* emotion.	This proves that the result might be more accurate when both cues are combined.
31	There was a joke in this segment that leads to positive valence, as well as a peak in audio valence.	Only one image has been classified as *"happiness"*.	This also proves that the result might be more accurate when both cues are combined.

Concerning **dimensional** emotions from audio, we used arousal —intensity of the emotion, to filter the valence of audio emotion cues. We defined a list of arousal thresholds (10, 20, 30, 40, 50). Namely, when a threshold is selected, only the time-segment entries with the highest corresponding arousal are considered of interest during a learning session.

Concerning **discrete** emotions from video, we used a time series analysis to synthetize the extracted emotions. We first divided the set of emotions: neutral, happiness, surprise, sadness, disgust, contempt, fear, anger, into three categories: neutral, positive (happiness, surprise) and negative (sadness, disgust, contempt, fear, anger). Second, we computed the scores of these emotions as the sum of the corresponding emotions for each category and thus, for each image (frame), with respect to the $(f3)$, $(f4)$, $(f5)$ formulas described below. Third, once the scores had been computed for each image, we identified the dominant type of emotion (neutral, positive, or negative) as the most frequent over all the frames of each segment.

$$score_{neutral}^{frame^n} = \sum_{i \in \{neutral\}} score^{frame^{n^i}} \qquad (f_3)$$

$$score_{positive}^{frame^n} = \sum_{i \in \{happiness, surprise\}} score^{frame^{n^i}} \qquad (f_4)$$

$$score_{negative}^{frame^n} = \sum_{i \in \{sadness, disgust, contempt, fear, anger\}} score^{frame^{n^i}} \qquad (f_5)$$

- $score^{frame^{n^i}}$ is the score of the emotion $i \in S$ in the result returned by the MS API for the $frame\ n$.

We then identified two patterns: the positive and the negative moments; neutral is dropped as it may be of little interest and also to avoid presenting too much information that might overwhelm the user. We then defined two measurements, *True positive* and *True negative* to track these two patterns. *True positive* detects the moments where the average of positive points is highest or equal to a fixed threshold. On the other hand, the *True negative* detects the moments where negative emotions are defined as significant, i.e., the gap between the average of negative points and the maximum score of the remaining emotions is less than or equal to a fixed threshold. We therefore consider the resulting positive and negative moments for both dimensional and discrete emotions as the most significant time-segments that are more likely to represent effective emotional reactions during the learning session.

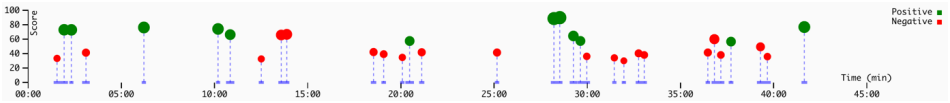


Figure 3.7: The bubble shape represents discrete emotion markers, while the star shape represents dimensional emotions. Bubble and star size depends on the computed score for each emotion, which is also represented by the height of the marker.

Summary Notes: Positive and negative emotional reactions detected can be used to design emotion *awareness cues*. Such cues can, for instance, be a timeline representation of the learning session in which we might combine both discrete and dimensional emotions, along with self-reported emotions and contextual information, as we did, for instance, in the design of Emodash described in chapter 4. Figure 3.7 portrays a simple example of a representation of a set of positive/negative markers for positive and negative time-segments, respectively.

3.8 Discussion

This work does not come without limitations. First, we did not emphasize how awareness cues captured here may inform learning or teachers' pedagogical practices. Second, we did not tackle the ground-truth of the algorithmic aspects in terms of accuracy of the APIs used to infer learners' emotions. Future work could consist, for instance, in designing an experiment to evaluate whether such APIs are accurate with respect to one's felt emotions. Third, the case study is learner-centric, and we believe it may prove useful to investigate whether there are correlations between the emotions expressed by the teacher and by the learner. Such a study may first show to what extent teachers enjoy interactions with learners through live learning sessions and whether they experience negative emotional experiences. Also, it may shed light on the emotional dynamics, patterns, contingencies, and labor of teachers and learners during live learning interactions.

Although we demonstrated our approach in a video-conferencing learning environment using visual and vocal expressions, the approach is, however, more or less generalizable to other types of information, such as textual or click-stream based. Besides, we opted for a post-processing (off-line) approach when capturing and quantifying emotions. Capturing real-time emotional awareness cues could be made possible by altering slightly the current architecture.

Despite the potential gain of using multi-modalities, the design and use, i.e., capture, quantify, and integrate, of such awareness cues comes with different challenges that need to be addressed properly. These include sampling when capturing and aggregating data, temporal alignments between cues, accuracy of the transformation (classification) of cues to drive other information, explainability of the cues, and complexity when encoding cues (Azevedo and Gasevic, 2019).

While focusing on fostering emotional awareness cues through universal discrete emotions, i.e., anger, disgust, fear, happiness, sadness, surprise; and dimensional emotions, i.e., arousal, valence; the learning settings might, however, be more concerned with learning-specific emotions, for example, pride, shame, boredom, frustration, pride (e.g., Pekrun, Goetz, et al., 2011). Despite their importance, designing around domain-specific emotions is still challenging, as they are almost entirely subjectively assessed. We rather sought to capture and quantify emotional awareness cues that may seamlessly integrate within the day-to-day practices of teachers and learners.

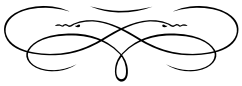
3.9 Summary

The current initial investigations prove, however, the feasibility of using cloud APIs for emotion recognition in video-based online learning environments. By applying both dimensional and discrete emotion models, we showed that we can detect different learners' emotions during such settings. This highlights the pertinence of multimodal approaches in capturing, quantifying, and integrating emotion cues. Our analysis shows that different emotions, e.g., positive and negative, are detected from different modalities, i.e., audio and video, in the settings of live learning sessions.

The design of multimodal emotional cues for teachers using representations that are easy to use and make sense of, may prove a challenging endeavor. One way to approach such a challenge would be to investigate the right balance between the amount of emotional cues gathered and represented to users, and the amount of contextual information made available to users to help interpret one's emotional state; this might mitigate emotion recognition accuracy. We describe in chapter 4 how we built upon our exploration described here to design around emotional cues. We illustrate our approach in the design of EMODASH, an interactive dashboard designed using an iterative design process, to foster retrospective emotional awareness of teachers. By doing so, we hope to investigate the impact of emotions on learning or pedagogical practices in naturalistic settings, an area awaiting further exploration.

Chapter 4

Supporting Teachers' Awareness of Emotions in Online Learning Using a Dashboard



> *A teacher who recognizes an individual student's emotions, understands the cognitive appraisals that may be associated with these emotions, and how these cognitions and emotions motivate the student's behavior can effectively respond to the student's individual needs.*

—Jennings and Greenberg, 2009, p. 493

4.1 Overview

Emotional awareness plays an important role in multiple parts of learning settings. However, they are harder to develop in video-conferencing environments due to technology-mediated interactions, technical challenges, and teachers' main focus on properly conducting learning activities. Building upon our technical exploration of emotional cues presented in chapter 3, we then designed EMODASH an interactive dashboard to assist teachers in developing retrospective awareness of learners' emotions in a video-conferencing learning environment. To evaluate the dashboard, we conducted an in-the-wild field study with five professional teachers. In particular, we sought to understand how teachers used EMODASH, while writing feedback to learners after language learning sessions. We found that EMODASH led teachers, who were already sensitive to learners' emotions, to incorporate more affective elements in their reports, thereby suggesting stronger awareness of learners' emotions. Teachers also wrote more formative and less summative feedback. Furthermore, our results suggest that glanceable visualizations of learners' emotions may be preferred and sufficient to foster teachers' awareness of learners' emotions. Finally, the dashboard led teachers to reflect on the way they conduct their lessons, using learners' positive emotions as a proxy evaluation of their teaching.

4.2 Motivation

Learners' emotions have a significant impact on learning processes and outcomes (Pekrun, Goetz, et al., 2011; Jennings and Greenberg, 2009). Emotions can also denote the presence of important or complex aspects of learning tasks that require sustained effort and attention (Boekaerts, 2010).

However, developing an awareness of learners' emotions in online learning settings is challenging. Interactions are mediated, sometimes asynchronous, and online spaces leave little room for informal exchanges (Dourish and Bellotti, 1992) that could help teachers develop an awareness of learners' emotions (Boden and Thompson, 2015; Lavoué et al., 2015).

Despite the importance of emotions, few "emotion-aware" dashboards have been designed and studied in a learning context. Such research can be grouped into two separate strands. One consists in supporting learners' self-awareness of and reflection on their emotions (Sun et al., 2019). Often, such dashboards are directed to capture and present learners' self-reported emotions to support their emotional awareness (Ruiz et al., 2016).

On the other hand, the second strand aims at capturing and reflecting back emotion cues to foster emotional awareness of teachers, advisors, etc. For instance, instead of using subjective (or self-reported) emotions, other dashboards have been directed to infer automated learners' emotions in online learning, such as using textual cues (Montero and Suhonen, 2014), learners' actions logs (Leony et al., 2013), and video cues (GhasemAghaei et al., 2016). Although few notable studies shed light on the importance of emotion in learning contexts, there persists a critical need to understand the impact of emotional awareness cues on typical learning activities.

We therefore designed EMODASH, an interactive dashboard, via an iterative design process (J. Nielsen, 1993) to foster teachers' awareness of learners' emotions of past learning sessions. The dashboard presents awareness cues about the emotional state of learners along with contextual cues, namely, interaction traces of both teachers and learners. EMODASH is part of the video-conferencing based learning platform for foreign languages SPEAKPLUS described in chapter 3, and integrates with the pedagogical tool that teachers use after each learning session to write reinforcing feedback reports for learners.

We conducted a field study of EMODASH for eight weeks with five teacher-learner pairs. We analyzed how teachers used the dashboard and the feedback reports they wrote *with* and *without* EMODASH. We sought to understand the type of emotion information teachers rely upon when writing feedback reports for learners, and investigate how awareness of learners' emotions impacts the feedback reports teachers share after learning sessions. We more particularly analyzed the content of the feedback reports to examine the extent to which emotion awareness led teachers to provide learners with formative feedback on why a particular task is correct or incorrect, and how to maintain or improve their performance (e.g. "I believe that success in interviews is attributable to practice and confidence"), compared to summative feedback that provides learners with an overall evaluation of their successes and failures (e.g. "You can make improvements"). In addition, we examined the impact of emotion awareness on the incorporation of motivational (e.g. "Good job") or affective content (e.g. "It was a pleasure to meet you") in their feedback.

We found that EMODASH led teachers to adapt the content of their feedback reports by integrating more affective language. Teachers also wrote more formative and less summative feedback reports. The teachers who changed their feedback reports most are those who interacted least with the dashboard, suggesting that glanceable visualizations of learners' emotions may be sufficient. Teachers appreciated the ability to identify learners' emotions at a glance, and to evaluate whether there were more positive or negative emotions during a session. Overall, visualizing learners' emotions would

seem promising to increase teachers' awareness of learners' emotions, as well as to develop their self-awareness. Teachers used EMODASH as a feedback mechanism as to how they are when conducting their lessons, considering that learners' positive emotions were a sign of lessons that worked well.

The contributions of this work are:

1. EMODASH, a dashboard designed based on interviews with five teachers and one learner and refined via an iterative design process, to facilitate teachers' retrospective awareness of learners' emotions, and support teachers' writing of reinforcement feedback reports to learners after video-conferencing language learning sessions.

2. An in-the-wild field study of EMODASH. This study extends our understanding of how emotional dashboards support retrospective awareness of learners' emotions in a video-conferencing language learning environment.

3. A qualitative and quantitative study comparing teachers' written feedback reports under two conditions *with* and *without* EMODASH.

4.3 Background

We focus on teachers' feedback after online synchronous sessions. Teachers' feedback can be effective for stimulating students' progress, if it is explicitly related to learning goals (van den Bergh et al., 2012). Learners are more likely to improve when teachers provide specific feedback on current performance, appropriate tasks and/or processes to improve upon, with clear goals to be achieved (Hattie and Timperley, 2007). Prior work has examined learners' receptivity to feedback, highlighting the importance of the source (Nguyen et al., 2017) (e.g. anonymous source vs. from a peer or an authority), as well as the nature and the content of the feedback (Tanes et al., 2011; Sun et al., 2019). For instance, formative feedback offers ways to maintain or improve performance while the task is still being performed or completed (Tanes et al., 2011). Affective language in feedback can also improve the receptivity of learners to this feedback (Nguyen et al., 2017).

Previous studies on the impact of emotion awareness focused on interactions between peers during collaborative activities. For instance, Molinari et al. (2013) showed that, in CSCL settings, learners benefit from being aware of what their collaborative partners feel during interaction. Learn-

ers that communicated their emotions to each other are more likely to build on their partner's ideas and to interact together in a transactive way. Lavoué et al. (2015) showed that students used emotional markers to reflect on their partner's activity, and to express their satisfaction when writing self-reports on online sessions. Another example is MoodMap (Fessl et al., 2012), which enables users to note and review their own mood over time, as well as to obtain insights into team mood according to a given meeting and/or a date. The authors showed that mood self-tracking improves cohesion within teams and leads managers to react proactively to changes in team members' mood (Rivera-Pelayo et al., 2017). Samrose et al. (2017) also showed that providing teammates with post-session feedback on group dynamics (including emotions) changes the way they participate.

Based on this previous research, our study focuses on the impact of retrospective emotion awareness on teachers' feedback to learners. Emotional information may support teachers in identifying successes and failures in their interactions with learners, as well as provide surface explanations on difficulties they encounter. Awareness of learners' emotions could also lead teachers to adapt their feedback to learners, i.e. by reassuring or motivating them (Linnenbrink-Garcia et al., 2016), or by integrating affective language. In our study, we examine the impact of emotion awareness on teachers' feedback. Building upon the feedback content categories proposed by Tanes et al. (2011), we investigate whether emotion awareness tools lead teachers to offer formative or summative feedback or to incorporate more affective or motivational content in their feedback.

4.4 Speakplus's Feedback Report Editor

As we described in chapter 3, the learning process on Speakplus is structured around three main activities:

1. Pre-session: before live learning sessions, teachers can create and customize learning materials (activities) for each learning session. Each learning activity has a duration and usually contains a set of documents, instructions for the learner or a set of personal notes for the teacher.

2. In-session: the teacher conducts a synchronous learning session with his/her learner. The pair (teacher and learner) communicates in real-time through video-conferencing. The teacher has the ability to share learning materials/activities with the learner (e.g. pdf, word, image, audio, and video), or to communicate via chat.

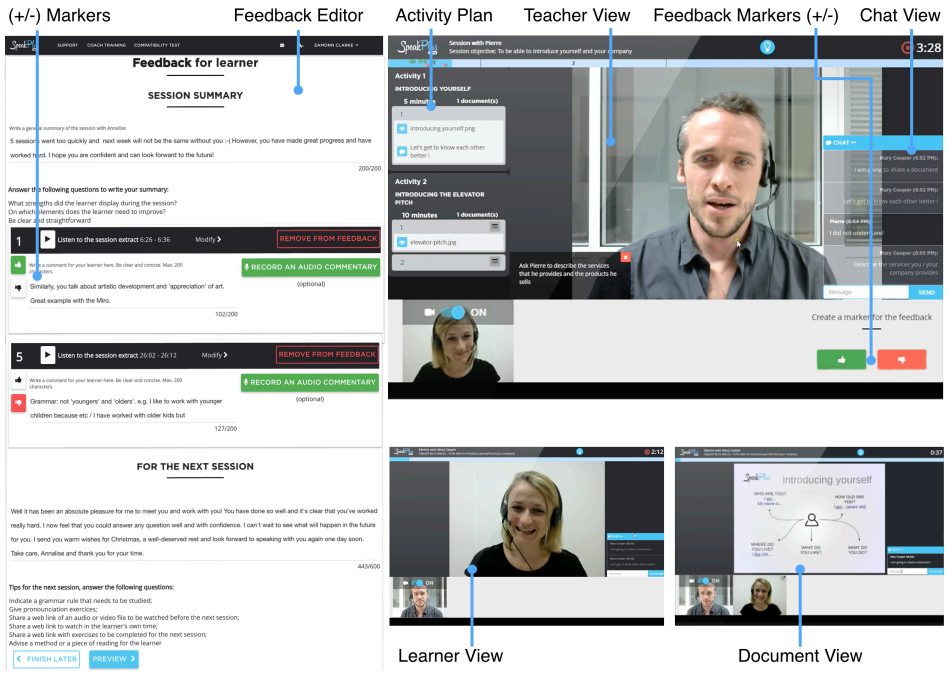


Figure 4.1: Example of a live learning session on SPEAKPLUS with the interface of the feedback report editor tool. We will explain the different components of the feedback report editor in the subsequent sections.

3. Post-session: the teacher writes a feedback report for the learner, including a summary of the session, as well as a set of instructions for the next session.

As illustrated in Figure 4.1, during the session, teachers can add annotations (called markers) to pinpoint learners' positive and negative moments. At the end of each session, teachers use these markers to write a pedagogical feedback report using a dedicated feedback report editor to reinforce learners' learning. Once the feedback report has been completed, the teacher shares it with the learner, who is notified by email. Learners can read the feedback on the platform, alongside an audio recording of the learning session.

4.5 EMODASH Design Process

Based on the role of emotion awareness in the learning process, we investigated various strategies to increase teachers' awareness of learners' emotions. Figure 4.2 presents the timeline of our design process. We went through six main iterations, over a period of almost two years, involving different stakeholders at each stage. We started with interviews of

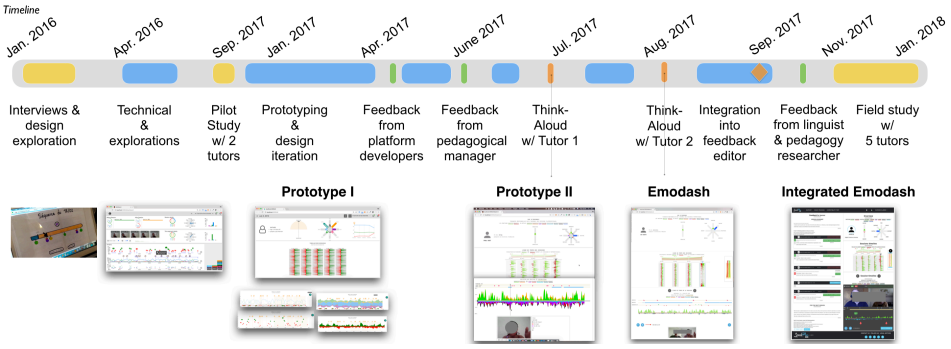
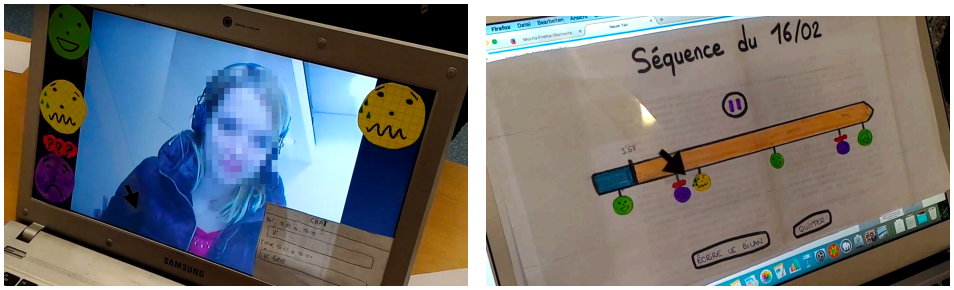


Figure 4.2: Key stages of EMODASH design process.



(a) Example of an early sketch of real-time emotional cues.

(b) Example of an early sketch of retrospective emotional cues.

Figure 4.3: Examples of early sketches.

teachers, as well as design explorations of live and retrospective awareness. We later explored and validated emotion recognition strategies, which we tested with two teachers. We further refined the design through several iterations geared towards feedback support, and finally conducted the study discussed in this paper.

4.5.1 Preliminary Interviews and Early Design Explorations

Preliminary interviews with five teachers and one learner highlighted the importance (and lack) of socio-affective development on the platform. This relates to the development of trust, mentioned by several teachers, the need to engage learners remotely, and the challenges related to cultural differences. Combined, these elements have a direct influence on how learners behave, for instance, how they will ask questions or interact with the materials.

Although synchronous video-conferencing enables participants to grasp facial expressions and gestures, remote learning is still poor in terms of socio-affective exchanges. Teachers' primary focus on their pedagogical goals, as they conduct learning sessions, could account for the lack of

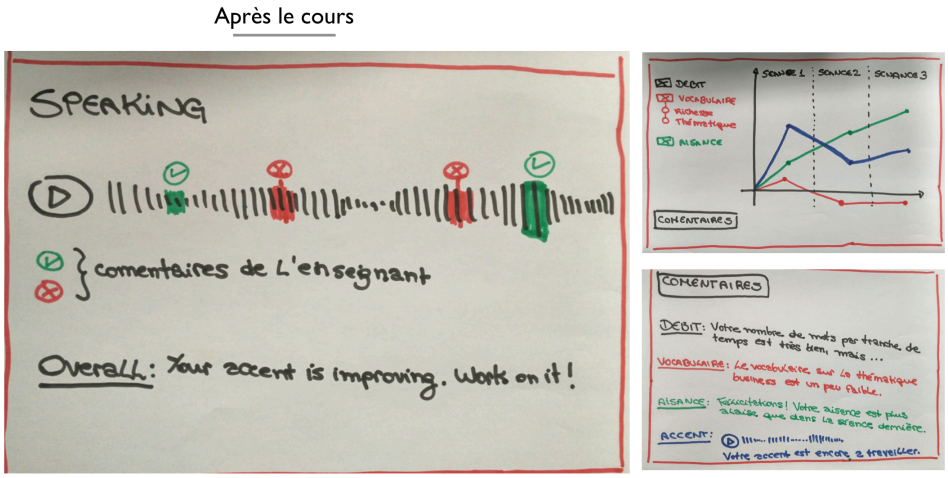


Figure 4.4: Example of a sketch of a dashboard.

emotion awareness. This problem becomes greater still in the context of SPEAKPLUS, with full-screen document sharing, often hiding the video feed and further inhibiting the development of emotion awareness during exchanges.

We explored early sketches of real-time and retrospective emotional awareness cue sharing through various forms of emojis, as shown in Figure 4.3b and 4.3a. During design explorations, opinions of the teacher, as well as the designers of SPEAKPLUS, gravitated towards the use of a dashboard offering an overview of learners' emotions over the past learning sessions (see example in Figure 4.4).

4.5.2 Pilot Study: Supporting Retrospective Emotion Awareness

We further explored fostering retrospective emotional awareness after each learning session through the design and use of dashboards. To that end, we built upon our investigations, presented in chapter 3, of collecting, processing, and representing emotions with a multimodal approach (audio + video) to ensure greater reliability of emotion recognition.

We converge in a **first prototype** building upon our technical exploration. The prototype, portrayed in Figure 4.5, is a web-based application enabling a teacher to explore the emotional state of his/her learners after each learning session. The dashboard was designed so as to present information from general-to-specific at different levels of abstraction: (1) overall information about the learning session conveys a summary of the emotional state of the learner during the learning session, combining both subjective and objective emotions (Figure 4.5.A); (2) temporal information integrates

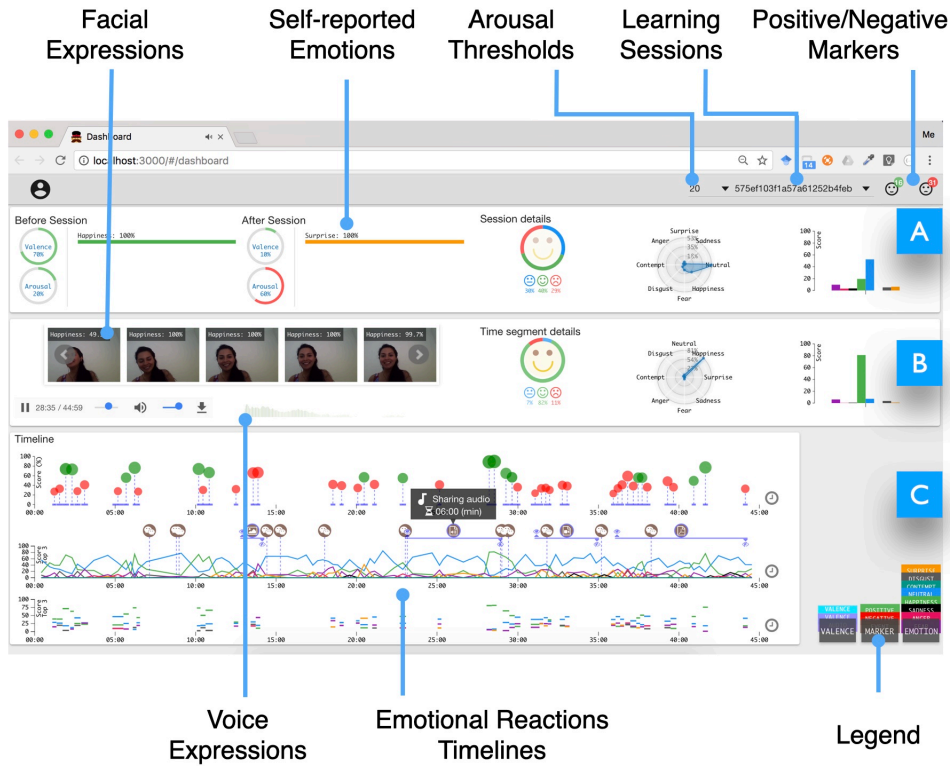


Figure 4.5: The interface of the first prototype. A positive marker is selected in the first timeline.

contextual information, together representing a timeline of the learning session seen through the emotional experience of the learner along the learning activities (Figure 4.5.C); (3) time-segment information designed to provide more details about each positive or negative emotion identified in the timelines (Figure 4.5.B).

We followed this up with a **pilot study** to gather initial feedback about teachers' perception/experience regarding the emotional cues captured and encoded on the dashboard. We conducted the pilot study with two teachers used to teaching on SPEAKPLUS, and structured this study in three steps. The first step consisted of a case study with 8 associated questions, e.g., *"At what time did the learner feel a peak of fear during the learning session?"*. The second step consisted of 6 questions based on a 7-Likert Scale regarding global perception on the dashboard, e.g., *"How do you find the display and organization of data on the screen?"*, *"To what extent do you think the dashboard will help you as a tutor to improve the learning experience?"*. The third step consisted of 7 grid questions with 7-Likert scales for each visualization, e.g., *"How do you find the learner image carousel displayed in part (II-a)?"*. We added to each question a text area requesting more explanations.

The **results** showed that teachers reacted positively to the dashboard. The answers to the questionnaire associated with the case study confirmed that

the dashboard is easy to use, as the tutors were able to answer all questions correctly without any help. Regarding global perception, it was clear that our dashboard has a problem in displaying timeline visualizations. Tutors state that *"reading is a little bit difficult"* —P1, and *"Conversely, timeline visualizations were not rapidly comprehensible [...] for me it would be better to keep only one simplified visualization"* —P2. This suggests that one simplified visualization of the three timeline visualizations would be easier to understand, for instance by adding a tooltip with the mouse hover events to the marker timeline. This apart, the perception of the dashboard was mainly positive. From the last part of the questionnaire, we note that the idea of combining emotions and events was very interesting for both tutors. In fact, they suggested further *"promoting"* —P1, this visualization: *"should be better valued"* —P2, *"the idea of combining events and emotions is interesting, but the chart is not large it's a good idea but should be simplified"* —P2. Therefore, more focus should be placed on the contextualization part as it provides essential understanding of the emotions experienced by learners.

Overall, the results of the pilot study emphasize three points:

1. While relevant, representation of multimodal data was too complex to explore and interpret. Teachers needed, however, a dashboard that is faster to grasp.

2. They needed a dashboard centered around one of their main tasks, apart from the learning session: writing feedback reports. This task required an awareness of learners' emotions over past learning sessions.

3. They needed emotions to be contextualized with learners' activities to better understand what could have triggered them. This last point is in line with several models of emotion regulation (Gross, 2008; Pekrun, 2006): Teachers can better understand learners' emotions and help them reappraise the situation (i.e., positively change the perceived value or control of the learning activity) if they are aware of the antecedents of their emotions.

Given the satisfactory quality of facial emotion recognition from the video feed, we decided to rely only on this source of data. The video stream enables us to extract learners' discrete emotions, e.g., neutral, happiness, anger, disgust, fear, sadness, surprise; which seem to be more understandable and familiar to teachers than valence and arousal. At the same time, we designed emotional awareness cues around contextual information, namely, a replay video recording of the learning session, along with the interaction traces of the teacher-learner pair for better readings.

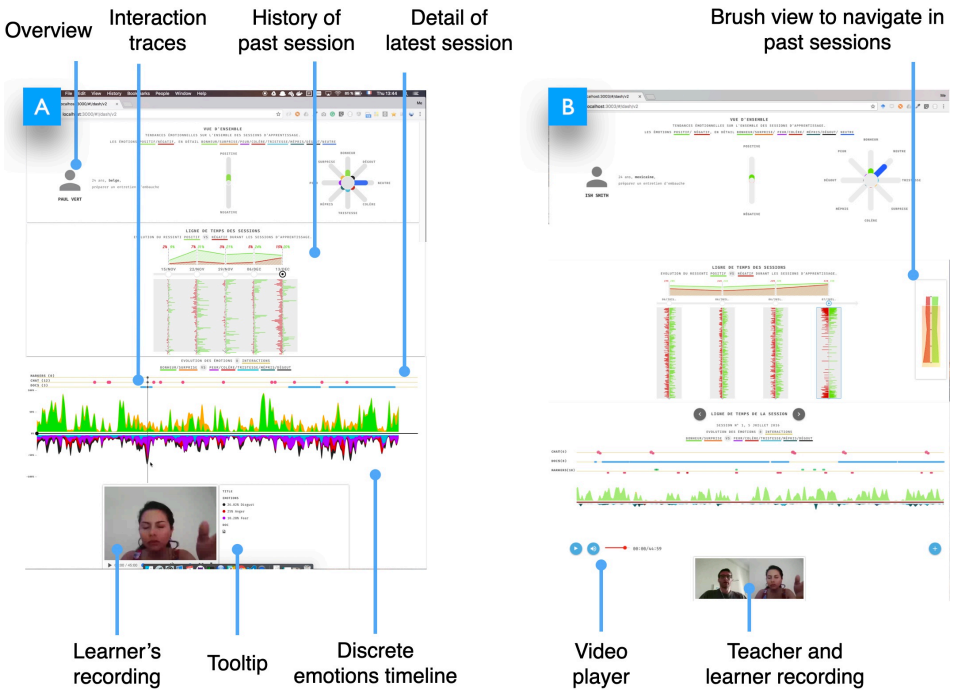


Figure 4.6: On the left (A) is interface of the first prototype on which we conducted the first usability test with Tutor 1. On the right (B) interface of the second prototype on which we conducted the second usability test with Tutor 2.

4.5.3 EMODASH Iterative Design

We iteratively designed EMODASH over three months until we converged toward a dashboard that could be tested (see Figure 4.5.3.A). During this iterative process, we presented the dashboard informally to stakeholders: a pedagogical manager of the platform and developers.

After converging towards a design (see Figure 4.5.3.B), we sought to identify usability issues involving two teachers. The first author conducted two tests as per the think-aloud method (C. H. Lewis, 1982). We conducted the tests on a 13" MacBook Pro. We first read a script to explain the purpose and organization of the test. Then, we asked teachers to imagine that they had conducted some learning sessions, and to use the dashboard to explore the emotional state of the learner involved. We asked to describe their thinking aloud as they interacted with the dashboard. Teachers gave their permission to record their voices and the laptop screen. The first session took 30 minutes, and the second one hour. After each test, we transcribed the audio, then summarized the issues highlighted by the teacher, as well as those we observed during the test. We then implemented important features and fixed the issues before moving on to the next usability test.

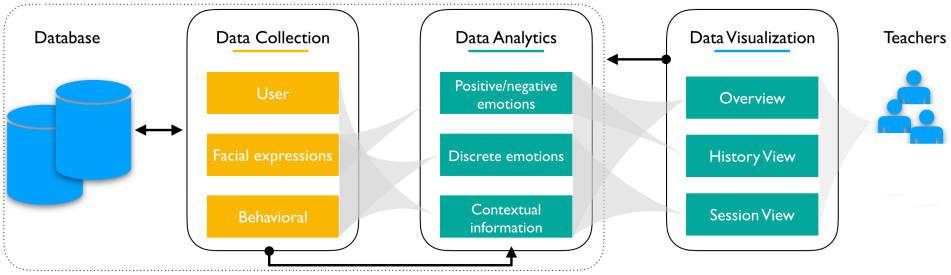


Figure 4.7: The architecture highlighting the types and relationships between raw data, indicators, and views of EMODASH.

We identified the following issues: difficulties in grasping information at a glance, for instance, whether a learner expressed positive or negative emotions during a session; navigation across views could be improved; lack of visual clarity on the session and history views; and that more details on each session (chat content, documents viewed) could help teachers contextualize. We present the changes on each view of the dashboard as we present them in the next section. Think-Aloud observations are detailed in our supplementary material (see *Think-Aloud*).

4.6 EMODASH User Interface

EMODASH augments SPEAKPLUS feedback report tool. It enables teachers to explore past teaching sessions through the lens of learners' emotions alongside the interactions of teachers and learners during the sessions. Figure 4.7 portrays the architecture of EMODASH, namely, the types and relationships between raw data, indicators, and views of the dashboard. The three views of EMODASH offer increasing levels of detail: from a general overview of learners to a detailed timeline enabling a replay of each session (Figure 4.8).

4.6.1 Data Collection and Analysis

EMODASH builds upon two types of data sources derived from the learning sessions: emotion recognition and users' interaction with SPEAKPLUS.

EMODASH relies on *video-based emotion recognition* to build its visualizations. We use the video stream as the main data-source for analyzing learners' emotions, as it enables us to collect data in an authentic and non-intrusive manner. We used the Microsoft Emotion Recognition API, which takes an image as input and returns normalized scores (the scores sum up to

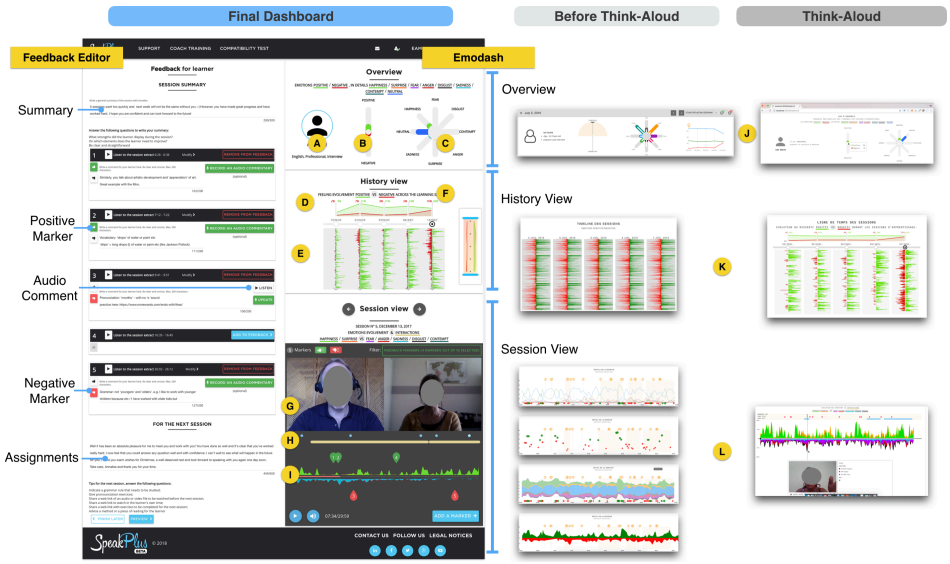


Figure 4.8: Left: EMODASH integrated in SPEAKPLUS feedback report editor. Right: Design Explorations.

1) for a set of emotions based on Ekman's classification (anger, contempt, disgust, fear, happiness, neutral, sadness, surprise) (Ekman and Friesen, 1976). In EMODASH we refer to positive emotion as the sum of the scores of happiness and surprise, and to negative emotion as the sum of the scores of the remaining emotions, without the neutral one. Details about the end-to-end automated process of emotion classification is provided online as supplementary material (see *Data*).

We combine learners' inferred emotions with interaction traces for teachers to contextualize and interpret them. We track the interactions with SPEAKPLUS performed by the teacher and the learner during the learning sessions. The data collected are: sharing files (pdf, video, audio, etc.), sending chat messages, and adding markers (positive or negative).

4.6.2 Overview

The first level is designed to summarize learners' emotions over all their learning sessions (Figure 4.9). It is composed of:

- (A) *Learner's profile* includes an avatar, a name, and a learning objective.
- (B) *Positive-negative emotion balance chart* provides distribution of learners' positive and negative emotions.
- (C) *Emotion radial chart* shows the overall distribution of the emotions: anger, contempt, disgust, fear, happiness, neutral, sadness, surprise.

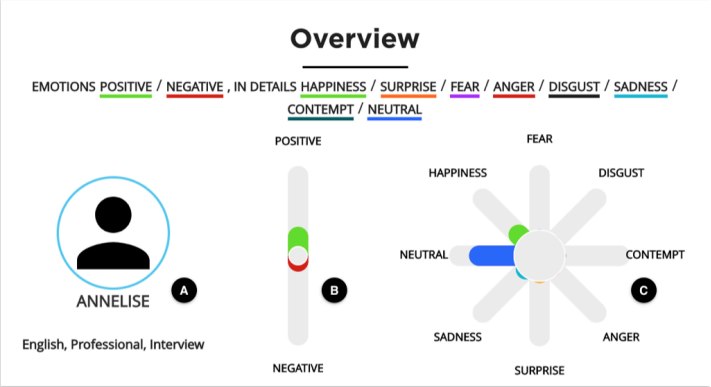


Figure 4.9: EMODASH Overview.

We used a color encoding scheme to distinguish the attributes, both in the positive-negative balance and in the radial column charts. We integrated the legend directly in the header, a fact that participants found attractive and easily recognizable in the think-aloud evaluations. Tooltips on both charts provide additional details. They display the name of the emotion with the corresponding score.

Design exploration: The first version of the overview conveyed more information, such as a timeline of positive, negative and neutral emotions, as well as a discrete count of the number of positive and negative intense emotions detected (see Figure 4.8.J). To facilitate reading, we simplified labels and expanded them. We also removed the timeline from the overview, and created a dedicated view (presented below), with more historical details on learners' emotions.

Think-aloud changes: Before think-aloud, the balance of positive/negative emotions showed the "sum" of positive and negative emotions, so that the chart shows the emotion with the maximal value (intensity), i.e., either the positive or the negative. Our reasoning was that it would make it quicker to grasp the overall direction of emotions. However, Tutor 1 found it confusing that no negative emotions were visible although some were present in the other charts. We thus modified the chart to display both positive and negative emotions.

4.6.3 History View

The second level of EMODASH displays distribution of the positive-negative emotions of the learner over time. At the top (Figure 4.10.A), a multi-line-area chart presents the evolution of the average positive-negative emotions

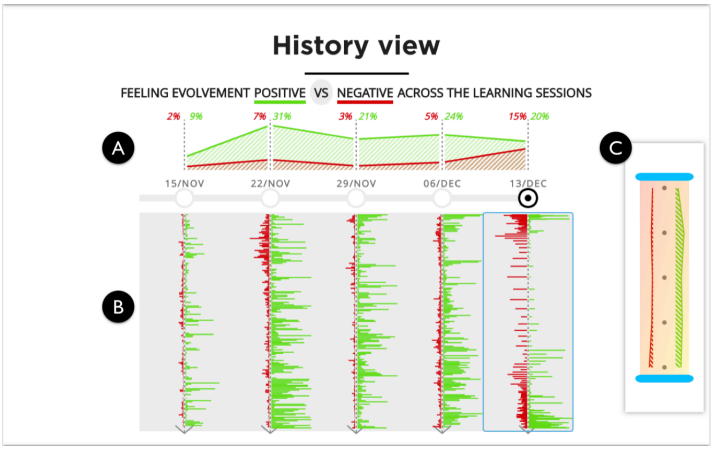


Figure 4.10: History View.

across the learning sessions. The teacher can thus compare the distribution of positive and negative emotions in place. Below is a vertical timeline for each session (Figure 4.10.B). The positive-negative emotions are split along a central axis to make it easier to identify positive and negative peaks. The vertical timelines display a fisheye distortion on hover and, when clicked, update the session timeline view (3rd level below).

The teacher can rely on the first chart (Figure 4.10.A) to visualize how emotions evolve from one session to another, and on the second chart to get more details (Figure 4.10.B). We found it interesting to see that some sessions obtained close averages despite having very different distributions. Similarly to the previous charts, the color encoding legend has been integrated into the view headings. On the right-hand side of this chart (Figure 4.10.C), a brush tool supports navigation across multiple sessions, which is useful when a large number of sessions are available (i.e. more than 5). The component is a miniature of the multi-line-area chart where dots represent sessions.

Iterative design: The first iteration of the History view provided dense timelines of positive/negative emotions for each session (see Figure 4.8.K). It was difficult to grasp the evolution of emotions across sessions. To provide a clearer overview, we simplified the timeline visually by averaging positive and negative signals over a larger time-span, so that users could more easily recognize whether a session involves more positive or more negative emotions. We also added the summary timeline of positive/negative emotions (Figure 4.8.D).

Think-aloud changes: The first think-aloud revealed that navigating across sessions was not obvious and would not scale well as the amount of recorded sessions increased. We added scrubbing (Figure 4.8.F) for nav-

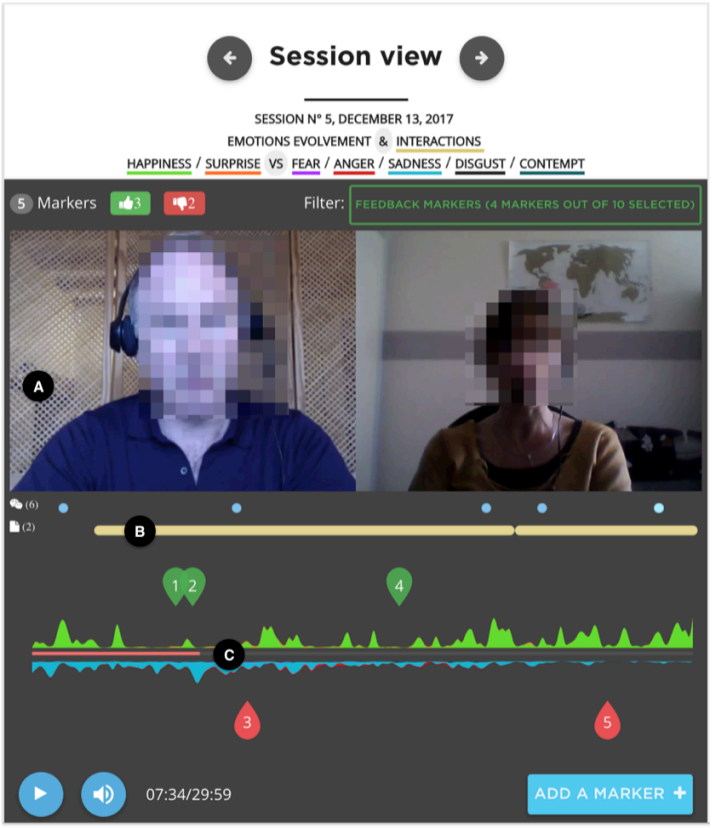


Figure 4.11: Session View.

igating across sessions. We facilitated navigation to past sessions so that clicking on one session of the history view would change the session view accordingly.

4.6.4 Session View

The session view is centered around a timeline of one learning session. Along with the emotions identified during the session, the interactions between users and the platform during the live learning session are visualized in this view (Figure 4.11). The main interactions are document-sharing (pdf, image, video or audio), chat messages, positive/negative markers. Teachers can add more markers if they want. These are then used to create the feedback report for learners.

All the emotions captured are displayed in this view, except the neutral one, which corresponds to the whitespace on the chart (Figure 4.11.C). Splitting positive and negative emotions along the horizontal axis makes it easy to interact with peaks of emotions on both sides (positive and

negative). Visualization helps the teacher recognize past sequences within the sessions, remember what happened, contextualize and reflect on the experienced emotions. The slider of the player in the middle of the chart can be used to navigate in the stream. The user can also directly click on the chart to navigate through the video, which is particularly useful to navigate to a specific pick of emotions. Two buttons support backward and forward navigation between sessions.

Iterative design: The first iterations did not incorporate a video replay of the session. They displayed the evolution of emotions alongside session interactions (e.g. chat messages, document-sharing), which overloaded visualization (Figure 4.8.L). To simplify the design: (1) We removed the neutral emotion, which is the dominant emotion recognized. (2) We centered positive/negative emotions along the horizontal axis, so that positive emotions would lead to peaks going up, and negative emotions to peaks going down. (3) We moved interaction events to a separate timeline and added video playback.

Think-aloud changes: Tutor 1 underlined the need for more contextual information to interpret learner's emotions as (1) the timeline presented only the interaction events, but not the object they were on (document, chat message, etc.), and (2) the playback displayed only the learner's video but not the teacher's one (Figure 4.11 – Prototype II). We added tooltips giving details on the emotions presented on the timeline, previews of shared documents, chat messages, as well as video-recording of both learner and teacher side-by-side. Tutor 1 also remarked on the need to navigate the video around peaks, so we added direct navigation using the streamgraph. In the second think-aloud, Tutor 2 found navigation with peaks interesting, as well as the contextual information we had added. During the test with Tutor 2, we observed that the video player was not positioned close enough to the emotion timeline. We decided to move the video player closer, so that the timeline would be more readily used as the controller. We switched to a dark-grey background for the session view to improve the contrast with the visual elements.

4.6.5 User Interface and Visualization Implementation

The user interface (UI) of EMODASH is built using JavaScript and Angular (Google, n.d.). The visualizations are implemented using D3js (Bostock et al., 2011). The views of EMODASH work in a client/server mode. When the UI starts loading, HTTP requests are sent to the server to pull the data. Data are pulled from the database and aggregated to produce JSON data objects for each view.

The video player is built on the top of HTML5 video specifications. It can play a video stream with play/pause/seek, volume up/down actions. We augmented it to render and visualize time series datasets with D3js.

4.7 Field Study Design

We conducted an in-the-wild study of EMODASH over eight weeks between November and December 2017. Our goal was to investigate the impact of retrospective emotion awareness in distance learning, i.e. delayed awareness of emotions felt by learners during synchronous interactions with a teacher. We focus more specifically on teachers' feedback, and the impact emotion awareness can have on written feedback reports to learners. In this context, we aim to answer the following questions:

1. ***Do emotion awareness tools have an effect on the type of feedback given, specifically formative or summative feedback reports?*** Could making teachers aware of learners' emotions help them interpret the way they learn and enrich their feedback with more formative than summative reports.

2. ***Do emotion awareness tools lead teachers to incorporate more affective and/or motivational content in their feedback?*** We believe awareness of learners' emotions could lead teachers to integrate more motivational or affective language.

3. ***What effect do different granularities of emotion visualization, via a dashboard, have on teachers' feedback?*** Progdash conveys information at different levels, ranging from an overview of learners' emotions to a detailed view of emotions identified during learning sessions. We explore teachers' use and perception of the different visualizations to provide design recommendations.

4.7.1 Participants

We recruited five professional teachers and five learners with English as a secondary language.

Teachers. We recruited five teachers (gender: [M=2, F=3], age: [Mean=42.8, Min=25, Max=58]) via an email to all SPEAKPLUS teachers. We relied on the SPEAKPLUS pedagogical manager selecting the final set of teachers, based on their experience and familiarity with the platform. The teachers

recruited had more teaching experience on Speakplus (P1: 181h, P2: 227h, P3: 66h, P4: 20h, P5: 91h), than the average teacher on the platform ($Mean_{participants}$ = 103, $Mean_{platform}$ = 26.45).

Learners. We recruited 5 learners (gender: [M=0, F=5], age: [M=32.6, Min=20 Max=37]) through a post distributed via emails and social networks. The post described the study in broad terms. It was addressed to people interested in benefiting from free learning sessions on a specific learning objective. No other credit was given to learners and all of them were first-time users of the platform.

All the learners were French native speakers. Their English level ranged from elementary (identified as "I know a little vocabulary but have difficulty putting words into sentences to express myself") to proficient ("I can take part in clear and coherent, well-constructed conversations on a variety of topics"). Learning objectives were: "improving oral fluency", "preparing a job interview", "working with American interlocutors", "applying for a job in an English-speaking country" and "preparing a work trip".

Both teachers and learners were informed about the logging and the possibility to withdraw at any time. They gave permission for recording and analyzing learning session data, and signed online for consent. They also filled out a preliminary questionnaire that collected demographics, motivation and English level (only learners).

4.7.2 Procedure

The deployment lasted for eight weeks and was divided into 4 stages: training, pair-formation, study, and wrap-up.

Training. During the first week, as our teachers resided in different geographical places and time-zones, we shared a training video using a private link hosted on YouTube[1]. In the video, we presented Emodash (see Figure 4.8), along with an explanation of the different components to help them get familiar with the tool.

Pair-formation. The pedagogical manager of the platform assigned the most suitable teacher for each learner based on both his/her English level and learning objective. We then asked participants to pick preferred time slots for the first session. Once a common time-slot worked for each pair, the learner received an automated email asking to fill out his/her profile

[1] https://www.youtube.com/watch?v=QwAS7 – 8JXQ

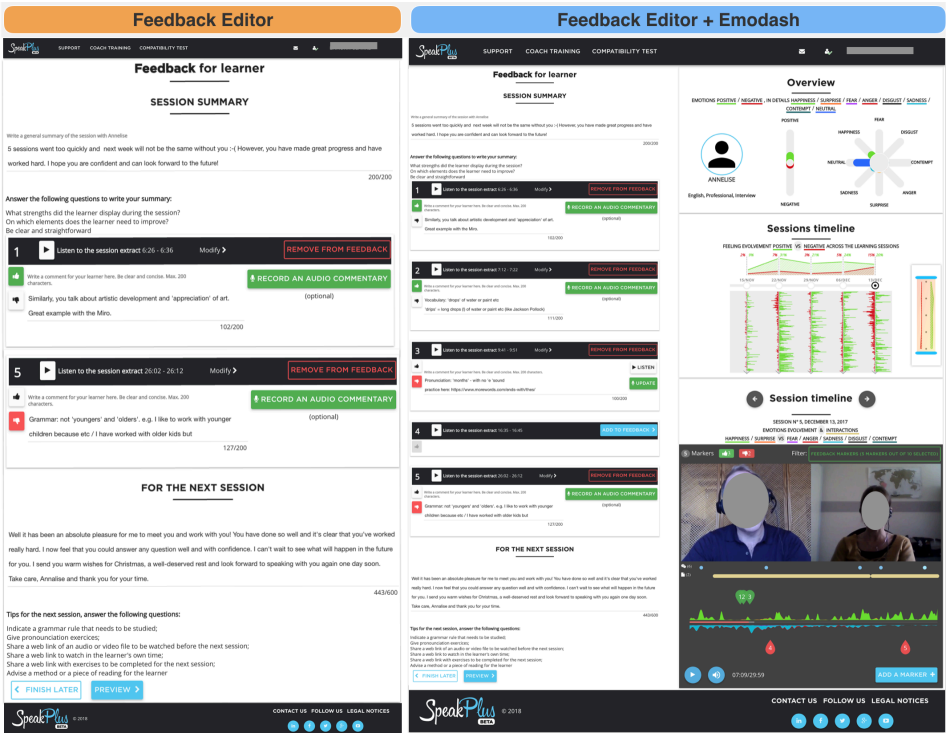


Figure 4.12: Left: SPEAKPLUS feedback report editor. Right: SPEAKPLUS feedback report editor with EMODASH.

and confirm the session. Moreover, the pedagogical manager contacted each learner by email giving them further details about how to get up and running with the platform.

Lessons and Feedback. Over the following weeks, each teacher met with his/her learner once a week for a learning session of 30 minutes (except one week where one pair met twice, at the beginning and at the end of the week). Learners booked their learning sessions with their teachers at their own pace.

Teachers were automatically informed by email when the feedback report was ready to be edited after the data processing and emotion analysis was complete (on average 40 minutes after the session). Teachers could then create their feedback reports using a dedicated editor (see Figure 4.12). Once the report was finished, teachers could share it with learners, who then received an email notification.

After the first session, we noticed that back-lighting issues had damaged the video analysis of one learner. We fixed this problem by post-processing the videos to improve contrasts, and re-ran the emotion analysis. We also sent an email to all learners with some tips to ensure proper video-

conferencing quality, mainly to avoid back-lighting, to make sure that their face was lit up, to sit facing the camera, and to make sure that the webcam and microphone worked properly before the session. We sent a similar email to teachers explaining the issue, and asking them to make sure to remind learners to avoid such issues.

Wrap-up. At the end of the experiment, a member of our research team set up online semi-structured interviews with teachers. These audio-recorded interviews lasted 20 to 30 minutes. At the end of the interview, each teacher was asked to fill out a SUS questionnaire (Brooke et al., 1996) to evaluate the usability of Emodash.

4.7.3 Apparatus

We deployed a dedicated web instance of the learning platform Speakplus with Emodash. Participants used their own devices from their location of choice (mostly home or work). Participants needed to have access to the Internet, a modern web browser (Chrome or Firefox), a camera, and a microphone for video-conferencing. More details about the (hard/-soft)ware used by participants are provided online in supplementary material (see *Setup*).

4.7.4 Data Overview and Analysis

Log and Screen Recordings. We collected teachers' interaction logs (clicks, scrolls, hovers, navigations, and playbacks) while they were editing the feedback report using the platform built-in trackers coupled with the *screen recording* tool Inspectlet[2]. We collected 20 (out of 23 sessions) screen recordings. One loss was due to an incompatibility of the recording script with the browser of one participant,while the other two losses were due to a storage issue.

Log and Screen Recording Analysis. One member of the research team analyzed interactions of teachers with Emodash while they were editing feedback reports. Observations were reported with a *title*, associated *level* of Emodash [Overview, History View, Session View], associated *visualization* if any [Balance, Radial, Multi-line, Mirror-bar, Brush, Stream, Marker, Chat, Doc, Navigation], associated *tag* [Scan (skimming), and Fixation]. A scan is a quick skim of a part of Emodash, while a fixation lasts more than 4 seconds. We used logs to measure interactions (clicks and hovers).

[2] https://www.inspectlet.com/

Feedback Reports. We collected the five feedback reports written by each teacher while using EMODASH, as well as the five most recent feedback reports they wrote *before* using EMODASH. The average length of feedback reports was 14.35 utterances (max=27, min=3, sd=7.56) per session with EMODASH and 13.56 utterances (max=25, min=1, sd=6.91) without EMODASH.

Feedback Report Analysis. We coded and analyzed the content of each feedback report with and without EMODASH:

- *Content Coding.* Our feedback coding is inspired by Tanes et al. (Tanes et al., 2011) content analysis. We analyzed the focus (summative vs. formative) of the content in order to study the extent to which teachers gave explanations in their feedback to learners. Summative feedback (e.g. "You can make improvements") provides information that situates individuals with regard to a set of criteria, whereas formative feedback ("I believe that success in interviews is attributable to practice and confidence") provides explanations on why tasks are correct or incorrect. We also focused on motivational and affective components of the content (Tanes et al., 2011). Motivational feedback (e.g. "Good job") provides positive or negative reinforcement regarding the learning activity. We distinguish affective feedback from motivational feedback as it is not directly related to the learning activity (e.g. "It was a pleasure to meet you").
- *Coding Unit.* Our coding unit was the utterance, considered as the smallest piece of a sentence with a clear ending (or pause), and a full meaning considered separately.
- *Inter-Agreement.* Two members of the research team coded separately a set of randomly selected feedback reports (20% of utterances), regarding the four categories: formative, summative, motivational and affective. Following which, both members discussed disagreement regarding utterance units before recoding the same sample with an inter-agreement of 66% for formative/summative utterances and 81% for motivational/affective ones. Then, they both discussed disagreements regarding each utterance to reach a consensus. Finally, one member coded the rest of the feedback reports.
- *Statistical Analysis.* We then conducted a factorial analysis on the feedback report content detailed in the results section.

SUS Questionnaire Analysis. We complied with Brooke et al. (Brooke et al., 1996) to compute the average of each participant before computing the overall average.

Interview Analysis. One member of the research team transcribed audio recordings of the interviews, before analyzing them to extract the main themes (Braun and Clarke, 2006).

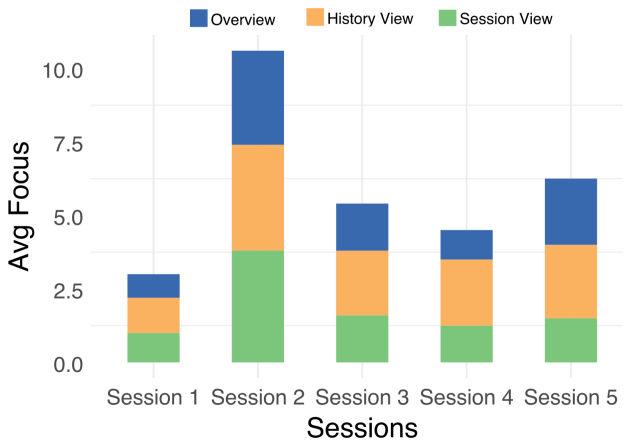


Figure 4.13: Average amount of focus (scans and fixations) per participant across sessions.

4.8 Field Study Results

4.8.1 Overall Use of Emodash

The results of the usability survey indicate that there was no significant problem with Emodash, with a SUS score of 80, 12 points more than the standard average score (68) (Brooke et al., 1996)

Teachers explored Emodash most during session 2 (measured in number of scans and fixations, Figure 4.13). This may be because of the email we sent between session 1 and 2 (see section 4.7.2).

Teachers' focus was spread across the three views of the dashboard. They interacted most with the session view, especially with the streamgraph, which controls the video player (Figure 4.14). The overview and history views are glanceable and did not offer significant interactive features. Participants interacted more in the first two sessions than at the end.

4.8.2 From Overview to Session Details: Interaction With Emodash

We compared teachers' perception of the three views of Emodash (overview, history view, session view), across the five sessions.

Overview. During the interviews, four (out of five) teachers highlighted the ease of use and usefulness of the overview to grasp learners' emotional

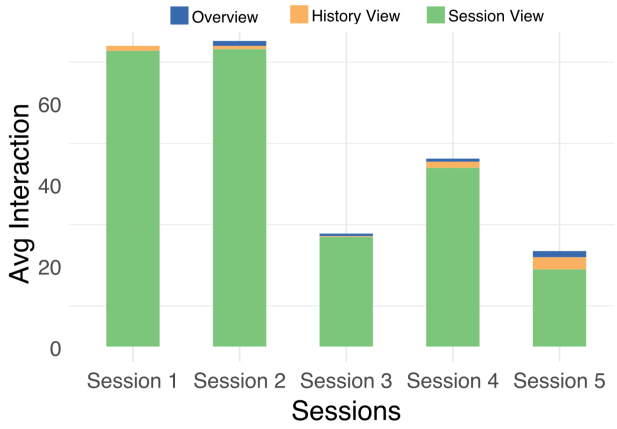


Figure 4.14: Average amount of interactions (hovers and clicks) per participant across sessions.

state (positive/negative) at a glance. On average, they focused on this 8.4 times (SD=6.98, P1: 7, P2: 9, P3: 3, P4: 20, P5: 3).

> The one I looked at most was the [overview] with the different emotions just because it is very easy to check. Is it more positive or negative? It's positive, great!. —P1

> I think the most [useful] is the overall map [meaning the Radial chart] because it's easy to understand. It shows the percentages, you know, it's color-coded, it's extremely easy to follow. I don't necessarily need the graphs of the session. —P5

History View. On average, participants focused 10.8 times on the History View, (SD=7.82, P1: 11, P2: 12, P3: 5, P4: 23, P5: 3). The history view was appreciated for providing a glanceable view summarizing the evolution of positive/negative emotions across all sessions, rather than for its details or for navigating across sessions.

> This is probably my favorite part [the History View], because it's very clear to see the positive and negative points of the session.—P3

> The history view gave you again the actual overall information of where the emotions went. Looking at it [...] shows quite clearly that there is a lot more look green on the second lesson than how it was in the first one, that said to me that I probably did better in the second lesson than in the first lesson. —P4

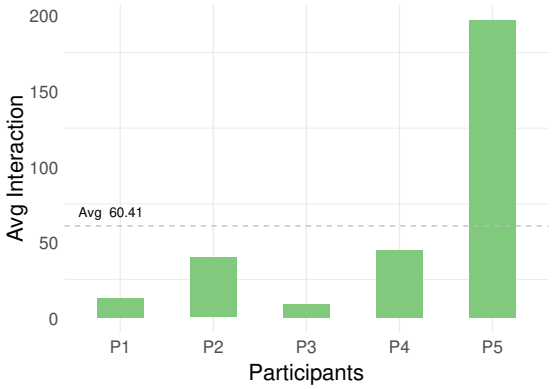


Figure 4.15: Average amount of interactions (hovers and clicks) per participant per session on the Session View of Emodash.

Participants did not interact with the history view to explore past sessions or navigate between them. Only one participant —P4, did so after editing the feedback report of his/her last session.

Session View. On average, participants focused 8.6 times on the Session View (SD=5.85, P1: 8, P2: 13, P3: 3, P4: 16, P5: 3). Moreover, they carried out 281 interactions in total (Mean=60.41, SD=70) with this view. Participant 5 interacted most (on average 181 interactions per session), P2 and P4 (on average 38, and 56 interactions per session respectively), and P1 and P3 only a little (on average 18, and 7 interactions per session respectively). Participants found this view interesting and relevant, even those who interacted little with the dashboard:

> The most interesting would be the Session timeline [...], because it makes it possible to go back to a particular point in time and look at that conversation again. —P1

Within the session view, participants clicked more on the streamgraph (compared to interactions with chats, docs, markers, and play/pause) (see Figure 4.15). In most cases (202 out of 211 clicks), participants navigated to emotional peaks (computed as clicks on areas of the streamgraph over the session's median value of the positive or negative emotions). Teachers used emotional data to navigate through the last recorded session, especially by clicking on these peaks, i.e. significant moments corresponding to high positive or negative emotions, and looked at the video to obtain further explanations on what happened at that particular moment.

Summary Notes: Participants appreciated the glanceable elements of the dashboard which gave them quick insights into learners' emotions. They

Table 4.1: Content analysis factors and levels.

Factor	Level
Condition (Emotional Dashboard)	- **With** (SPEAKPLUS + EMODASH) - **Without** (SPEAKPLUS)
Category (Feedback Content)	- **Motivational** - **Affective** - **Summative** - **Formative**

used the detailed session view to navigate in the video. From the interviews, we noticed a difference between the useful aspects (the overview), and the interesting aspects (the session view). Participants did not navigate to past sessions, and were mostly interested in the session they had just conducted.

4.8.3 Feedback Report Content

We analyzed the content of the feedback reports written by teachers after each session according to four categories of utterances: formative, summative, motivational, affective (see section 4.7.4). For each category, we compared reports written with EMODASH as well as the most recent feedback reports teachers wrote before using EMODASH.

We conducted a factorial analysis on the feedback report content. The study was a 2x4 repeated measures design (within-subject). The factors and levels are explained in Table 4.1. The measured variable was the number of utterances. We identified 287 utterances with EMODASH and 276 without it (Figure 4.16).

We analyzed the number of utterances using a non-parametric generalized mixed-effects model (GLMM) analysis (Dean and J. D. Nielsen, 2007; Bates et al., 2015), with fixed effects of *Condition* (emotional dashboard) and *Category* (Feedback report) and a random effect of *participant*. GLMM does not require the assumption normality of data (Wobbrock and Kay, 2016). We used each feedback report from each participant as a trial, which gives us enough data points to fit the model. GLMM manages the dependencies of the data within subjects. It also deals with missing data as we do not have a full balancing, for instance, a particular category (motivational, affective, summative or formative) may not exist for a given feedback report.

When building the models, we verified the five assumptions required to validate the good fit of the GLMM model (Bolker et al., 2009) using **R**. First,

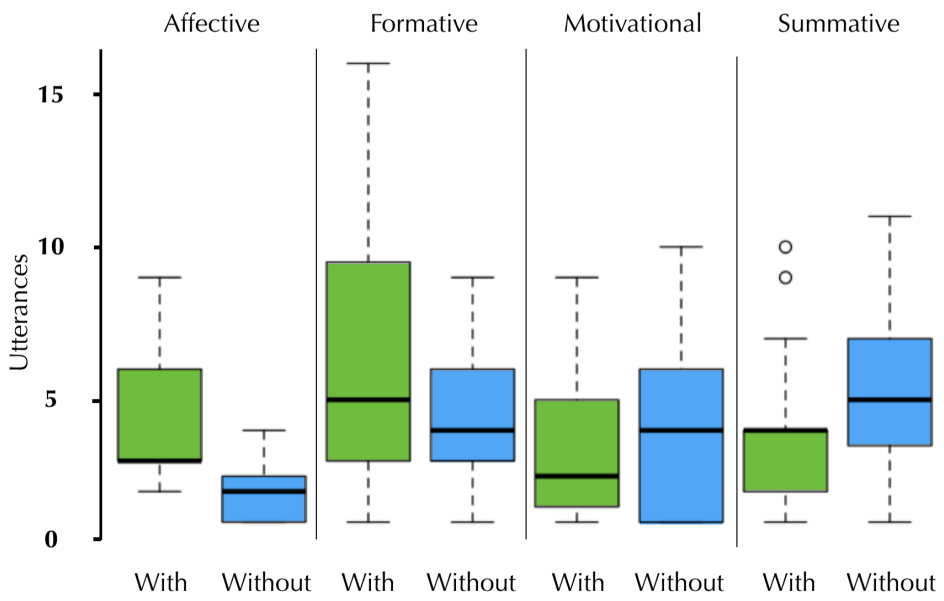


Figure 4.16: Boxplot comparison of feedback categories under both conditions with and without Emodash. With median, 1st quartile (25%) and 3rd quartile (75%). The whiskers represent the minimum and maximum.

we validated the goodness-of-fit of distributions of the data of each *Category* and under each *Condition*, which showed a tendency towards a *Poisson* distribution, which is usually the case with count data. We validated the goodness-of-fit graphically using histograms and statically using Chi-square test. Second, we graphically verified that standardized residuals of the GLMM model showed no departure from the normal using a QQ plot. Third, we graphically inspected the absence of autocorrelation in the residuals of the model to ensure that the amount of information that we have on the relationship between the number of utterances and our two factors (*Condition* and *Category*) is maximal. Fourth, we verified that the variance of the residuals is equal across levels of our factors, graphically using boxplots, and statically using Levene's test. Fifth, we verified that the model was not overdispersed using a Chi-square test, to ensure that the residual variance is not larger than the estimated mean. We provide the detailed statistical analysis as supplementary material (see *Statistical Analysis*).

We statically tested the significance of the fixed effects (*Condition* and *Category*) in the model using the Wald Chi-square test (Bolker et al., 2009). We find a significant effect of *Condition* ($\chi^2(1, N=157) = 4.21$, $P = 0.040$) as well as a significant effect of *Category* ($\chi^2(3, N=157) = 8.87$, $P = 0.031$) on utterances.

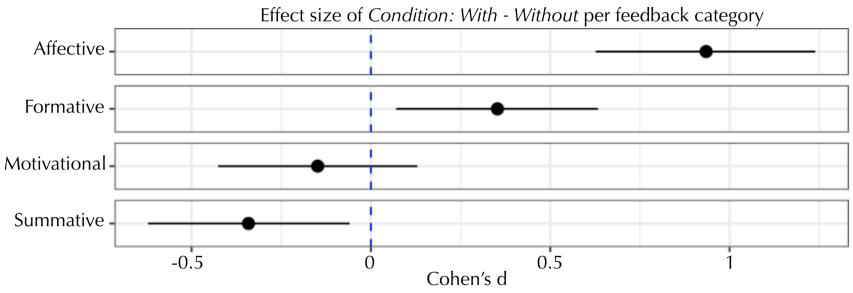


Figure 4.17: The effect-size of *Category*'s levels under *with - without* condition with the associated confidence intervals.

We conducted post-hoc pairwise comparisons to compare the amount of feedback per *Category* before and after the introduction of EMODASH (*Condition*) (see Figure 4.16 for a visual summary). We used Holm-Bonferroni for adjustments (Holm, 1979). We found significantly more *formative* utterances with EMODASH than without (Estimate = 0.351, SE = 0.141, z = 2.494, P = 0.0126, respectively 44% vs. 38%), and less *summative* ones (Estimate = -0.341, SE = 0.159, z = -2.139, P = 0.0324, 28% with vs. 43% without). Regarding *affective* content, the number of utterances increased with EMODASH (Estimate = 0.935, SE = 0.294, z = 3.176, P = 0.0015, 19% with vs. 10% without), whereas we did not observe a significant effect on the *motivational* level (Estimate = -0.146, SE = 0.176, z = -0.827, P = 0.4082, 23% with vs. 31% without).

We calculated Cohen's d, the standardized mean difference (Cohen, 2013) by dividing estimated differences by the residual standard deviation suggested as effect-size (Feingold, 2013). As portrayed in Figure 4.17, the results show a large effect for affective (d = 1.030), a medium effect for formative (d = 0.387), a (negative) medium effect for summative (d = -0.376), and a (negative) small effect for motivational (d = -0.160).

The differences between formative and summative, with and without EMODASH, appear to be stable across sessions (see Figure 4.18).

Figure 4.18: Percentage of Motivational, Affective, Summative and Formative Utterances per session, without and with EMODASH

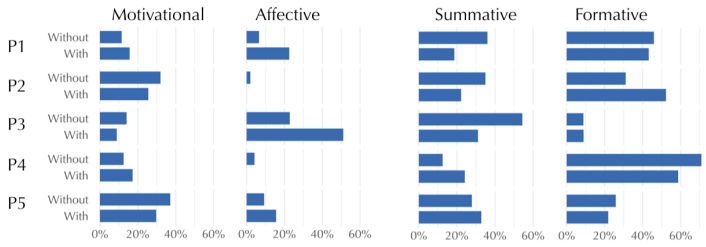


Figure 4.19: Percentage of Motivational, Affective, Summative and Formative utterances per Participant, Without and With Emodash.

The percentage of affective utterances tends to increase with Emodash over time, while the percentage of motivational utterances with Emodash tends to decrease over time.

Focusing on individual teachers' feedback, we observed different ways in which Emodash influenced (or not) their reports. P1 and P3 are those for whom Emodash had the strongest impact in terms of affective utterances per report (twice as many with Emodash) (Figure 4.19), where the amount of affective language in their reports increases over the sessions (Figure 4.18). This relates to far less summative utterances and to stable formative ones. We notice that these two teachers —P1 and P3, with P5, already wrote affective utterances in their feedback reports before using Emodash. It is also noteworthy that P1 and P3 are the teachers who interacted least with the streamgraph. P2 and P4 wrote few affective utterances before and even less with Emodash. These two teachers, with P5, also interacted most with the dashboard, but these interactions did not have the same effect. P2 wrote less summative and more formative feedback, whereas P4 wrote more summative and less formative feedback. For P5, the content of the feedback is more stable with and without Emodash.

Summary Notes: Teachers wrote more formative (and less summative) feedback reports with Emodash. They also wrote more affective feedback reports. This is especially true for teachers, already sensitive to learners' emotions, who incorporate more affective elements in their reports, suggesting a stronger awareness of learners' emotions. We did not observe changes in motivational utterances.

4.8.4 Impact of Emodash on Teachers' Practices

Besides feedback reports, during our interviews teachers reported on the impact of Emodash on their teaching practices, especially in considering their own emotions and assessing their own work.

Self-Awareness. Although the dashboard focused on learners' emotions, teachers reported an increased emotional self-awareness after using EMODASH. Indeed, P1, P2, P4, and P5 used the dashboard while editing their feedback and reported looking back at parts of the session to focus on their own behavior.

> It was quite useful to see the video as I was making the feedback comments and to see myself again on the screen. —P2

But more broadly, EMODASH also served as a reminder that learners' emotions have an impact on learning. Which could be forgotten, as P1 remarked:

> It actually reminded me of the emotional contingency, emotional parts of the student [...] usually I'm listening to accent pronunciation, grammar, their learning skills, but I didn't really think much about the general emotion, are they nervous, are they scared, are they worried, I just usually disregard that, but EMODASH reminded me, but wait a minute this a living, breathing, exchanging with emotion. —P1

Self-Evaluation. Teachers started reflecting on their teaching practices and how they were conducting their lessons, for instance, to get feedback on how they were conducting their class:

> As a teacher I always want to have some kind of feedback to show that I'm doing a good job and it kind of clarifies that I'm not doing too bad [...] —P4

Similarly, two participants mentioned using EMODASH as a way to evaluate their performance or as a warning when negative reactions are expressed by learners:

> I will definitely be concerned if I saw some, you know, a huge percentage of negativity. I do want to do better as a teacher. —P5

> I wanted my students to be happy all the time and to be very satisfied with the class and when I saw negative reactions in the overview or in the [History View], it worried me. —P1

Perceived Limitations. Participants identified two classes of limitations. The first one is the ability to identify emotions with a high intensity without the support of technology. P2 was most negative regarding Emodash, considering it "interesting but not particularly useful". She elaborated further:

> As far as teaching is concerned, I don't think it is particularly useful. Because I think [...] if there are any extreme emotions, if someone is feeling very angry or disappointed you should get that feedback without having to see this [meaning Emodash]. —P2

The second significant concern formulated by other teachers was the amount of work that Emodash involved. They expressed shortcomings regarding the time and efforts required:

> It was really extra time on the top of teaching. —P1

> What you show on the dashboard, it's quite detailed information. Also I don't think you will have the time, because already you have to plan lessons, do the feedback, and that intensity of feedback takes your time, it will be very heavy workload. —P2

The increased focus on emotions, and its relationship with teaching quality also led to "*concern*" with a negative emotional reaction from learners — P1. Moreover, the fact that teachers worked on a platform probably further increased the role of emotions, with teachers wanting to please learners.

4.9 Discussion

4.9.1 Learners' Emotion Awareness and Teachers' Formative Feedback

Regarding the impact on summative and formative feedback (research question #1), we observed that using Emodash led teachers to take into account learners' emotions and adjust their feedback by reflecting on the emotions identified during the learning session. Emodash led teachers to give proportionally more formative than summative feedback, providing learners with explanations on their performance or cues on how to improve themselves. This may help learners to maintain or improve their performance during and across the learning sessions (Tanes et al., 2011), as formative feedback is more effective for learning due to its corrective

nature (Chen, 2001; Higgins et al., 2002). This observation raises important questions as to the impact of teachers' awareness of learners' positive or negative emotions. Also, more studies should be conducted to identify the factors at stake on larger scales.

For instance, as suggested by Montero and Suhonen (2014), there could be *human biases* in how teachers interpret learners' emotions and their personal opinion of them. A learner who felt mainly negative emotions could be perceived negatively by the teacher, even if these emotions are not due to his/her interactions with the teacher but to other contextual reasons. Given that some emotions are difficult to identify based only on facial recognition (Visschedijk et al., 2013), computational emotional information should be treated with caution.

4.9.2 Learners' Emotion Awareness and Teachers' Affective Feedback

Concerning the influence of emotion awareness on affective and motivational feedback (research question #2), we did not observe a significant impact on the proportion of motivational feedback. This result could be accounted for by the fact that all teachers had already incorporated motivational utterances in their reports and may not need to be aware of learners' emotions to integrate this kind of feedback. It could also suggest that teachers saw no difference in learners' outcomes by integrating more motivational feedback. In fact, Tanes et al., 2011 showed that students who received considerably less motivational feedback had no significant difference in their outcomes, compared to students who received more motivational feedback.

Teachers wrote significantly more affective feedback, which should improve learners' receptivity (Walz, 1982) and language acquisition. Specifically, for the second language acquisition domain, Dulay and Burt, 1977 first introduced the concept of the "socio-affective filter" to explain that, in language learning, the acquirer must not only understand the language to acquire it but must also, in a sense, be "open" to it. According to this concept, more affective language from the teachers could help learners acquire a lower filter and so acquire more of the language directed at them. Our results would thus suggest that better awareness of learners' emotion would lead to better language acquisition.

Examining the impact of the type of feedback during an online writing task, Nguyen et al., 2017 also showed that positive affective language increased positive emotions and reduced participants' annoyance and frustration, thus leading to an increase in work quality. Previous studies also showed

that positive affective language can increase positive emotions and reduce negative feelings when receiving criticism (Neuwirth et al., 1994). This would suggest that integration of more affective language from teachers in their feedback could lead learners to feel more positive emotions.

Finally, the proportion of affective language increased over the sessions, thus suggesting that the teacher-learner pair developed socio-affective relationships, which was our original motivation when developing Emodash. More specifically, the impact was observed for teachers who already wrote affective parts in their reports without Emodash. This would suggest that Emodash reinforces existing practices rather than creating new ones. Further studies would be needed to explore this in greater depth.

4.9.3 Designing Dashboards for Emotion Awareness

For teachers, paying attention to learners' emotions comes on top of many other tasks: evaluating learners' interventions, structuring feedback, following the lesson plan, guiding learners, sharing documents, marking moments for later feedback, etc. As mentioned by one participant, one would ideally expect that emotion awareness happens "naturally", but in practice this is challenging.

Participants showed favorable views towards Emodash, although they used it in a variety of ways, exploring different levels of the dashboard while writing their feedback, or exploring the data out of curiosity. Regarding our third research question, we can draw some conclusions as to the appropriate level of detail and timescales of learners' emotions.

First, our results suggest that an extensive historical record and visualization of emotions may not be useful in our context. Teachers seemed more interested in reflecting on the last session and used the dashboard for that purpose. This means that collecting and analyzing emotion data over the extended period of a course may not be necessary or appropriate in all contexts, although some identify it as a future theme in this area (S. K. D'Mello, 2017).

Second, teachers tended to want their feedback to reflect the general emotions identified for the learner during the last session, where general emotions mean "positive" or "negative", rather than more detailed discrete emotions. While this leads us to question the need to provide teachers with details on specific emotions, it rather suggests offering broader affective states as in (Noteborn et al., 2012; Mega et al., 2014).

Third, participants interacted mainly with the streamgraph to navigate in the video recording by clicking on emotional peaks, whether positive or negative. However, it would seem that this navigation did not lead teachers to change their way of writing feedback. In fact, teachers who navigated most in the session view —P5, had the same kind of content in their reports with and without EMODASH, whereas teachers who interacted least —P1 and P3, adapted the content of their feedback reports with EMODASH.

These elements suggest that teachers may need more at a glance emotional information than detailed navigation in the past learning session. Precise and information-rich visualization of emotions may not lead to changes in the way teachers write feedback. Emotion awareness may not be directly linked to a specific kind of data or visualization but rather to the general availability of information about emotions. The presence of EMODASH reminds teachers that emotions should be considered in the learning process.

Emotions in online learning environments should be studied not only as insights into cognition, commitments and concerns (Boekaerts, 2010) but also in light of the concept of "motional presence" introduced first by Cleveland-Innes and Campbell (2012) and defined as *the outward expression of emotion, affect, and feeling by individuals and among individuals in a community of inquiry, as they relate to and interact with the learning technology, course content, students, and the instructor*. Emotional presence should be considered as a distinct construct, apart from cognitive, social, and teaching presence, that encourages social interactions between learners and teachers (Stenbom et al., 2016). P1 illustrated the relevance of emotional presence by explaining that:

> Usually, I focus on progress: Did she understand? Did she learn from that point? Which is not connected with emotion, but EMODASH made me aware of the other parts of teaching which is, is she pleased, is she happy to learn. —P1

4.9.4 Self-reflection and Evaluation

We designed EMODASH to support teachers in developing a stronger awareness of learners' emotions. We found that it also increased teachers' awareness of their own behavior, leading them to reflect on the feedback they gave in reports and in other sessions.

Given the lack of direct feedback from learners to teachers, teachers also used EMODASH as a proxy to evaluate their work, to know that they are "*doing a good job and [...] not doing too bad*" —P4. We find nonetheless that

emotion awareness tools should be treated with caution. Teachers often related positive emotions to the quality of the teaching session, although the literature insists on the benefits of negative emotions such as confusion (S. D'Mello, Lehman, et al., 2014) or anxiety (Pekrun, Goetz, et al., 2011) on learning. Incorporating some guidance on the role of emotions in education could mitigate these problems.

4.10 Limitations

We conducted the study with five pairs of learners and teachers over the course of two months. This study is limited in the number of participants, and the five learners were female. A larger-scale between-subject experiment could complement our results. While limited, this is the first in-the-wild study to shed light on the use of learners' emotions in learning dashboards.

Prior research focused on awareness of learners' performances and/or behaviors, that are easier to capture and to remediate (Schwendimann et al., 2017; Verbert et al., 2014). We are opening up research on the use of emotions to enhance teachers' reflection on learning processes. Yet further research is needed to explore how such dynamics differ and evolve over time and whether there is any correlation with learners' receptivity of the feedback.

Our work builds upon emotion recognition algorithms. Emotion recognition is still a very active area of research, in which contextual, cultural and technical challenges remain to be addressed. As our focus is on the impact of retrospective emotion awareness on teachers' feedback, we were less concerned with the correct identification of precise emotions at any given time, but rather with providing trends and emotional indications at different levels and timescales. This explains our use of a standard emotion recognition API for reliability purposes, over emotions more related to learning situations, such as Pekrun's achievement emotions (Pekrun, 2006; Pekrun, Lichtenfeld, et al., 2017).

More profoundly, Boehner et al. (2007) criticized the use of purely automated emotion recognition techniques, and emphasized the risks of disconnect between what is detected and what is actually felt by people. They argue that emotions are experienced through interactions and are influenced by cultural and social factors, i.e., emotions are built and evolve through interactions between people. As such, the interpretation of emotions evolve, change and differ depending on people's context. We partly agree with this remark, but argue that some strong emotional reactions can

be properly interpreted with little context. Nonetheless, our work adopts a hybrid approach: while emotions are inferred from the recordings, and conveyed visually, we preserve and provide the context (e.g. video recording, the associated events, etc.) of the experienced emotions of the user in order to enable a better interpretation by teachers.

Finally, throughout the trial, and in our final interviews, none of the participants flagged recognition problems, although some teachers —P2, P3, discussed recognition and expressed curiosity as to how emotions were interpreted by the application. Similarly, in a previous study with an emotional dashboard using the Microsoft Azure API, teachers did not raise issues related to the use of such an API (Ez-zaouia and Lavoué, 2017).

Finally, EMODASH could lead to anxiety on the part of the learners, bearing in mind that every emotion they express will be recorded and analyzed. From observing the videos, we observe that the awareness of the recording disappeared after a few minutes. Considering larger-scale deployments, our results suggest that detailed recordings may not be necessary, and that aggregate summaries may be enough to raise teachers' awareness.

4.11 Summary

The development of richer socio-affective relationships between teachers and learners is an important component of successful learning processes (and outcomes). However, developing such relationships is challenging in online-learning. To tackle this challenge, we aimed at fostering teachers' awareness of learners' emotions. We designed EMODASH a dashboard visualizing learners' emotions alongside the activity of past learning sessions. The design process involved several iterations and the involvement of experts: professional teachers and pedagogical managers. We integrated EMODASH in the feedback editor of a video-conferencing learning platform, to support the emotional awareness of teachers as they provide feedback to learners.

We conducted a field study for 8 weeks with 5 pairs of learners and teachers to investigate the role EMODASH played on teachers' feedback. We focused on three questions: 1) Do emotion awareness tools have an effect on the type of feedback given, specifically formative or summative feedback reports?; 2) Do emotion awareness tools lead teachers to incorporate more affective and/or motivational content in their feedback? and 3) What effect do different granularities of emotion visualization, via a dashboard, have on teacher feedback?

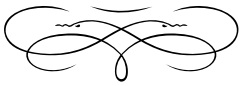
EMODASH led teachers to write more formative and less summative feedback and to incorporate more affective language in their report. Teachers who interacted least with the dashboard are those who changed their reports most, suggesting that an overview of learners' emotions may be sufficient to better consider learners' emotions. Teachers seemed to favor quickly graspable visualizations, although important information might get lost. New kinds of visualizations could provide an overview and highlight important moments without overwhelming details. Indeed, teachers also relied on EMODASH richer streamgraphs, pinpointing *positive* and *negative* emotional peaks, to navigate in the video recordings of past sessions. Finally, teachers also described how visualizations led them to reflect on their practice, and acted as a way to acquire information on the quality of their teaching. Overall, it led them to adjust the way they shared feedback with learners.

Our study opens up questions for future research. In terms of visualization: How to design glanceable summaries that still convey nuances about learners' emotions? And how to better place emotions in context, i.e., compared to past sessions or other learners, to support teachers' interpretation of emotional data? While we observed some changes in teachers' practices, these observations need to be studied and analyzed more thoroughly: to what extent awareness tools reinforce existing practices rather than create new ones? Also, what is the interplay between teachers' self-reflection, awareness of their own emotions, and awareness of learners' emotions? Finally, we focused on teachers although our end goal is better learning processes and outcomes for learners. Further studies on how learners receive affective feedback fostered by emotion awareness tools would close the loop from learners to teachers and back.

Use of a video-conferencing based learning platform enabled us to have good quality videos of learners' faces. Advances in emotion recognition could provide further opportunities to deploy affective computing tools in schools: for instance, to inform teachers about the affective state of students and help them provide personalized support to students in a timely fashion. In addition, affective states can also be used as a proxy to infer indicators about students, such as involvement and engagement, as well as learning-related emotions such as frustration, confusion, boredom while students are engaged with learning tasks (Pekrun, 2006; B. Kort et al., 2001). This could have side benefits for both teachers and schools, such as monitoring students' satisfaction and well-being during courses, providing directions to improve course design and learning activities, and for teachers' own professional development (Montero and Suhonen, 2014).

Chapter 5

Assisting Teachers' Data-informed Practices in Online Learning Using a Dashboard



> *Visible teaching and learning occurs when there is deliberate practice aimed at attaining mastery of the goal, when there is feedback given and sought, and when there are active, passionate, and engaging people (teacher, students, peers) participating in the act of learning. [...] What does matter is teachers having a mind frame in which they see it as their role to evaluate their effect on learning.*
>
> —Hattie, 2012, p. 53:55

5.1 Overview

Teachers blending in-class and remote online learning activities face numerous challenges in monitoring learners' online activity and progress, especially to inform their in-class interventions. We present a design study of PROGDASH, a dashboard enabling teachers to monitor learners' activity in an online grammar and spelling learning platform. We designed PROGDASH based on interviews with seven teachers to understand their needs and collaboratively iterate on design prototypes. We implemented and integrated PROGDASH in a commercial online learning platform. We conducted a three months-long field deployment with 29 teachers collecting logs, diaries, and interview data to evaluate how PROGDASH informed their practices. Our findings extended prior work on teachers' practices using dashboards: teachers found PROGDASH actionable to devise informed pedagogical practices: monitoring, conducting assessments, planning interventions, sharing in-class, providing debriefing and feedback. Based on our findings we suggest directions for dashboards to effectively support teachers in schools using online learning platforms.

5.2 Motivation

Schools are increasingly integrating web-based learning environments with in-class learning activities (Technavio, 2016). Teachers leverage these platforms, to support in-class learning, help learners develop skill-sets that should already be mastered, or to provide personalized support to learners with specific needs, e.g., dyslexics.

Teachers mixing in-class and remote online learning activities face numerous challenges in monitoring learners' online activities, especially when informing their in-class interventions. Dashboards can be efficient tools for facilitating sense-making, reflection, and making informed decisions (Verbert et al., 2014). Prior work on virtual learning environments, used mainly in-class, has shown the potential of dashboards to support teachers, for instance, to improve their awareness of their classrooms (Holstein et al., 2018), to support their planning tasks and interventions (Xhakaj et al., 2017), to support learners with lesser abilities (Molenaar and Campen, 2017), and to provide personalized support to learners in a timely fashion (Aslan et al., 2019).

However, with online learning, the individualized pacing of learners makes it difficult for teachers to monitor learners' progression before engaging in formative in class interventions. The way teachers leverage

dashboards in the long run and how dashboards shape their pedagogical practices is still unclear. Two recent systematic reviews of more than 150 learning analytics dashboards emphasize the need for more longitudinal studies to investigate how a dashboard might inform and impact teachers' practices (Bodily et al., 2018; Schwendimann et al., 2017).

In this chapter, we present a design study (Sedlmair et al., 2012) and a longitudinal field evaluation of PROGDASH, a dashboard to assist teachers in monitoring learners' progression on an online grammar and spelling learning platform. We designed PROGDASH based on interviews with seven teachers to understand their needs, and we collaboratively iterated on design prototypes (Dow et al., 2011). We implemented and integrated PROGDASH in a commercial online learning platform called PROJET-VOLTAIRE[1]. We then conducted a three-month field deployment with 29 teachers, collecting logs, diaries, and interview data to evaluate how PROGDASH informed and impacted their practices in blended learning. Our results show that most teachers reacted positively to the dashboard. They found it actionable to inform several pedagogical practices: monitoring, conducting assessments, planning interventions, sharing in-class, providing debriefing and feedback. Our results further highlight a divide between teachers specialized in the concepts of the learning platform (i.e. French language) and those in disciplines requiring the skills (e.g. Marketing), even if both were concerned with achieving the same objectives, mainly helping learners master spelling and grammar. Based on our findings, we provide design implications aimed at improving dashboards to bridge online and in-class learning and foster learners' self-reflection.

The contributions of our work are:

1. A design study of PROGDASH, a dashboard designed based on interviews with seven teachers and refined through collaborative design prototypes, to bridge between learners' use of an online grammar learning platform and teachers' data-informed pedagogical practices.

2. An in-the-wild field study of PROGDASH. This study extends our understanding of how teachers integrate a dashboard in their practices to articulate remote and in-class learning.

3. Design implications for dashboards to bridge online and in-class learning and facilitate data-informed pedagogical practices.

[1] https://www.projet-voltaire.fr

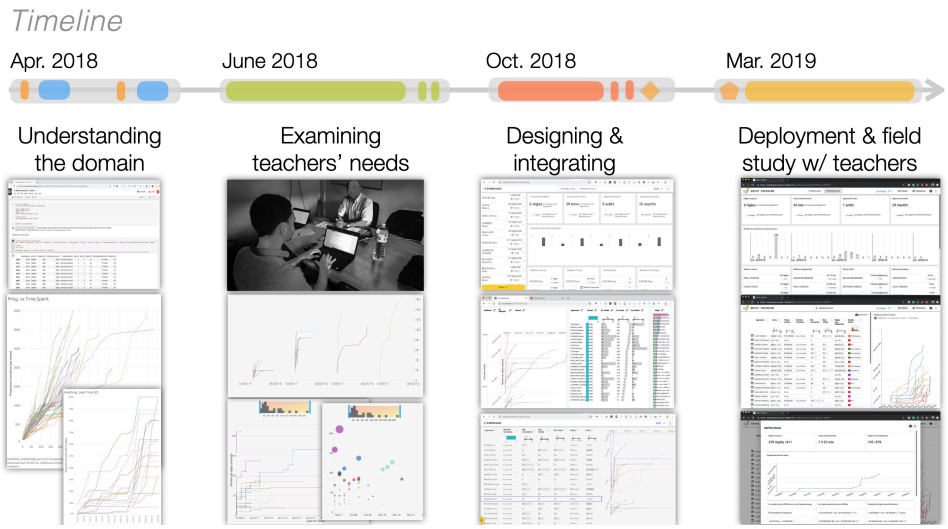


Figure 5.1: Key steps of PROGDASH design process.

5.3 Design Process

As Figure 5.1 illustrates, we followed a teacher-centered design study process (Sedlmair et al., 2012) in four steps:

1. **Understanding the domain.** We discussed with PROJET-VOLTAIRE designers to understand the the platform, the exercises, the learning process, and its data, see Section 5.4. We gathered real data to prototype realistic visualizations.

2. **Understanding teachers' needs.** We interviewed seven teachers to identify their needs for monitoring and understanding learners' progression, and to explore design alternatives, see Section 5.5.

3. **Design, implementation, and integration.** We iteratively designed PROGDASH, which we then enhanced through several informal discussions with members of the company and teachers from our first interviews. We deployed a stable design of PROGDASH to the company's R&D team for usability feedback, see Section 5.6.

4. **Field deployment.** We conducted a three-month field study with 29 teachers to capture teachers' experience/perception of PROGDASH, and how it may assist them in informing their practices, see Section 5.7.

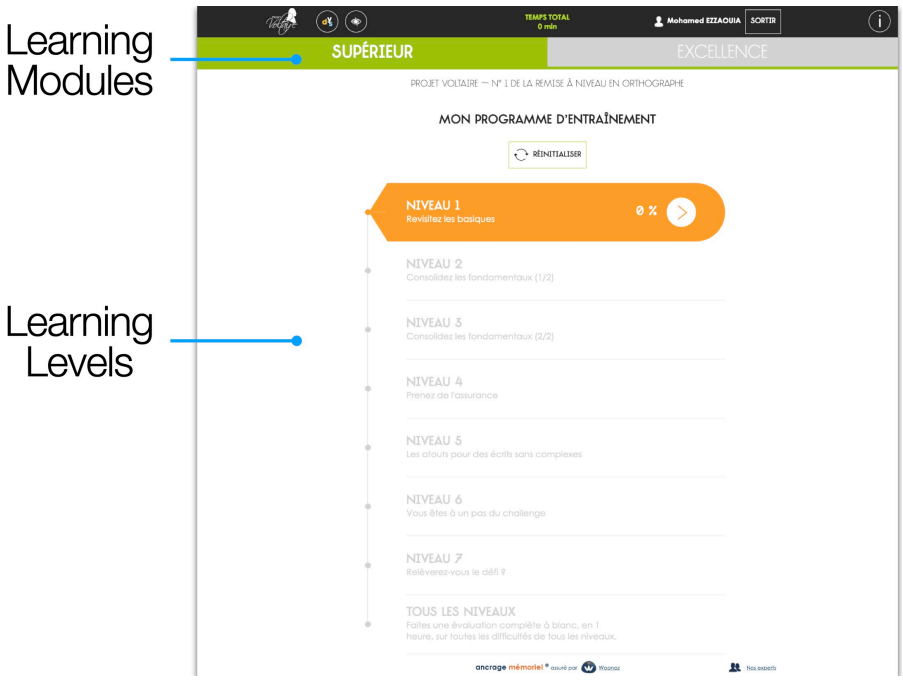


Figure 5.2: Learning modules in Projet-Voltaire.

5.4 Projet-Voltaire's Platform

We collaborated with Projet-Voltaire, an online platform for learning French vocabulary, grammar and spelling rules (Figure 5.2). The platform is widely used by French public and private schools (—Primary, middle, high school and university). It can be used autonomously or in-class, either projected on a whiteboard with class exercises, or on personal computers.

5.4.1 Learning Process

The platform centers around an exerciser that adapts to learners' skills in vocabulary, spelling, and grammar. Learning is structured around modules corresponding to a given level of expertise. Each module has a set of stages that learners must go through to master the level. A stage has a set of grammar/spelling rules, corresponding to the concepts that learners need to master. Each rule has a set of exercises. Figure 5.3 shows typical questions on Projet-Voltaire. Teachers can also set up class quizzes, i.e., short evaluations, typically carried out at the beginning or the end of a course. Projet-Voltaire collects the pertinent learning traces on learners (see Table 5.1), which we used to build Progdash.

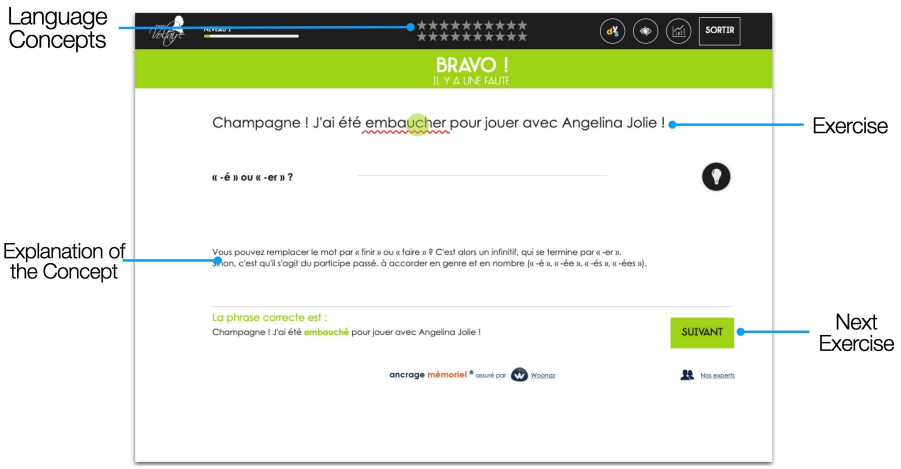


Figure 5.3: Example of exercise in PROJET-VOLTAIRE.

Table 5.1: Data collected on PROJET-VOLTAIRE.

Attribute	Description
Module	Learning module(s) a learner is enrolled in
Stage	level of difficulty of the module
Rules	to master in each module
Exercises	associated to a rule
Answers	exercises' answers (correct/incorrect)
Time-spent	Time-spent to answer each exercise of a rule

5.4.2 Learning Data

Prior to our study, teachers had access to a reporting table (see Figure 5.4). The table listed learners with a set of indicators split across multiple columns: time spent, score (%) of learner's mastery, initial level, scores and time spent on the programmed quizzes, last connection, and rules acquired. Teachers can also download a pdf report about each learner or export the entire table as a spreadsheet file.

Figure 5.4: Example of a class report.

5.5 Teachers' Needs, and Design Goals

We conducted nine interviews with seven teachers to understand (1) how they use the learning platform, and (2) how they monitor learners' pro-

gression. We recruited teachers via the company's education customers' team, and they received a book on the French language as a thank-you gift. Interviews were audio-recorded, and later transcribed and analyzed by the main author to identify requirements in supporting teachers.

Context of use We identified three ways of how teachers integrate Projet-Voltaire in their classrooms. First, as a *learning material*, they incorporate platform-based activities in their curriculum with a dedicated time-slot. Second, as a *prerequirement aid*, where teachers push learners to use the platform alongside their course. And, finally, as a *personalized aid* for specific learners, such as learners with grammar-related disabilities, e.g., dyslexia, or with very poor grammar skills.

Learning progression Teachers characterize learners' progression as (1) the *levels and modules* that learners have reached, (2) the *regularity* of their practice, and (3) how *skills* develop over a period of time. They highlighted the temporal nature of the learners' progression, this is "hidden" in the existing reporting table which displays only learners' current state. This causes learners' progression —in sense of, e.g., unique pathways, discrepancies, common strategies, similar struggles; to become difficult for teachers to grasp. Teachers reported relying on informal discussions with learners to infer their progression and regularity of using the platform.

Informed interventions Teachers often conduct interventions in-class to congratulate, encourage, remind, or alert learners. They highlighted the complexity of using the existing reporting table to keep track of learners' online activity and to pinpoint gaps in learners' progression, which hindered engaging in informed interventions with learners, e.g., acquiring arguments for debriefs and feedback. Also, when preparing a French class, teachers expressed interest in knowing which grammar concepts were not mastered by learners or those that proved most tricky in order to tackle them in class.

We derived the following design requirements from our interviews:

R1: Offer at-a-glance indicators to facilitate monitoring of learners' activity at class level by providing summaries of learners' overall practice, engagement, and regularity in using the learning platform.

R2: Provide indicators about learners' practice to facilitate close monitoring at learner or group of learners' level. Teachers emphasized the need for practicable information enabling them to keep track of the status of learners' online activity on the platform.

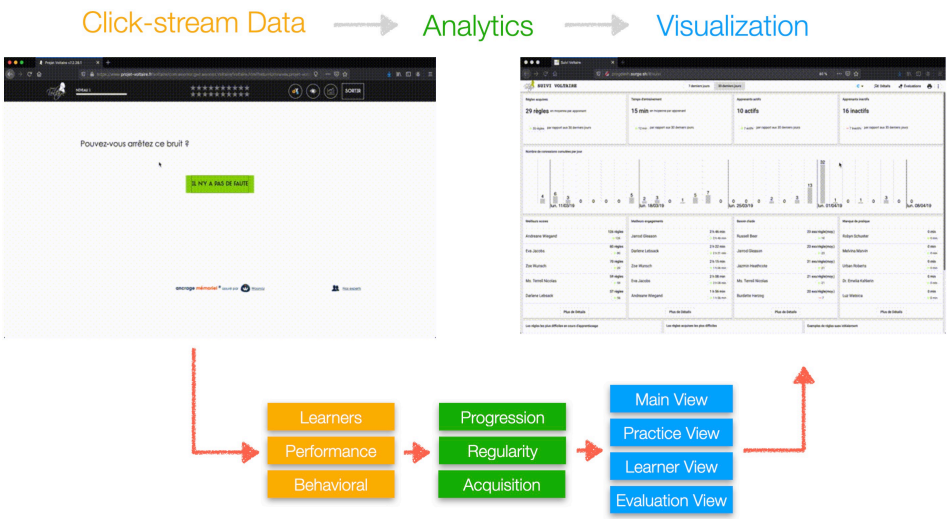


Figure 5.5: Overview of PROGDASH's architecture.

R3: Incorporate indicators about skills, i.e., grammar concepts that are already mastered or prove challenging. In order to prepare class interventions or to help individual learners, teachers expressed the need to understand what they should focus their interventions on, based on learners' difficulties.

R4: Provide information about learners' progression over time. Teachers expressed the wish to grasp the progression and regularity of practice over time. Learners' progression is multifaceted and depends on how teachers and learners engage with the learning platform. Relevant indicators include grammar concepts mastered, time spent, and amount of exercises practiced.

5.6 PROGDASH User Interface

Figure 5.5 illustrates the architecture of PROGDASH, namely, the types and relationships between learning data collection, analytics, and visualization. The main challenge in designing PROGDASH was to visualize all the information required in a small visual space in order to clearly and directly inform teachers about the state of learners' progression. Following *"overview first, zoom and filter, then details on-demand"* (Shneiderman, 2003), we structured the information in four different views so that teachers can drill down for more information as needed. We adopted statistical and visual summarization techniques (A. Sarikaya et al., 2018) to present relevant indicators.

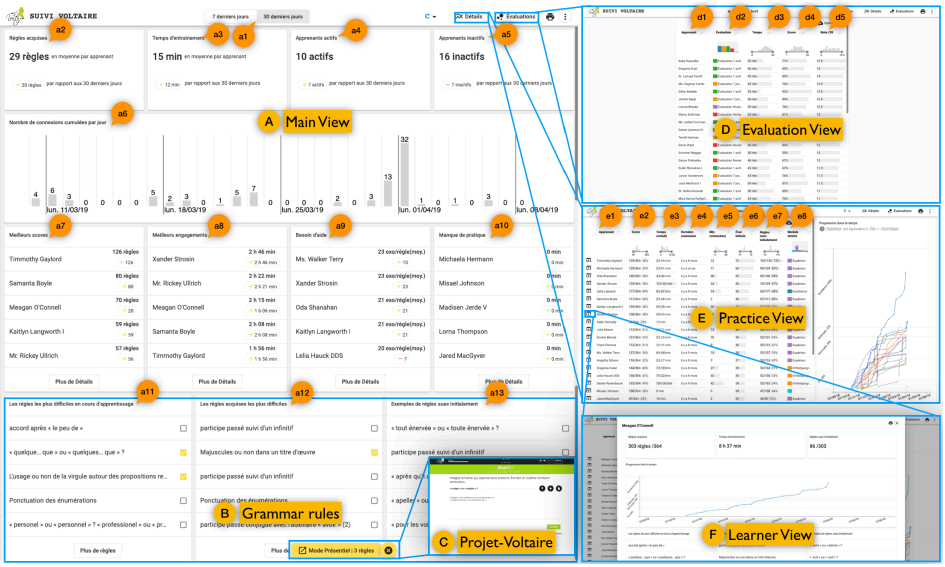


Figure 5.6: Progdash User Interface.

Progdash compiles metrics over the last 7 or the last 30 days, and teachers can pick between these two temporal windows. We combined textual, visual, and color encoding when designing the dashboard to facilitate glanceable reading of the indicators. We provided comparison and contextualization of Progdash 's indicators by using bars to encode cells in tables and line-charts to visualize learners' progression paths. We also added histograms and bar-charts in the headers of tables' columns to facilitate filtering and searching. We provide supplementary online material about the dashboard in this link[2].

Main view —as portrayed in Figure 5.6.A, presents overall information to monitor learners' activity at class level (**R1**). It presents 12 metrics in a grid compiled for the selected time window —7 or 30 last days, see Figure 5.6.a1. The grid is structured around three topics on learning regularity, learners' progression, and grammar rules acquisition, described in Table 5.2. The first row of the grid provides the average number of acquired rules per learner, the time spent as well as the total number of active/inactive learners. The same widget is used to show the name, the value, and the associated (+/-) delta —change since the previous time window (7 or 30 days), of each metric. The second row shows a timeline bar-chart of learners' connections per day. The third row displays learners' related metrics: score, commitment, help, and dropout. The same widget is used to present the full name of the learner, as well as the associated value and (+/-) delta. The last row, as Figure 5.6.B portrays, presents grammar rules related metrics: in focus —currently in learning by learners, acquired (already learned), and known (already known). The rules are sorted by their

[2] http://progdash.surge.sh/

Table 5.2: Main View's indicators within 7 or 30 days.

Attribute	Description
	Learning Regularity
Grammar Rules	Avg. number of mastered grammar rules per learner, with the associated (+/-) delta of the selected time (Figure 5.6.a2).
Time Spent	Avg. time spent learning per learner, with the associated (+/-) delta of the selected time window (Figure 5.6.a3).
Active Learners	Number of active learners, with the associated (+/-) delta. Learner is considered active if s/he makes at least one (in 7 last days) or 4 (in 30 last days) connections (Figure 5.6.a4).
Inactive Learners	Number of inactive learners, with the associated (+/-) delta. Learner is considered inactive if he/she makes zero (7 last days) or less than 4 (30 last days) connections (Figure 5.6.a5).
Connections Timeline	Total some of unique connections per day (Figure 5.6.a6).
	Learners Progression
Score	Top five learners in grammar rules acquisition (Figure 5.6.a7).
Commitment	Top five learners in practicing on the platform (Figure 5.6.a8).
Help	Top five learners struggling the most (Figure 5.6.a9).
Dropout	Top five non-practicing the most (Figure 5.6.a10).
	Grammar Rules Acquisition
In Focus	Top five rules currently in acquisition, and leading to struggles to most learners (Figure 5.6.a11).
Acquired	Top five rules already acquired, although led to struggles to most learners (Figure 5.6.a12).
Known	Example of five rules already known by most learners (Figure 5.6.a13).

level of difficulty exhibited by learners' practice. Teachers can select a set of rules and launch **Practice view** to practice them with learners live in-class (**R3**), as Figure 5.6.C portrays.

Practice view —as portrayed in Figure 5.6.E, enables close monitoring of the most important indicators on learners' activity at an individual and group level (**R2**). It displays the list of learners in the class with a set of associated metrics, namely learners' Fullname, Score, Time spent, Connections, Initial level score, Initial evaluation score, Reached module; arranged in a tabular form, as described in Table 5.3. The column headers are displayed either as a histogram or a bar-chart according to whether the column is categorical or nominal. Similarly, the cells are encoded using bars to support comparison between learners. Interaction with the headers through filtering and sorting is provided to support searching and to make scanning information easier. The table is augmented with a line-chart of learners' learning progression pathways in a timeline of all learning modules in which learners may enroll (**R4**).

Evaluation view —as Figure 5.6.D portrays, displays the list of learners' metrics related to their evaluations, i.e., quizzes programmed by teachers (**R2**), in the same tabular form as the Practice view. This includes the learner's name, the time spent on the evaluation, the score obtained, as well as a corresponding derived score on a scale of, see Figure 5.6.d1–5.

Table 5.3: Practice View's indicators.

Attribute	Description
Learner	Learner's full name (Figure 5.6.e1).
Score	Learner's current score, that is, the number of rules known + acquired over the total number of rules in his/her learning path, as both a percentage and a total number (Figure 5.6.e2).
Time Spent	Time spent practicing on the platform (Figure 5.6.e3).
Connections	Total number of connections on the platform (Figure 5.6.e4).
Last connection	Last connection, as: since x hour/day/month(s)(Figure 5.6.e5)
Initial level	Grammar rules that are initially known by a learner, as both a percentage and a total number (Figure 5.6.e6).
Initial Evaluation	Score at initial evaluation as a percentage of correct answers (Figure 5.6.e7).
Reached Module	Last module reached by learner in his/her learning path (Figure 5.6.e8).

Learner view —as Figure 5.6.F) portrays, is shown as a dialog box that presents the main metrics of the learners in a grid form, in a similar way to the Main view (**R2, R3**). This includes the current score, the time spent, the total number of initially known rules, the learning progression line-chart, as well as grammar rules acquisition related metrics (top five rules in focus (currently in learning by learners), acquired (already learned) and known (already known).

Implementation: The Progdash user interface is built using (Microsoft, n.d.) and Angular (Google, n.d.). The visualizations are implemented using D3js (Bostock et al., 2011). Progdash works in a client/server fashion over rest web services. Data are pulled from the database and aggregated to compile the metrics of the different views, before being served back to the browser using JSON format. For optimization purposes, data are cached on the client for 1h before refreshing.

5.7 Field Study Design

We deployed Progdash for three months between March and June 2019. We aimed at investigating how teachers used the dashboard to monitor learners' progression. Specifically, we sought to answer two main research questions:

- Does Progdash provide teachers with useful information about learning progression?
- How does Progdash inform teachers' practices in assessing learners' online learning and conducting formative interventions?

5.7.1 Procedure

Deployment lasted three months for each teacher, split into 3 phases: opening, deployment, and closing.

Participants: We recruited 29 teachers via the company's newsletter (gender: [F=18, M=11], age: [min=30, max=60], school: [Middle=7, High=5, Vocational=7, University=10]). They all used PROJET-VOLTAIRE with their learners. We informed teachers of the logging implemented on the dashboard and their ability to withdraw at any time. They gave us permission for collecting and analyzing data.

Opening: We started the study after we had recruited 11 participants. We continued to enroll participants for two weeks afterwards. In the first week, we set up a webinar as teachers resided in different locations. We explained the main objectives of the study as well as the different parts of PROGDASH to help them become familiar with it. The webinar was recorded and later shared, along with a pdf user guide, with all the participants when they joined our study.

Deployment: We informed the teachers that they could use PROGDASH for a three-month period using dedicated web instance. Teachers used their own devices to access the dashboard from their location.

Closing: At the end, we asked the participants to fill out a questionnaire to collect feedback about the usefulness and effectiveness of PROGDASH, as well as teachers' practices. Later, a member of our research team set up online semi-structured interviews with six teachers. The interviews lasted 30 to 50 minutes. The questionnaire and interviews focused on (1) teachers' overall experience and perception in using PROGDASH and (2) how teachers used the tool in practice.

5.7.2 Data Collection and Analysis

Logs: We captured teachers' actions on PROGDASH with the associated timestamps.

Questionnaires: During the opening, we asked the participants to fill out a profile questionnaire to collect demographics information. During the deployment phase, each week, we asked teachers to fill out a simple diary questionnaire, asking if they had used PROGDASH the week before, when

and what for. We gathered 24 (out of 29) participants' responses to the profile questionnaire, a total of 135 diary entries where teachers reported whether or not they had checked Progdash during the week [many times = 48 entries, once = 33, no checks = 54], and 22 (out of 29) responses to the ending questionnaire.

Interviews: We audio-recorded the interviews and used Braun and Clarke (2006) thematic analysis process:

1. ***Familiarizing with the data.*** We manually transcribed and pseudonymized the interviews. We then allocated notes to interesting excerpts.

2. ***Conducting initial coding.*** We transferred excerpts to a spreadsheet and assigned them codes. We kept enough context to ensure proper understanding. We duplicated excerpts with multiple codes. We iterated on codes to converge towards coherent sets.

3. ***Identifying relevant themes.*** We iteratively piled and organized codes into themes, resulting in 20 initial themes.

4. ***Reviewing relevant themes.*** We created a theme map from the initial themes, which we discussed among co-authors as suggested by (McDonald et al., 2019). We decided to split some themes and to merge others, resulting in 6 main themes. We kept representative excerpts for each theme to ensure coherence.

5. ***Naming relevant themes.*** We named and defined each theme by its set of code excerpts. We then reviewed coded excerpts to refine the reporting on each theme.

5.8 Field Study Results

5.8.1 Teachers' Overall Use

Participants' interaction with Progdash evolved over the course of the deployment. As Figure 5.7 portrays, log data shows less activity at the end than at the beginning of the experiment, with school holidays in April accounting for a lesser use for two weeks.

Connections. Log data showed that, three participants —P11, P17, P27, dropped out as they never used PROGDASH. The remaining 26 participants used PROGDASH for 472 times in total [min=1, max=72, mean=18.2, sd=19.6] (identified as accessing the dashboard and selecting a class from the menu to explore *AND* counting only single login sessions). Two participants asked to withdraw —P1, P29, after using it 17 and 46 times, responsibility. P1 explained that the dashboard did not suit her/his needs of dealing with 25 groups of learners, which s/he preferred to handle directly using spreadsheet exports. P29 explained that the dashboard did not suit her/his needs. Four participants used it once —P12, P18, P24, P4. Twenty-one teachers used PROGDASH three or more times.

Table 5.4: Use of PROGDASH's views, as count (#).

View	Count	Min	Max	Avg	SD	Vocational	High	Middle	University
Main view	512	1	73	19.7	20.7	100	81	119	212
Practice view	556	1	82	23.2	25.1	112	77	120	247
Learner view	360	1	137	18.1	32.9	37	33	137	153
Evaluation view	267	1	57	10.5	14	55	63	32	117

PROGDASH Views. Teachers used mostly the Practice and Main views, followed by the Learner view, than the Evaluation view (Table 5.4). Teachers in university and middle schools used the *Learner view* more than in high or vocational schools, mostly due to extensive use of it by some teachers, e.g., P9, P10.

Overall, teachers in the university used PROGDASH more than other schools, mostly because they manage large groups of learners and also often they manage the PROJET-VOLTAIRE for the whole university. Therefore, we decided to focus on the middle, high, and vocational schools in order to gather realistic practices of teachers who use PROJET-VOLTAIRE in blended settings, meaning articulating face-to-face and online learning activities. We present in following sections, the results of 17 participants from middle, high, and vocational schools, without counting two dropouts —P11, P27. Fourteen out of the 17 participants filled out the post-questionnaire and made a total of 76 diary entries where they reported whether or not they had checked PROGDASH during the week [many times = 24 entries, once = 15, no checks = 36]. Figure 5.8 shows the interest in PROGDASH's views and metrics as reported in the post-questionnaire by the 17 participants from middle, high, and vocational schools.

Teachers' Overall Experience. We asked teachers in the post-questionnaire: *"Would you like to continue using the dashboard?"*. All 14 out of the 17 (middle, high, vocational) who filled out the questionnaire responded *"Yes"*. Similarly, we asked participants *"Would you recommend*

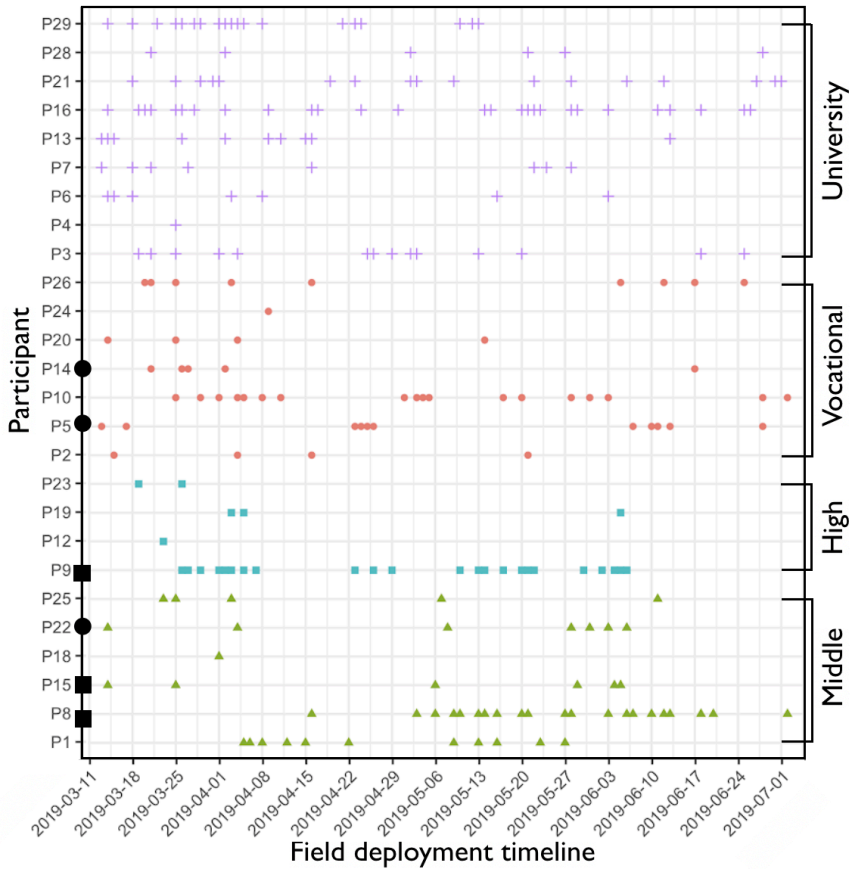


Figure 5.7: Teachers' sessions using Progdash. Interviewed teachers are marked with squares for French language teachers, and with dots for other teachers.

the dashboard to a colleague already using Projet-Voltaire?", and their average rating (out of 14 responses) was above four on 5-point Likert scale. Overall, teachers were satisfied. They highlighted in their diaries different features of Progdash, such as the (1) visual aspect of the dashboard: *"I like the graphics"* —P5, *"visual aspect of the class"* —P20, (2) learners' progression: *"progression line-charts are very interesting"* —P22, *"quick way to visualize learners' trajectories or their connection to the platform"* —P12, and (3) grammar acquisition: *"ability to see easily the levels learners reached and the number of concepts acquired by each learner"* —P1, *"many more details and vision about learners who do nothing to target them with exercises to do"* —P15.

Teachers' Pain-points and Suggestions. In their diaries and interviews participants noted some usability issues and reported several suggestions. Teachers wanted to identify learners who missed an evaluation session. They had to look for blank cells in the table to find those learners, which became tedious with multiple evaluations per learner.

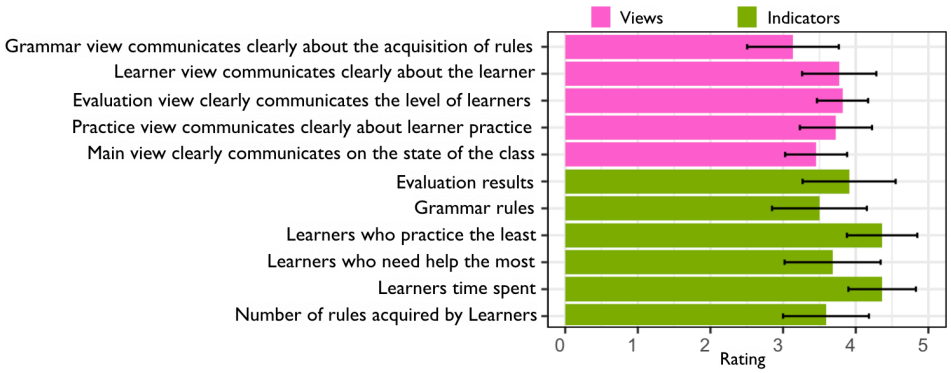


Figure 5.8: Teachers' interest in PROGDASH's views and indicators on a scale of 0 to 5 with the confidence interval.

P22 stated that teachers must often cope with technical problems in the classroom, turning on the computer, connecting, setting up PROGDASH, or any other tool becoming too time-consuming. S/he suggested simplifying it by making two modes: *"a simple mode and an expert mode. The simple mode needs to be really ultra simple and ultra direct to track learners' results, and when you want to move on to something more advanced you can."* —P22. S/he added by pointing out the need for an automated weekly report that could be sent by email to teachers as it might be more practical to access and to use on his/her mobile phone.

P14 highlighted his/her frustration with the design of the grammar acquisition view as s/he needed to be able to assign asynchronously specific concepts so that learners could practice those concepts when they launched the learning platform.

P8 and P9 disagreed with the sampling choice of indicators in the main view. For instance, P9 stated: *"top 5 is not enough because it needed to be balanced to 35 [number of learners in the class], it must be a top 10"*. Similarly, regarding the top 5 samplings of grammar concepts, P8 stated: *"I really like to be able to choose what I give to learners using PROJET-VOLTAIRE live in the classroom"*. Drawing a parallel with learners' progression line-charts, P8 and P22 requested a timeline for learners' time spent on the platform.

5.8.2 Teachers' Pedagogical Practices

Table 5.5 provides an overview of the pedagogical practices teachers developed with PROGDASH, and considered helpful.

Table 5.5: Summary of teachers' pedagogical practices through the use of Progdash.

Teachers' practices along with illustrative quotes
Monitoring learners' progression
I really rely on the line-charts, I find them very telling [...] and immediately I see who dropped out or not, after I look at the numbers, but I start by looking at the line-charts. —P8
Often when I conduct my monitoring review, once a month, actually, I will dig deeper [...] I will rather be in the exploratory but regarding a particular learner. —P14
Formative, Summative Assessments
Now, when we have the detailed list of all the [grammar] concepts, we know precisely that if the learner no longer has this concept which is displayed it means that it is acquired [...] . Before s/he [a learner] used Projet-Voltaire yes! but what value? does s/he learn no! we had to evaluate and conduct an evaluation. —P15
When we fill out learners' transcripts in the school, there is a criterion: the ability to ex-press oneself in writing with clarity and accuracy; of course, here we have objective elements [meaning using Progdash] to check the boxe. —P9
Planning, Adjusting Interventions
We had planned to work on the homophones "en" and "on" and when I saw that it was acquired by everybody, we did not do it. —P15
I use [Progdash] to adapt things when I realize that a language [grammar] element is not mastered at all, I will change my plan to incorporate it. —P8
Sharing Progdash In-class With Learners
They [Learners] were pleading to have such information, to see how they progress. —P8
I showed [Progdash] to learners [...] telling them that I could see everything that happened. And, when they knew it, it changed [learners' pace], because they didnt think we had access to it. —P15
Providing Feedback to Learners
It is rather to motivate them [learners], but it is also to sanction those who do not do their job. —P8
I didnt do a lot of close or individual monitoring with them [learners] except reminding them that they had to use the app. —P14

Monitoring Learners' Progression. According to log data, 13 (out of 17) teachers checked the dashboard three or more times to monitor learners' activity on the learning platform (see Figure 5). Overall, in their diary reporting, those teachers reported a total of 35 checks of Progdash before a class [min=1, max=8, mean=3.18, sd=2.08]. They reported doing so to check learners' status and to look for outliers, either to encourage those practicing most or to talk to those not practicing enough. Teachers used such a strategy to provide the whole class with feedback . In post-interviews, P8 stated: *"I start by looking at the line-charts, and I see who is doing very well to be able to congratulate and encourage them and who is not doing too well. It allows me to talk to them the next day when I see them"*.

Eight teachers reported in their diaries that they checked Progdash during a class a total of 27 times [min=1, max=8, mean=3.37, sd=2.26]. In post-interviews, two French teachers —P9, P15 emphasized checking the dashboard during a class, while learners were working on the learning platform, to check discrepancies between learners' progression in grammar concept acquisition. P9 stated: *"during the class when they are in training, at a given*

time, I tell everyone to stop where they are. I now have the goal to harmonize the progression of the class, so I point to the gaps between learners".

Additionally, 13 teachers reported in their diaries checking PROGDASH after a class for a total of 26 times [min=1, max=5, mean=2.00, sd=1.29]. Teachers did so to inform follow-up debriefs with learners, or just to check if learners were progressing as expected. In post-interviews, P15 stated looking at the dashboard to check whether a learner had practiced a list of grammar concepts using the platform: *"I check the next week, if s/he still has the same list [of grammar concepts], then we will talk"*. Teachers looked at different dashboard indicators depending on the type of interventions they anticipated. Learners' time spent on the platform was one of the most important indicators for all teachers (see Figure 6). French teachers who teach material related to grammar concepts or who even use the learning platform in class, were interested in monitoring a combination of indicators about learners' progression, namely, the time spent on the platform, the regularity of learning, pacing, and skills development —P8, P9, P15.

Furthermore, teachers relied primarily on the line-chart timeline to conduct overall monitoring first before drilling down for more details. In post-interviews, four (out of six) teachers —P8, P5, P14, P22, highlighted the effectiveness of the line-charts in comparing learners' progression and identifying extreme learners: *"the line-charts representation is interesting because in the blink of an eye we can have the progression of learners compared to each other"* —P14. Three teachers —P8, P5, P14, stated that they started their monitoring from the line-charts and then looked at the numbers: *"I really rely on the line-charts, I find them very telling [...] and immediately I see who dropped out or not, after I look at the numbers, but I start by looking at the line-charts"* —P8. Finally, two teachers —P14, P22, stated in post- interviews that they conduct exploratory analyses using the dashboard. P22 stated that s/he did so to explore out of curiosity: *"it was mostly to see and do some monitoring, not remediation. It was mostly to see [PROGDASH] out of curiosity"*. P14 stated that s/he often drilled down regarding a specific learner: *"often when I conduct my monitoring review, once a month, actually, I will dig deeper [...] I will rather be in the exploratory but regarding a particular learner"*.

Formative, Summative Assessments. The dashboard was instrumental to teachers as it provided ongoing feedback about learners' online activity. This, in part, supported teachers in conducting quick formative assessments about learners' progression. In post-interviews, P8 stated comparing without and with the dashboard: *"when I land on the dashboard, I see the [grammar] concepts that are the most challenging for learners [...] before I had to infer it, as learners were the ones who informed me [during discussions]. Now I have it directly on the dashboard: s/he [a learner] doesnt need to explain, and I see which concepts s/he is struggling with"*. Similarly, P15 stated: *"[before*

Progdash] we did not have the impression that we have a value [from using Projet-Voltaire]. Now, when we have the detailed list of all the [grammar] concepts, we know precisely that if the learner no longer has this concept which is displayed it means that it is acquired [...]. Before s/he [a learner] used Projet-Voltaire yes! but what value? does s/he learn [grammar concepts] no! we had to evaluate and conduct an evaluation".

Additionally, the dashboard supported teachers in conducting summative assessments of learners' progression at the end of the school year at the end of some instructional unit. P9 described how the dashboard provided objective facts for filling out evaluation reports: *"when we fill out learners' transcripts in the school, there is a criterion: the ability to express oneself in writing with clarity and accuracy; of course, here we have objective elements [meaning using Progdash] to check the boxes"*. Similarly, P8 stated using the dashboard to prepare a class council meeting: *"at the time of preparing the class council, it allowed me to see [learners' progression] in the blink of an eye. It is a great tool to see whether learners' work is steady over time, where s/he managed to progress, where s/he didnt"*. On a class level, P15 reported using Progdash to evaluate grammar fundamentals among classes: *"all the bottom part [of the Main view] about [grammar] concepts [...], especially the concepts that are already known [by learners], enabled us to see which classes had good fundamentals, and that was interesting"*.

Planning, Adjusting Interventions. Teachers used the dashboard to inform their in-class activities. Among the 13 teachers who checked Progdash periodically (three or more times), seven reported in their diaries that they checked the dashboard to prepare a debrief. They also used it to adjust their teaching plans, assign work, or structure their classroom activities (e.g., peer-tutoring).

In post-interviews, all the French teachers stated that they used Progdash to adjust their lectures. P14 stated: *"we had planned to work on the homophones "en" and "on" and when I saw that it was acquired by everybody, we did not do it"*. Similarly, P8 stated: *"I use [Progdash] to adapt things when I realize that a language [grammar] element is not mastered at all, I will change my plan to incorporate it"*. Only two teachers —P15, P25, reported in diaries that they used Progdash to prepare lesson content.

Teachers also used Progdash to target groups of learners based on their difficulties. In fact, among teachers who checked Progdash frequently (three or more times), eight teachers stated that they look at the dashboard to check the status of individual (or groups of) learners. For instance, P14 stated: *"what we found interesting is having real information to know that this learner has a problem with this rule, to be able to reuse them in more individualized exercises in personalized aid"*. P9 stated *"I will note on our digital working space*

in bold characters, to tell them: I warn you, I checked the dashboard. And, I will mark with a 56 font-size: work on PROJET-VOLTAIRE."

In the post-interview, only P15 explained using PROGDASH to set up peer-tutoring sessions: *"I take active learners [...] and I pick one learner with difficulties and ask Paul [active one] to work with Pierre [struggling], to show and explain"*.

Sharing PROGDASH in-class. The six teachers who participated in post-interviews, stated that they used PROGDASH to facilitate in-class debriefs, by showing it to individuals or to the class. Two teachers —P8, P9, stated that they shared the dashboard in-class to tell learners about their indicators, and to enable them to see their progression, as learners did not have access to such information. P8 and P9 further reported that learners *"were pleading to have such information, to see how they progress"* —P8, and they *"ask to see their line-charts"* —P9.

In post-interviews, five teachers —P8, P9, P15, P14, P5, explained that they aimed at improving learners' pacing by showing that they had indicators on their practice on the learning platform. P15 stated: *"I showed [PROGDASH] to learners [...] telling them that I could see everything that happened. And, when they knew it, it changed [learners' pace], because they didnt think we had access to it"*.

Additionally, in post-interviews, all teachers reported that they shared PROGDASH in-class to facilitate motivational debriefs with learners. Teachers described different strategies to ensure that learners maintain a steady practice. One was simply to allow learners to see their names among the top 5, or bottom 5 learners, which, in part, helped stimulate learners' self-reflection and motivate them to practice more. Similarly, P9 stated that showing the table in the Practice view ordered by learners' time spent online, *"triggered some sort of challenge, and they [learners] worked harder"*. Moreover, in post-interviews, all teachers stated that they showed the progression of line-charts to motivate learners. Its main role was in *"demonstrating the correlation"* —P9, P22, between practicing on a regular basis and learners' progression. Hence, showing progression over time was important to emphasize the longitudinal nature of the learning process. However, one teacher decided not to display the information to the whole class, and preferred to provide one-to-one debriefs with learners, as s/he thought that showing PROGDASH to the whole class would stigmatize learners: *"[PROGDASH] is something I keep for myself. I have to show it individually [...] it stigmatizes learners, and right away they would mock each other."* —P5.

We did not anticipate that teachers would share the dashboard in-class. We discuss the ethical implications of such a practice in the discussion section.

Providing Feedback to Learners. We identified different feedback strategies building on Progdash, both at individual and class level. All the teachers stated in the post-interviews that they reminded learners of the need to practice, often lecturing them briefly about the importance of practicing on a regular basis to maintain steady progress. Teachers either checked learners' status using Progdash before or during a class to congratulate learners who were practicing and provide public or private reminders to others who were not. Moreover, French teachers in particular, conducted individual debriefs with learners. For instance, P9 stated: *"I have split-classes, so I can do individual debriefs. Learners come at the end of the class and ask to see their line-charts"*. Besides, feedback also took the form of guidance regarding how to best use the learning platform. P5 stated: *"I tell them, take your time [to read and understand grammar concepts], we give you the concepts and examples so you can train on short sentences, and after you can try again"*.

In the post-interviews, teachers who did not teach French —P14, P5, P22, reported that it was the learners responsibility to practice. They mostly provided feedback regarding the efforts made by learners online on the platform, e.g., time spent, and the number of connections. Also, they often reminded learners of the importance of mastering the language for their future, or as stated by P14: *"I explain to them their [future] job, I train them to be assistant managers, so: you will have to write letters, emails [...] they will have a lot of written production"*. Non-French teachers used the learning platform in addition to their classes, e.g., marketing and management. They stated having little time to provide formative feedback, and also that they did not feel competent in providing formative feedback to learners regarding grammar concepts, as they are not specialized in French grammar. For instance, P14 stated: *"I do not feel qualified to explain a [grammar] concept, or other things, because I do not necessarily master it myself"*.

5.9 Discussion and Design Implications

In this section, we discuss our findings in relation to previous work. We also suggest design implications for dashboards in online learning settings.

Providing Insights into Learning Progression. Our findings show that Progdash was instrumental to teachers in accomplishing different tasks. First, in terms of monitoring, the dashboard surfaced information previously invisible, and confirmed teachers' intuitions about learners' online progression that they built from formal or informal discussions with learners in the classroom. This corroborates prior results, particularly in improving teacher *"visibility"* of (Groba et al., 2014) and *"knowledge"* about (Xhakaj et al., 2017) learners' online learning.

Additionally, the dashboard supported teachers in conducting formative and summative assessments of learners' activity both during and at the end of instructional units, which suggests that dashboards may help reduce assessment time as found in (Groba et al., 2014; Govaerts et al., 2011). PROGDASH particularly supported French teachers in identifying those grammar concepts where learners struggle most, and in adjusting learning materials accordingly as found in (Xhakaj et al., 2017; Molenaar and Campen, 2017).

Understanding Individual and Group Progression. PROGDASH highlighted the large variability of learners' progression in the classroom, enabled by indicators that categorize learners (e.g., top learners) and by temporal visualizations (e.g., line-charts), which highlight patterns of tendencies among learners. Also, this corroborates prior results on dashboards as enablers to identify trends, for instance in forum discussions (Mazza and Dimitrova, 2007), learners' notebooks (Yuan et al., 2008), and learners' regulation in a social network (Groba et al., 2014).

Additionally, teachers emphasized the effectiveness of line-charts in understanding the temporal progression of learners, which supported the identification of outliers, as in (Mazza and Dimitrova, 2007; Molenaar and Campen, 2017). Importantly, line-charts enabled teachers to grasp and understand the unique progression of each learner. Teachers started their exploration of the dashboard by looking at line-charts and then at numbers, as reported in post-interviews. Line-charts also supported the comparison between learners' progression —P8, P5, P14, P22.

Supporting Effective Data-informed Practices. Our study underlines critical implications of use of teachers' dashboards in contexts like ours, namely, specialized online learning platforms. That is, even if all teachers in this study used PROJET-VOLTAIRE to help learners improve their grammar skills with associated grades, there was a divide in PROGDASH use. Teachers who used the platform in their French language courses, relied on different indicators of learners' progression, including grammar concept acquisition, and conducted different informed pedagogical interventions: debriefing, feedback, peer-tutoring, and lessons. On the other hand, teachers who used the platform while teaching other disciplines, relied mainly on indicators about learners' efforts, such as the amount of time invested online and the number of connections. They were reluctant to provide formative feedback to learners, e.g., related to grammar concepts, and thus mainly lectured and reminded learners to use the learning platform on their own.

As a result, for instance, the type of feedback given by teachers could have a *mixed-effect* on learners. That is, in post-interviews, all teachers emphasized that they "congratulate" top learners, while all non-French teachers relied

mainly on "timespent" to devise their feedback. For example, praising learners for task performance is shown to be ineffective as *"it contains such little learning-related information"* (Hattie and Timperley, 2007). In part, this suggests the need to train teachers to better support the intended use of such learning platforms through dashboards. Also, in part, it suggests the need for dashboards to facilitate finding teachable elements so that a teacher can provide effective and actionable feedback about learners' progression. In fact, teachers' feedback is more likely to work when it is supportive, timely, and, importantly, specific to the learning activity (Shute, 2008).

Anticipating, Facilitating in-class Use. Progdash supported teachers in engaging in informed debriefs, e.g., *"positive reinforcement"*, *"verbal warning"* to push learners to do their best, as was found in (Aslan et al., 2019). Importantly, teachers shared the dashboard with the whole class to support their debriefs for three main reasons: *transparency*, to share information with learners as they do not have access to such dashboards —P9, P8; *control*, to let learners know that teachers have precise indicators about their progression and pacing —P8, P9, P15, P14, P5; and for *motivation*, to foster self-reflection and stimulate challenge by showing learners' indicators about their progression. This is very important as the dashboard fostered the dialog between teachers and learners in both directions, as in some classes, learners asked teachers to debrief them regarding charts on the dashboard. Such dialog improved the learning experience in the classroom, which would be lost if, for instance, the dashboard was designed only for learners (Roberto Martinez-Maldonado et al., 2015), or if the feedback was automated, e.g., programmed to learners, a practice found to be ineffective (Hattie and Timperley, 2007). However, we did not anticipate that teachers would show the dashboard in-class for debriefing with learners, or show information on all learners to the whole class, and Progdash was not designed for this purpose. This raises design and ethical concerns, especially regarding showing full names, rankings and learning data of all learners.

Additionally, in post-interviews, teachers emphasized their use of the dashboard in-class to foster learners' motivation and self-reflection. This is very important as most teachers' dashboards, even those designed to be used synchronously in classrooms, are not designed to be shared with learners to support debriefs or feedback, even though dashboards can have a positive effect on learners, as was found in (Aslan et al., 2019; Holstein et al., 2018). Besides, although we did not intend to support self-reflection by design through Progdash, this is in-line with the impact model of learning dashboards proposed by Verbert et al. (2014). These findings suggest the need for dashboards to support in-class sharing with individuals and groups, or for focused teaching interventions. However, dashboards need to be carefully designed to anticipate this kind of emerging teachers' practices.

5.10 Limitations

Although we aimed at including in the study teachers from different disciplines, and levels, e.g., middle, high, and vocational schools, to gather realistic insights into teachers' practices in using web-based online learning platforms, teachers were self-selected to participate in this study. Therefore, the results may reflect practices of motivated profiles of teachers. Additionally, we asked teachers to fill out a diary by sending them a weekly email, which may have pushed teachers to check the dashboard. Finally, our study is exploratory in nature, as the learning platform we collaborated with has a diverse *context of use* in schools, thus making it difficult to derive representative pedagogical practices of teachers in their use of the dashboard. The next step would be, for instance, to conduct a study centered around each profile of teachers with respect to their context of use of the learning platform. Further work is needed to understand: how our results may generalize to other online learning platforms, what is the impact of the dashboard on learners, and investigate the design of dashboards that connect teachers, learners, and others.

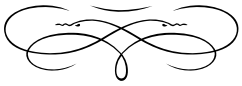
5.11 Summary

We presented a design study of a dashboard to support teachers in monitoring learners' progression on an online grammar and spelling learning platform. We iteratively designed the dashboard with teachers, and we integrated it on a widely used platform. The result of this work is PROGDASH.

We presented the results of a three-month field study of a dashboard in-the-wild, with 29 teachers, from different schools. Our results show that teachers used the dashboard at different times (before, during and after a course) and for different purposes: monitoring, formative, summative assessments, and planning interventions. Teachers shared the dashboard with learners for transparency, control, and motivational purposes, to support debriefs, provide feedback and in-class interventions, at both class and individual levels. Our results highlight a divide between the practices of teachers teaching French language and those using the platform for extra-curricular learning. Based on our findings, we suggest directions on how teachers' dashboards could better bridge online and in-class learning.

Chapter 6

Social, Pedagogical, and Technical Underpinnings of Teachers' Dashboards



> *Today, we have been predominantly model-focused, and the current models we are producing really failed to account for the complexity of the education system. How are the different agents in our system interacting? That's a key question that we need to start to address.*
>
> —Dawson, LAK'20, Keynote

6.1 Overview

The design and use of teachers' dashboards draw social, pedagogical, and technical ramifications. Such dimensions are both crucial and challenging. On the one hand, to truly elicit reflections on social factors, we need a critical inquiry into the pedagogical and technical underpinnings from dashboards in use. On the other hand, to properly design such technologies, we need in-depth understanding and consideration of the social factors of users' interactions in the context. We build upon our empirical results based on two longitudinal design studies of EMODASH and PROGDASH, both tailored towards teachers' pedagogical practices. We conduct a reflexive inquiry to examine the social, pedagogical, and technical dimensions interacting in the design and use of teachers' dashboards. We engage with these three dimensions, to examine the opportunities and challenges that we may face in the development of teachers' dashboards. Together, these dimensions serve (1) as a conceptual umbrella for the design space of classroom technologies, and (2) as macro-implications for the design of dashboards best fitting teachers' and learners' everyday situations.

6.2 Motivation

Education research has shown significant interest in dashboards for a wide range of applications. For instance, to make sense of learning data (Verbert et al., 2014), support monitoring (Holstein et al., 2017), moderate learning (Roberto Martinez-Maldonado et al., 2015), facilitate coaching (Aslan et al., 2019), initiate feedback (Roberto Martinez-Maldonado et al., 2015), foster awareness (Holstein et al., 2017). Many contributions in this direction demonstrate an interesting, challenging, and interdisciplinary area of research (Schwendimann et al., 2017; Rodriguez-Triana et al., 2016; Sedrakyan et al., 2019; Bodily et al., 2018).

However, research into education dashboards remains disjoint. We are confronted with a lack of conceptual implications spanning the design space and accounting for the interactions between various users and stakeholders (Shum et al., 2019). We lack studies examining the social, pedagogical, and technical ramifications of design and use of teachers' dashboards.

In this chapter, we build upon our two longitudinal design studies of EMODASH and PROGDASH in-the-wild. We designed each dashboard to serve teachers' pedagogical needs on online language learning platforms. We examined each dashboard in use over a long period. Therefore, in this chapter, we take a reflexive stance on our empirical results to articulate factors to inform the design and use of dashboards in specific and information-driven technologies in general for teachers' use. Several cru-

cial decisions are made and need to be made when designing teachers' dashboards. It can be difficult to separate decisions from the technical, pedagogical, and social contexts of the classrooms, teachers, learners, and the learning platforms. From a social perspective, we engage with the three concepts: monitoring, awareness, and reflection. Monitoring is often structured around the informational model of collecting data, describing aspects of learners' activities. We examine other assumptions of the concept, such as socialization, rule, and reward-based monitoring. While awareness is crucial for leveraging the educational and social interactions of teachers and learners, there is little guidance in education literature suggesting how to design for awareness. Reflection is a key concept in education, but we know little about the concept in the context of dashboards in education settings. From a pedagogical perspective, we engage with planning, feedback, and coaching. Planning refers to all the activities when teachers engage with learners' data before and after the school or classroom or in their home. Feedback takes different forms and refers to all the talks, debriefs, and discussions that teachers engage with learners. Coaching refers to all practices when teachers engage with learners to demonstrate parts and aspects of the learning. Together, they serve teachers most when bridging online and face-to-face learning. Last but not least, from a technical perspective, we engage with steps, challenges, and risks that we have learned from our experience in prototyping dashboards for teachers. We engage with the social, pedagogical, and technical dimensions as concepts and constructs, and drive implications of each dimension to provide a holistic view of the design and use of teachers' dashboards.

6.3 Social Perspectives of Teachers' Dashboards

Dashboards carry forth several social characteristics and contingencies. In the spirit of what we have discussed in chapter 4 and chapter 5, a dashboard conveys experiences, evokes emotions, elicits responses, and catalyzes reflections. A dashboard embodies an individual's and group's inter-personal aspects, beliefs, feelings, values, behaviors. It is critical to consider the social dimension when informing the design and use of teachers' dashboards.

6.3.1 Monitoring

Monitoring is one of the main design goals of dashboards in online education settings (Schwendimann et al., 2017; Bodily et al., 2018). Robert (2016) defined monitoring as the "act of observing the actions of others to determine whether they comply with a predefined set of behaviors". Rodriguez-Triana et al. (2016) described monitoring "as tracking learners' activities and outcomes". These definitions articulate an interesting facet

of the concept of monitoring, namely the informational-based tracking and evaluation of the learners' activities or outcomes of their activities. However, little is known about the implications of the concept of monitoring in education as a socio-interaction practice.

First, monitoring has a **representational** aspect. It deals essentially with what and how information about learners can be represented and reviewed. We can broadly categorize monitoring information into two categories: learners' behaviors and the outcomes of their behaviors. They are both complementary and activity-dependent. If the learning activity is easy to control, then the monitoring of users' behaviors can be easily and readily assessed. Otherwise, if the activity is less easy to control, then the monitoring of learners will be only achievable using their outcomes. In the setting of EMODASH, we were able to monitor learners' facial expressions to infer learners' emotions. In the setting of PROGDASH, we were able to monitor learners' outcomes, such as grammar concept acquisition, learners' progression, as well as some of their behaviors, such as their use and time spent on the learning platform. Representational monitoring may serve three roles: diagnostic, performance, and compliance (Richards, 1988).

Diagnostic monitoring focuses on ways to improve learning. This includes pinpointing strengths and weaknesses of teaching and learning, along with their associated causes. The grammar view of PROGDASH, for instance, enabled French teachers to evaluate grammar fundamentals among classes and identify concepts that learners were struggling with or had not yet mastered.

Performance monitoring focuses rather on ways to stimulate learners to become more effective. This includes promoting comparative information to stimulate challenges. We did not capitalize on such design goals in our dashboards. However, some teachers underlined the need for stimulating challenges and competitions among learners and appropriated some views of PROGDASH to achieve such goals. One teacher, for instance, speaking on projecting the dashboard to learners stated: "Telling them [learners] grammar is important, it's a positive social marker, it will also make you credible or not in your professional life! They are not at this level. Learners are at a much more basic level: *challenge between each other*. For me this is a very important point." —emphasis added, P9, PROGDASH.

Compliance monitoring focuses on ways to ensure meeting and maintaining enthusiasm and commitment, as well as predefined levels of quality and standards. One way in which the teachers we engaged with, maintain their standards is through distributing **rewards and sanctions** (Sanacore, 2008). In reward-based monitoring, teachers build up a (reciprocity) relationship between themselves and their learners. That is, teachers distribute rewards emanating from learners' outcomes. One of our teachers speaking

on how s/he addresses learners: "[I say] if you don't *practice an hour and a half* you will have a bad *mark*" —emphasis added, P5, Progdash. Learners also moderate their efforts, trying to maximize their outcomes with respect to the forthcoming rewards: "We have learners in front of us, we must not be utopian, if there is no *marks* for learners, if there is not a *carrot* at the end, it [learners' commitments] does not last over time, it doesn't last over time." —emphasis added, P22, Progdash.

Alongside rewards and sanctions, teachers maintain learners' commitments to learning activities through **rules** (V. G. Johnson, 1994). Predefined rules are established by teachers, and monitored with respect to learners' behaviors and/or outcomes: "[...] in contrast, we force them [giving learners a seatwork] in a computer classroom saying, here it is, if you don't practice an hour and a half you will have a bad mark; well I don't like to do it, it is 'threatening', but, threatening is enough sometime." —emphasis added, P5, Progdash. Teachers weight the dashboard indicators, with respect to their rules for distributing rewards or losses accordingly: "[...] the way I do it, it means, *imposing a mandatory 20 minutes* per week on all learners, and the one who doesn't do it, s/he doesn't do it, but s/he will have a bad mark, it will lower his/her grade, the learners know it." —emphasis added, P22, Progdash. Teachers, from time to time, conduct "control" and "surveillance" to ensure that learners coalign with the predefined objectives: "When I check [dashboard] every week, the goal is to verify that the learners have signed in, to verify that they actually did the work, it's just the *surveillance* [...] I don't necessarily watch the results every week" —emphasis added, P8. Rules are fixed by teachers as standards against which they make comparisons and judgments as to whether a learner is complying or not; such as using the online materials for a certain amount of time per week, e.g., "an hour and a half", "20 minutes", in the quotes above.

Second, monitoring has a **socio-interactional** aspect (Eisenhardt, 1985). It deals essentially with teachers' socialization with learners to build up a common ground of teachers' objectives so that learners identify themselves with or internalize these objectives. In this way, teachers work on fostering a high level of engagement from learners during moments of feedback, debriefs, talks, discussions, and coaching. During such formal and informal moments, teachers try to moderate learners to behave in ways that co-align with the teachers' objectives, by fostering values and beliefs that might prove appropriate for the task at hand, as well as for the future of learners, or as highlighted by one of our participants: "I harp on them, I explain to them the job, I train them to become managers, assistants: [...] 'you [addressing learners] will have to write emails, you will have lots of written productions, there are lots of examples of learners who graduated and they couldn't find a job [...] they made too many mistakes [...]' I'm rather more into lecturing" —P14, Progdash.

Socialization is, however, often conducted alongside the representational aspect of monitoring. In other words, instead of conveying values and beliefs, teachers use learners' behavior or outcome facts to support their social interactions in ways to convince learners by showing them, for instance, indicators about their learning, cases of excellent learners, success stories, and so on: "When I redistribute the tests and the marks to them [learners], I correlate, I show them the correlation between each person's personal investment in using platform and the progression, and in fact, it is almost mathematical, i.e., , it is either the amount of work [...] by far or the [learning] regularity" —P9, PROGDASH. The goal is, in part, to foster transparency, but also in part, to foster trust among learners so that they can identify —if not internalize, with the objectives of the teacher: "I told them [learners], we can see all those who practice seriously have progressed well, it was my intent to show them visually those who did the 20 minutes per week, showing the correlation between really serious practice and progression curves [...]; mathematically, the effort of practice and the progression curve is correlated" —P22, PROGDASH.

Rules are recalled from time to time during classrooms activities, during formal and informal discussions, and, more particularly, during moments of socialization with the goal of reinforcing the attitudes and values that will contribute to and gear towards the success (or failure) criteria. In other words, teachers try to remind everyone in the classroom of what must be done: "When I check [the dashboard] after a class, my goal is to control [check on] a class that I don't necessarily have during the day, to see where they [learners] are. Because, I actually know when I'm going to see them in the class, I'm going to issue the wake-up call. Or, I will write on the digital workspace [...] in large font, in character 56, telling them [learners]: 'I warn you, I checked the dashboard! Practice on PROJET-VOLTAIRE"' —P9, PROGDASH.

Third, monitoring has a **commitment** aspect (Ouchi, 1979; Fikes, 1982). Commitment seems to be both teacher- and learner-dependent. Monitoring requires a commitment from teachers to successfully moderate learners' learning and stimulate their engagement. At the same time, teachers aim for a high level of commitment from learners to make the most of online learning materials. From this standpoint, different levels of commitment are possible, among which internalization —is the highest, followed by identification, with compliance —as the lowest (Kelman, 1958). Internalization occurs when one commits to an activity, practice, conduct, and so on, because of a full belief in its "content" and its intrinsic reward. Identification occurs when one frames commitment to an activity due to a reciprocal response without a full belief in the relevance of its content. Compliance occurs when one commits to some practice to gain a reward, and at the same time avoid loss and sanction.

The study of Progdash highlights that teachers appropriated the dashboard differently. Some appropriated the dashboard for both diagnostic and performance monitoring. They engaged with different informed pedagogical interventions, such as debriefing, feedback, peer-tutoring, and lessons. They were eager to help make progress and maintain a learning environment conducive to learners' engagement. Others appropriated the dashboard for compliance monitoring. They fixed rules, e.g., practicing for a certain amount of time per week, and distributed rewards and sanctions. However, this might not foster excellence, as it is conducted withinsufficient focus on the learning activity. Also, it might not foster a high level of learners' commitments.

The interaction between teachers' and learners' levels of commitment and monitoring, is, however, much more complex. Teachers tend to engage with and invoke different forms of practices, e.g., imposing, motivating, coaching, to moderate learners' behaviors and outcomes. Learners also adjust their efforts knowing they are monitored, and some appreciate being so. However, monitoring seems to be a more complex construct where the different levels of commitment, e.g., internalization, identification, compliance, might all occur together at the same time. In this respect, P8, reflecting on the relationship s/he was able to build up with his/her learners, stated: "They [learners] like to see that I am watching —in quotes, that I am attentive to that. They like it a lot too" —P8, Progdash.

Summary Notes. Classroom monitoring is about emphasizing learners' behaviors and the outcomes of their behaviors, namely, learning process, progress, and skills to enable teachers and learners to see "what is going on" and gain insights to take informed actions or inform moments of socialization.

The practice of monitoring has three main aspects. First, irrespective of the form and level of monitoring teachers deal with, monitoring is a deliberated practice that requires a commitment by allocating expectations to learners and frequently assessing their associated status. Second, the representational aspect of monitoring requires capturing, assessing, and encoding learners' behaviors and outcomes to enable teachers to improve learners' learning, stimulate learners to be more effective, and ensure that learners behave in concerted ways according to an agreed-upon strategies. Third, the socio-interaction aspect of monitoring requires harnessing morals, attitudes, values, and beliefs by making explicit, communicable, and appropriate —with respect to the classroom objectives, actions that learners can readily discern, recognize, and appreciate.

6.3.2 Awareness

Education research emphasizes the role of social awareness in promoting teachers' "well-being, motivational orientation/efficacy, and mindfulness" (Jennings, Snowberg, et al., 2011, p. 38) and in maintaining a "healthy classroom climate" auspicious for learners' involvement. However, there is little guidance in education literature suggesting what awareness is about in the context of teachers' information-driven technologies and how to design for such a concept.

> *Example —knowledge*: Usually, I focus on progress: Did s/he understand? Did s/he learn from that point? Which is not connected with emotion, but EMODASH *made me aware of the other parts of teaching* which is, is s/he pleased? Is s/he happy to learn [...]? S/he was a very easy learner, s/he showed me his/her emotion. And also, I could see it on EMODASH, it reflects that feeling about him/her [...] of course nobody is happy all the time, I may be confused if I don't understand something —emphasis added, P1, EMODASH.

> *Example —feedback*: As a teacher, I always want to have some *kind of feedback* to show that Im doing a good job and kind of clarifies that Im not doing too bad. —emphasis added, P4, EMODASH.

> *Example —context*: In my school, I send a phone message to the parents, it's *the parents who receive the message* saying your son/-daughter did 10 minutes on PROJET-VOLTAIRE, s/he must finish the 20 minutes, every week the parents receive the message. —emphasis added, P22, PROGDASH.

Teachers' dashboards shape both teachers' and learners' interactions. Situations like those illustrated by the four examples above arise in teachers' dashboards. They highlight all different situations and contexts for teachers' information needs and uses. Clearly, the design and use of dashboards entail various factors for context, information, feedback, and knowledge sharing. These factors are central for classroom interactions at both individual and group level.

Such illustrative situations underline at least two main aspects. First, we need to seek beyond transmission of information in teachers' dashboard design (Markopoulos and Mackay, 2009, see chapter 5). Second, we need to consider carefully the social interactions taking place in classrooms and, beyond this, to leverage classroom technology for seamless coordina-

tion (Harper et al., 1989; Heath and Luff, 1992), communication (Dourish and Bellotti, 1992), and socialization (Heath and Luff, 1991; Galegher and Kraut, 1990) between the different interactant parties.

Both aspects above contribute to classroom awareness through information-driven technology. The former addresses designing the "right" time-space information, while the latter addresses placing information within naturalistic human needs, practices, and situations. In this sense, we leverage information for teachers to overcome challenges within particular situations and contexts of their classroom and school and to manage the emotional, social, and instructional contingencies of their living classrooms Jennings and Greenberg, 2009. We motivate our assumptions of classroom awareness from different perspectives, including, education, cognition, and sociology. The educational perspective relates teachers' social awareness to their mindfulness (Jennings and Greenberg, 2009) and "habits of mind" (Costa and Kallick, 2008). Such habits include teachers' "tendencies to gather data through all of the senses, to be aware of and reflect on experience in a nonjudgmental manner, to be flexible when problem solving, to regulate emotion and be resilient after setbacks, and to attend to others with empathy and compassion" (Roeser et al., 2012, p. 167). In turn, social awareness fosters teachers' capacities to build sustainable relationships and environments conducive to learners' engagement and learning.

The cognitive perspective relates awareness to self as well as other concepts, including regulation, metacognition, knowledge, inter-personal, attention (Morin, 2009). In this view, awareness is seen as a function of the human brain —consciousness, or a tool for consciously adjusting and reshaping the self and perhaps others, controlling inner behaviors in response to a perpetually changing world around us (Lou et al., 2017).

The sociological perspective understands awareness in the context of social interactions. Namely, as a dual or a reciprocal challenge —sometimes a problem, of knowing the latent traits of the "identity" of others, as well as the perception of the identity of self in others' eyes (Glaser and Strauss, 1964; Glaser and Strauss, 1965).

Building upon these perspectives, we can articulate at least three implications for the design and use of awareness in dashboard settings. The first implication is connecting teachers and learners through data. We, in part, inform 'connecting' from the educational perspective to support teachers' and learners' intuitions, intentions, and attentions towards each other. The second implication is rendering socio-organizational interactions through data. We inform 'rendering' from the cognitive perspective to foster users' "feelings" of presence and alertness. Finally, the third im-

plication is accounting for reforming in teachers'/learners' practices and conducts. We inform 'reforming' from the sociological perspective to support the rational/emotional negotiation of interactant parties with respect to information sharing. Different forms of awareness stand out from these views. Representational and socio-interactional forms are broad forms of awareness (Schmidt, 2002). The former focus on informant cues that we can make available to users, while the latter focus on leveraging users' social interactions. We focus here more on the latter for two main reasons. One is that we articulate classroom awareness upon the concept of monitoring (Rodriguez-Triana et al., 2016). Second, in Jennings et al.'s prosocial classroom model, "social awareness" is the impetus of teacher-learner interactions to "build strong and supportive relationships through mutual understanding and cooperation" and "effectively negotiate solutions to conflict situations" (Jennings and Greenberg, 2009, p. 495). As "most technologies and prototypes focus on reporting and recording" information in comparative work research (Bødker and Christiansen, 2006, p. 4), we prefer to use the terms of connecting, rendering, and reforming to underline the social aspects in our discussion of classroom awareness.

Connecting teachers and learners through data, highlights the management of information flow at individual and group level of the classroom and beyond. When using the term "connecting", we seek to emphasize that the information flow in the classroom awareness settings is a means to an end. It is inherently not a one-way channel. In this sense, we record and report background information about interactant parties involved in classroom conduct so they can make reciprocal sense of one another's conduct and situation, and provide the appropriate context for their past, current, and future learning activities.

In the design of PROGDASH, we engaged with teachers to understand their needs and explore the design space of the dashboard. However, we lacked the dynamics of teachers' work practices and classroom conduct. The one-way channel of information flow was a salient design factor in our dashboard. Namely, our understanding was that teachers would check the dashboard and transmit information, through discussions, debriefs, and so on. We even added an option to print the indicators presented on the dashboard. Clearly, the solution we came up with was less sensitive to reinforcing social interactions by "connecting" teachers and learners with respect to mutual conduct. To our surprise, the results of the post-interviews after three months' field deployment of PROGDASH show that all the teachers shared (projected) the dashboard to their learners, either to the whole classroom or individually. The results show that teachers did so for at least three different reasons. First, teachers want to be "transparent", so they shared the dashboard with their learners as they did not have access to the information that teachers have. Second, teachers want to "moderate" learners' behaviors and outcomes, so they shared the dashboard with

learners, thus letting them know that the dashboard facilitates knowledge of how learners use the learning platform (presence), what activities they perform (conduct), and so on. Third, teachers want to mutually "motivate" learners, so they shared the dashboard in the classroom to stimulate challenges among learners by showing them their mutual progression and conduct.

Rendering the socio-organizational interactions highlights the representation of information. By rendering, we empathize displaying teachers' and learners' online activities so as to convey the naturalistic organizations of their reciprocal practices. The rendering aspect of information-driven tools is not solely a descriptive report of what's going on. It is rather an emphasis that taps into the socio-organizational interactions of the conduct of the classroom members and beyond, fellow teachers, parents, and so on.

Dashboard information is, to a great extent, not used as-is. In the study of Progdash, teachers used, shared, and interpreted information through the lens of different informal cues publicly available in their immediate environment, such as social and physical interactions with learners, informal discussions in/outside classrooms, moments of recreation and socialization, class council meetings, meetings with parents, and so on. Teachers' sharing of the dashboard with learners is clearly a tacit example of "rendering": making private —invisible, activities of learners visible to all learners to shape the conduct of the classroom members. By doing so, teachers enabled learners to view information about one another's conduct, use such information to provide a shared context —a feeling of presence, and convey the behaviors that serve the appropriate conduct for the ongoing learning activity. Namely, teachers seek to build learners' engagement and involvement. Also, teachers triggered learners to grasp, interpret, and externalize their opinions, which in turn provided more context for appropriate interpretations of the information recorded and reported via the dashboard. Such appropriation of Progdash by teachers underlines the importance of rendering through information-driven tools as a means to drive and catalyze one another's conduct not only via "oversee" but also via "socialization".

Reforming teachers' and learners' interactions highlights two aspects. One consists in capturing the ephemerality of teachers' and learners' practices, while the other consists in capturing how such ephemeral (temporary) practices inform, shape, or even (re)create better practices and conducts. In fact, although online learning may seem individual and private, most teachers' and learners' activities are mutually informed, structured, and organized. The conduct of a learner or a group of learners affects the conduct of fellow learners as well as the conduct of the teacher. Such practices and conducts tend to be ephemeral, i.e., they fade out as time passes. Accounting for

the ephemerality of such practices in a dashboard continuously provides useful, enriching, and up-to-date information. Teachers and learners will continuously find a need for and pay attention to the rendered information.

The results of the study of PROGDASH highlight how capturing information, which is less publicly visible, such as a group of learners struggling with specific grammar concepts, means that the teacher has to tackle such concepts in a lesson, a debrief, and so on. The more the reforming is not accounted for in the setting of a particular classroom, the more awareness information will likely have no impact on the conduct of teachers and learners. The results from the study of PROGDASH show that the dashboard impacted French language teachers the most, far more than other teachers who were using the platform as a prerequirement aid where learners are encouraged to use the learning platform on their own. For French language teachers, awareness information directly reformed their day-to-day practices. To foster classroom awareness, it would appear necessary to capture how teachers' and learners' practices continuously reform their current and future interactions.

Summary Notes. Classroom awareness is about emphasizing the emotional, social, and instructional situations within the classroom and beyond, i.e., what teachers and learners are doing, if the materials are engaging, if they are struggling, if they are showing appropriate conducts, if the socio-organizational interactions are effective.

Awareness builds upon monitoring. The main objective of the monitoring is capturing relevant information for the conduct of teachers and learners, encoding, and reporting information so that it is readily available. Awareness aims at placing information in the dynamic of social and physical contexts and contingencies of the classroom, with respect to teachers' and learners' interactions. Designing for connecting teachers and learners through data allows for reflecting the mutual flow of interactant parties. Designing for rendering the socio-organizational interactions allows for capturing the naturalistic negotiations, responses, and tradeoffs of interactant parties with respect to a shared "knowing". Designing for reforming teachers' and learners' practices and conducts allows for structuring their ongoing learning situations, which in turn catalyze their alertness towards the contingencies flourishing in a complex environment where intertwined learning activities, co-located, physically distant, and so on, are all intermixed. Last but not least, taken together, designing for social awareness allows for teachers and learners to, more likely, perpetually identify with the dashboard and continue to pay attention to the recorded and rendered information in the long run.

6.3.3 Reflection

Experiences of learning and teaching, by their very nature, engage some form of reflection. Teachers reflect on their day-to-day practices; some may reflect more deeply, seeking for rational principles and engaging with practical solutions for their puzzling issues, while others may reflect less deeply engaging with "quick fix" solutions, perhaps constrained to some extent (Korthagen and Vasalos, 2005, p. 48). Our aim here is to examine ways to leverage teachers' and learners' reflections through information-driven technologies and, in particular, dashboards.

> *Example —self-reflection*: It [dashboard] would certainly make me reflect on my teaching if I have a lot of negativity [negative emotion]. So, positive then it makes me happy, because as a teacher, I always want to have some kind of feedback to show that Im doing a good job and kind of clarifies that Im not doing too bad. —P4, Emodash

> *Example —self-evaluation*: [I check on dashboard] whether s/he [learner] is attempting questions too quickly, almost mechanically. But how to help him/her to change the way s/he is doing things like that [...] maybe I could better coach them, if I did the work in class, if I was with them, and if it was not remote, but it is my teaching practice that I question. —P8, Progdash.

Reflexive situations like the ones illustrated by the examples above arise in teachers' dashboards. The design and use of dashboards denote factors to catalyze, delve into, and reassess our perceptions, beliefs, feelings, actions, and knowledge. These factors are central for both teachers' and learners' reflective thinking and for professional development.

Reflection is a key concept in education. For instance, the work of Dewey, Schön, and Kolb aims at informing professional development in education through reflective thinking. Dewey, in his work on reflective thinking, defines the concept of reflection as a "function" by which we "transform a situation in which there is experienced *obscurity, doubt, conflict, disturbance of some sort*, into a situation that is *clear, coherent, settled, harmonious*" (Dewey, 1933, p. 100:101, emphasis added). He advocates reflection as a "logical" practice of knowledge inquiry in education, suggesting that teachers perpetually seek the right balance between proceeding with immediate actions of transmitting information —through "showing and telling", and practicing "stop and think", which is, in Dewey's eyes, one of the main "practical problems of the teacher" hampering the stimulation of reflection (Dewey, 1933, p. 270).

Building upon Dewey's stance on reflection as a method of inquiry, Schön, in his work on the epistemology of practice, introduces the *reflective practicum*, or in Schön's words: —"reflection on the reflection-in-action of practice", as a method for leveraging teachers' and learners' learning by doing via "on-the-spot experiment", each working as a "reflective practitioner", namely someone "attentive to patterns of phenomena, skilled at describing what he *observes*, inclined to put forward bold and sometimes radically simplified *models* of experience, and ingenious in devising *tests* of them compatible with the constraints of an action setting" (p. 322, added emphasize, Schön, 1987).

Building on Schön's stance of learning by doing, Kolb's experiential learning model underlines a cyclic nature of reflection in learning as "the process whereby knowledge is created through the transformation of experience" (Kolb, 2014, p. 49). In Kolb's view, concrete learning experiences are followed by reflective observations, and then by conceptualizations, leading to active experimentations, which result in new experiences and learning.

Building upon these stances, we can reconcile at least three implications for the design and use of reflection in dashboard settings. One consists in highlighting situations of contingency as auspicious situations to trigger teachers' and learners' reflective thinking most. The second implication consists in facilitating the exercise of stepping back from day-to-day practices and shifting gears to critical thinking to see conditions for what they really are. Finally, the third implication consists in stimulating the effort and curiosity of seeking to make sense out of the emotional, social, and instructional experiences of teaching and learning.

Situations of **contingency** in dashboards trigger reflective thinking. We reflect mostly on part of what we know, what we are aware of, or in Schön's words, on the "knowing-in-action". We reflect when our thoughts are stimulated by informant cues. Whether or not reflection is intended by design, the information recorded and rendered through dashboards offers opportunities to stimulate reflection.

Perplexed, troubled, or confused situations are the most auspicious for catalyzing reflective thinking (Dewey, 1933, p. 106). In the post-interviews of EMODASH, all the teachers reported that high positive or negative emotional reactions from learners triggered teachers' curiosity to inquire into those moments. A teacher elaborated how a high emotional pick triggered his/her curiosity: "There was one part that was *a surprise, then I thought that's strange*, why did it peak on the surprise, I think I want to *understand why*, I did have a look at that [...], I didnt understand why! I was just curious about how the chart is such a big [referring to a pick]. I just wanted

to listen again, to what [the learner] was telling me, and to see, because it seems to be quite positive feeling" —emphasis added, P2, Emodash.

Similarly, one teacher, for instance, from the study of Progdash, was mainly "lecturing" learners based on their "time spent" on the learning platform. S/he elaborated saying: "I did not learn enough on the grammar concepts [...] I know that is not necessarily what will make [learners] improve their writing but for now I have based my feedback on that, and I know it's not enough and indeed your *dashboard enables many more things* [...] it is my mistake, *it might be necessary that I refine my practices too*" —P14, emphasis added, Progdash. Along the same lines, another teacher was conducting mainly remote learning with learners. S/he elaborated saying about learners who game materials: "[I check on the dashboard] whether s/he [a learner] is attempting questions too quickly, almost mechanically, *but how to help him/her to change the way s/he is doing things like that* [...] maybe *I could better coach them*, *if I did the work in class*, *if I was with them*, and *if it was not remote*, but *it is my teaching practice that I question*" —P8, emphasis added, Progdash. Emodash had very much the same effect on teachers. It raised teachers' concerns about the way they conduct their lessons. Most teachers appropriated (Dourish, 2003) the dashboards like a "job performance review" of their teaching in the eyes of learners' exhibited emotions: "I will definitely be concerned if I saw some you know a huge percentage of negativity, I do want to do better as teacher" —P5, Emodash.

Situations of contingencies, in Dervin's theory of sense-making, create gaps in our knowledge, where sense is broken, which leads to information seeking and inspection in order to construct a deep understanding (Dervin, 1998). Such situations are likely to trigger teachers most, encouraging them to embrace reflective thinking about their practices or as stated by: "it [Emodash] actually *reminded* me of the *emotional contingency*, emotional parts of the learners, because again usually I'm listening to accent pronunciation, grammar, there's learning skills, but I did not really think much about the general emotion, are they nervous, are they scared, are they worried. I just usually disregard that" —P1, emphasis added, Emodash.

Dashboards help to **step back** from day-to-day practices, or as Dewey said, to "stop and think" (Dewey, 1933, p. 272). They enable the pausing immediacy of actions and the *delving* into situations of particular interest to see how one's actions fit short/long-term commitments, effects, conducts, and so on (Merritt, 2018, p. 2).

When a teacher, for instance, sees a negative emotional reaction, s/he may wonder *why*. The dashboard will serve its purpose when it makes it possible to "go back to a particular point in time and look at that conversation again" —P1, or to a particular moment, event, piece of data, to assess, con-

struct, revise, or reclaim sense out of it. EMODASH facilitates revisiting past sessions, especially when experiences start fading as time passes. In this sense, teachers find the dashboard to be a "useful tool to *look after 3 or 4 lessons to see* how *things are moving* emotionally" —P1, emphasis added.

Stepping back from immediacy can support both retrospection and introspection thinking. Some teachers did engage with EMODASH to retrospect their interactions with learners, or as, stated: "I can *look back and see* from previous weeks and compare what s/he [learner] had said" —P4, emphasis added. Along the same lines, another teacher stated: "I guess it helps to go back and watch the session again and really find out what s/he [learner] struggles with as far as his/her emotions". Others engaged with past moments to introspect their practices: "I think it was quite useful to see the video as I was making the feedback comments, and *to see myself again* on the screen, it was quite interesting [...] to see the peaks and *to see where the peaks were*" —P2, empathize added.

Dashboards facilitate **sensemaking** by leveraging teachers' and learners' curiosity to make sense out of their mutual experiences. They support teachers' and learners' efforts to make connections between sequences of their interactions.

During the analysis of teachers' experience in using EMODASH while writing their feedback reports, we notice, for example, that some teachers stop writing and shift to check certain moments of their past lessons. They did this several times, jumping back and forth between writing and the exploration of certain moments of the recording of the lesson. One teacher, for instance, was editing the feedback report. While s/he was doing so, s/he constantly shifted his/her attention to EMODASH and fixatated on a document. EMODASH allows teachers to see all the contextual interactions of their lessons, including the documents that teachers share during the live lesson. In the post-interview, we asked the teacher why s/he fixated on such a segment of the lesson. The teacher answered: "I want to make sure s/he [learner] is reading it [document] correctly and if s/he is doing the activity correctly. *I find out* that Im not looking at him/her [learner] as much, unless I try to kind of show a word like with my hand or try to get him/her clues with my expressions. *I find out* [that] Im more focused on the document [..]. *I find out* [that], well the reason, I use more documents with him/her is because his/her level of English is extremely low [...]" —P5, emphasis added.

In the above quote, we notice that the teacher almost got "trapped" into the immediacy of running a live session. Because the learner had a very low English level, the teacher immediately "jumped to concrete action" (Kolb, 2014, p. 106) or a "quick-fix". S/he shared a document so that the learner could read through it. The teacher realized that, during almost the entire

lesson, s/he was sharing and focusing on a document and was having very few direct conversations with the learner, which ,however, is the main goal of the learning sessions.

Summary Notes. Classroom reflection is about emphasizing situations of contingencies to stimulate the exercise of stepping back from the immediacy, leading teachers and learners to seek to make sense out of their mutual experiential learning.

Reflection thus builds upon both monitoring and awareness. We cannot reflect on what we don't know. In monitoring, a key concern is to capture and report relevant information. On the other hand, awareness leverages relevant information through the lens of connecting, rendering, and reforming with respect to emotional, social, and instructional classroom situations. Reflection aims at leveraging information to foster an effort of looking over and rethinking "knowing-in-action', so that teachers and learners become acutely aware of their effects on the conduct they engage with, the help they offer, and the mutual interactions they engage with. Designing around classroom contingencies makes it possible to encode cues and clues to individual problems and situations, which, in turn, stimulate teachers' and learners' curiosity to explore their reciprocal experiential learning. Designing around stepping back from immediacy of actions makes it possible to facilitate the process, for teachers and learners alike, in order to enlarge their repertoire of images of contexts, actions, behaviors, and outcomes. Designing around making sense out of classroom practices allows teachers and learners to (re)frame their experiences, to confirm, modify, or refute facts about their conduct, to refine and link up their actions, behaviors, and outcomes, leading to reflective practitioners capable of devising informed responses to changes, discrepancies, and conflict situations in classroom settings. Last but not least, taken together, designing for reflection, makes it possible to catalyze teachers' and learners' reflective thinking by offering stimulating opportunities for stepping back from the immediacy of actions, delving into past situations to gain a deep understanding of important aspects of day-to-day classroom practices, and projecting such knowledge for effective future conduct.

6.4 Pedagogical Perspectives of Teachers' Dashboards

When we look at what online learning really means, and at what roles teachers play in the settings of our studies, we might find that the real aim of teachers is learners' involvement rather than direct teaching, and that the contribution of teachers to learning is to stimulate learners' involvement. Assessment, feedback, and coaching, —using information dashboards, are therefore some of the main interventions more likely to support teachers in bridging online and face-to-face learning.

6.4.1 Teachers' Routines in Using Dashboards

The practices of planning (Harris and Hofer, 2011; Sadler, 1989), feedback (Hattie and Timperley, 2007; Shute, 2008), and coaching (Stix and Hrbek, 2006; Schön, 1987) are important in educational research. They are, however, mostly conceptualized as practice (1) derived by an expert —the teacher, (2) aiming at improving motivation, conduct, and regulation of learners' learning. We instead discuss other assumptions that may prove useful in informing such pedagogical practices through the design and use of dashboards, and how they relate to the sociotechnical dimensions of teachers' dashboards. We build upon the results of our studies, to articulate and discuss the circumstances under which teachers may engage in interventions with learners; circumstances under which learners may prefer to seek interventions from their teachers; and circumstances under which teachers and learners may rely on some sort of information for their interventions, in particular, dashboards.

Planning learning is an important activity among the many things that teachers do (Harris and Hofer, 2011). Teachers we engaged with refer to "planning" or preparation as all the activities that take place before and after the school or classroom, as well as in teachers' homes. In the setting of online learning platforms, planning takes many forms. Such forms depend on teachers' context of use of the online materials. Teachers, for instance, use the PROJET-VOLTAIRE to level up the writing skills-set required for the course they teach. They try to push learners to use the platform on their own. They require/expect more autonomy from learners than teachers who use the platform as part of their course. Teachers, therefore, assess and plan things differently. We highlight below some of the planning routines of the teachers we engaged with.

After familiarizing themselves with the learning platform, teachers explain to learners how learning is managed by it. However, this can be difficult at times as some teachers do not master the inner working of the learning platform. This is especially the case of PROJET-VOLTAIRE, for instance, where learning is powered by machine learning. As a result, teachers have problems explaining discrepancies and differences between learners' progression and how it develops and changes over time.

Planning has an instructional aspect. It deals with the activities that teachers engage with to devise their instruction moves, such as monitoring, class intervention, and fostering learners' reflection. The teachers we engaged with have established their planning routines. For instance, to plan for a debrief, teachers, rather than checking on all learners, which would take up a considerable amount of their time and energy, check only on a sample, such as top and bottom learners, and use that sampling to convey their

message to the whole class. In the same way, teachers adjust their time and energy in planning by taking advantage of particular situations of interest. Teachers might provide feedback with reduced planning by building, for instance, on learners' mistakes when distributing exam paper corrections. Corrections of papers with too many mistakes show that learners are not practicing enough, thus catalyzing debriefs. Teachers check —often quickly, from time to time, on learners' progress, and issue emails, schedule homework, and so on. They adjust their course content by identifying and reusing the trickiest concepts.

Planning has an assessment aspect. It deals with the validation activities that teachers have to conduct based on learners' behaviors and outcomes when using the learning platform. Teachers used the dashboard for both formative and summative assessments. In the former, teachers evaluate learners' acquisitions from time to time along the course. In the latter, teachers evaluate learners at the end of instructional units. Assessments serve teachers in devising their next instruction moves. Teachers program online evaluations for their learners. They also reprogram evaluations for individuals or groups of learners when, for instance, learners miss a scheduled evaluation. They check and distribute the results of the evaluations. They validate and select learners to pass certificates. They prepare class council meetings. They grade learners and fill out their transcripts in the school.

Feedback plays an important role in teaching and learning (Hattie and Timperley, 2007). For the teachers we engaged with, feedback refers to all talks, debriefs, and discussions with learners. It takes different forms depending on teachers' context of use of the online learning materials. Moreover, it aims at moderating learners' future thinking, behaviors, and outcomes.

For our teachers, feedback to learners could be provided in the school, before, after, or in the classroom. It could be written —delivered asynchronously, or verbal —delivered synchronously. Teachers provide feedback to individuals or groups of learners or to the whole class at once. This could be formal, such as when a teacher takes 10 minutes at the beginning of the class to provide a debrief, or informal, such as when a teacher engages with a learner as they pass each other in the school.

Teachers use different strategies of feedback (Evans, 2013). One strategy is to provide directive feedback to learners, such as evaluating, correcting, or clarifying learners with respect to their behavior and outcomes from using the learning platform. Another strategy is to provide facilitative feedback to help learners frame their own feedback, such as motivating learners with respect to their future professional life. Teachers provide different types

of feedback to learners: lecturing, sanctioning, congratulating, motivating, evaluating, correcting, clarifying, and socializing. Feedback seems to be almost always given by teachers and not sought by learners themselves. Only two teachers mentioned that learners sought feedback from them regarding their progression with online materials. Overall, teachers seem to engage with feedback, but with little content learning-related information. However, this might have mixed effects on learners. For instance, one learner may perceive congratulations from a teacher as a motivation to do more or maintain a continuous effort, while another learner may perceive the same feedback as a license to take it easy (Hattie and Timperley, 2007; Nelson and Schunn, 2008). Teachers and learners need to be aware of feedback as a highly valuable tool for sharpening their learning. They need to master both the provision of as well as the search for mutual quality feedback.

Teachers' **coaching** is an important practice for improving the learning experience (Stix and Hrbek, 2006). For the teachers we engaged with, we refer to coaching as the practices when teachers engage with learners to demonstrate parts and aspects of the concepts of the learning materials. In Schön's words, teachers' coaching is a form of the reciprocal dialog with three essential features: "it takes place in the context of the student's attempts to [learn]; it makes use of actions as well as words; and it depends on reciprocal reflection-in-action" (Schön, 1987, p. 100).

Our results highlight some of the qualities of the teacher-coach. Teachers shared the dashboard with learners either as a group or individually to catalyze a dialog with their learners. Some teachers also demonstrated different aspects of learning, such as tackling concepts that are not mastered by learners or those that proved most tricky. Teachers also reported that they shared the dashboard to foster learners' self-reflection on their own or other people's behaviors, outcomes, and progress, namely, what Schön referred to as a reciprocal reflection-in-action in the coaching practice (Schön, 1987, p. 100).

Conversely, teachers, for instance, were using the PROJET-VOLTAIRE as a prerequirement aid to support their course, but not as a part of it. As these teachers were not specialized in the concepts of the learning platform, they were reluctant to intervene to demonstrate parts or aspects of the learning materials, i.e., grammar concepts, and thus mainly lectured and reminded learners to use the learning platform on their own. However, lecturing or reminding learners to do this might not be very effective in practice. The design of the dashboard could be more helpful for teachers, not specialized in the concepts of the learning platform, if it provided specific content to inform teachable moments. This could help teachers better coach learners by striking the right balance between "telling and showing".

Summary Notes. These routines derive from the different teachers we engaged with. The context of use of the online learning materials influences teachers' practices. There is a clear variability among teachers, as concerns their ways of planning, providing feedback, and coaching. Each teacher has "routinized" such practices in unique ways. We have yet to explore the impact of such routines on learners. Such variability among teachers' routines shows that the use of online platforms and the associated dashboards in school may be in an experimentation phase. Dashboards need to incorporate learning-related contents, as well as guidelines, to assist teachers in how to best make sense and use of the dashboard.

In the following section, we unpack the factors affecting teachers when using dashboards. Such factors are, in part, related to the interplay between the learning activities taking place in the classroom and those taking place online, as regards the needs of both teachers and learners.

6.4.2 Antecedents and Consequences of Teachers' Use of Dashboards

We build upon our results to discuss factors that influence teachers' use of dashboards in authentic settings. Such factors can trigger, support, or even hamper teachers' use of dashboards.

Teachers' **teaching model** affects their information needs and uses to inform their practices. Teachers have different models for how things should work and what they expect from their learners. Some teachers may be more into cognitivism, others may be more into socio-constructivism,while others may develop their own models for the classroom (Kalina and Powell, 2009). During design exploration of Progdash, we asked, for instance, one teacher if s/he is interested to know learners who get stuck in Projet-Voltaire as they game the system, attempt questions, and avoid hints. Teachers reacted that they would not have any influence over such learners. Some teachers may seek to behave in such a way as to avoid confrontations, feel they have no control over the class, or believe they have no influence over learners (Korthagen and Vasalos, 2005). Inter-personal aspects, beliefs, feelings, behaviors, are but a few factors playing out in teachers' teaching models. They have an enormous impact on their interactions with learners.

Teachers' **perceived responsibility** of learners' online learning greatly impacts their practices, that is, teachers' perception as to what extent learners are personally responsible and intrinsically engaged with the learning activity at hand. Teachers associate online learning responsibility with learners' personal characteristics, namely, with either learners' effort or ability, or both. Teachers who associate learning responsibility with learners' abil-

ity, are, to a great extent, aware of the different aspects of online learning, what learners are doing, whether the materials are engaging, what concepts are causing the most problems. They are aware of the capacity and limitation that learners experience on the online activity. They are willing to engage with learners to deploy their repertoire of strategies, such as feedback, hints, lessons, and so on, to provide learners with the help they need. One teacher from the study of PROGDASH elaborated in the post-interview on this, saying: "it [PROGDASH] is part of my work as a teacher in face-to-face sessions, because we have one [session] per week for each class, to support learning the concepts and optimizing the use of this program [PROJET-VOLTAIRE]. Therefore, that's my job as a teacher, the computer does not replace the teacher" —P9, PROGDASH. On the other hand, teachers who associate online learning responsibility with learners' effort think that learners are not trying hard enough to achieve the progress required. Such teachers may use the dashboard mainly for monitoring purposes and for deploying some of its three components discussed earlier, such as rule or reward-based monitoring.

Perceived affectivity is another factor influencing teachers' practices. The kind of relationship teachers have with their learners influences their practices as well. Teachers who have a good socio-affective relationship with their learners may be more eager to provide the assistance learners need. Learners also may feel more eager to seek assistance from their teachers. Teachers' interactions with learners showing good performance and/or conduct —whether through feedback, coaching, lessons, or some other practice, may not be the same experience as in their interactions with average or below average learners. Two teachers from the study of PROGDASH reported how learners stay at class during breaks, asking for a debrief regarding their progression on the dashboard. Teachers may be more ready to engage with learners showing appropriate conduct and with better performance, and more reluctant to do so with bottom learners. In the same way, providing feedback to top learners engenders positive feedback, which may be more pleasant to share, while lower performance engenders negative feedback. Also, it may even require more effort from teachers to frame feedback in a way that reflects performance and is still perceived as constructive feedback by the learner.

Proximity between the materials taught in face-to-face and in online learning affects teachers' practices. Teachers are not really aware of learners' online activities. We need to make learners' online activities and experiences more visible in order to elicit teachers' interventions. The more a dashboard is designed to bridge between off- and in-class learning, the more likely it is to ensure teaching and learning continuity between the online and interactive learning taking place inside the classroom. The more a dashboard is designed to compensate for such proximity, the more likely it is to elicit teachers' planning, feedback, and coaching. To appreciate this

concept, consider, for instance, P22, a math teacher who participated in the field study of Progdash. Although s/he was a math teacher, s/he was using the Projet-Voltaire as an online learning platform to level up learners' skills in spelling and grammar for day-to-day classroom writing. Many teachers do the same with this learning platform, actually as a "quick-solution" for the writing problem of their learners. This teacher was the one using the dashboard the least to inform class interventions. The post-interview shows how the teacher was passionate about his/her subject, math, and absolutely not all grammar. S/he elaborated that the main feature s/he liked most on the dashboard was the line-charts. S/he shared Progdash with his/her learners in-class once. S/he reported doing so to explain the "correlation" between steady time spent on the platform and consistent progress. S/he continued explaining how, mathematically, we can prove such a correlation. The teacher was clearly more enthusiastic about the subject s/he taught than about grammar concepts. We can imagine the result if we designed in a way to connect more closely the information recorded and rendered on the dashboard with the subject the teacher is teaching in the classroom. The dashboard would more likely catalyze the teacher's interest.

Typical or **distinctive situations** of learners' online learning impact teachers' use of dashboards. Such situations are easy to grasp because they are different from others. The more the situation is distinctive, the more likely it will be to catch teachers' attention and elicit some sort of interaction to devise either planning, feedback, or coaching. In the post-interviews of the study of , for instance, all teachers emphasized "congratulating" learners with a higher performance or "lecturing" those with a lower performance. In almost no cases teachers mentioned engaging, through the use of the dashboard, with average learners. The results of the study of Emodash highlight the same pattern. Teachers showed more interest in peaks of learners' emotional reactions than in normal moments of the live learning session. Regardless of the effectiveness of these practices, such distinctive situations, however, seem to elicit use of teachers' dashboards the most.

Reciprocal reflection impacts teachers' and learners' interactions through the use of data. The more the dashboard elicits teachers to engage with data-informed practices, e.g., planning, feedback, coaching, the more likely it is to trigger their observation and inquiry of their impact on learners. It has very much the same effect on learners: for instance, a feedback or a reciprocal dialog with the teacher will more likely trigger learners' reflection on their learning, conduct, behaviors, outcomes. Past interactions either initiated by teachers or sought by learners will more likely influence their current and future interactions. The reflective effect shapes not only teachers' and learners' practices and interactions, but also their mental model, responsibility, and affectivity. A teacher, for instance, may realize, either through facts presented on the dashboard or his/her reflective observation

and inquiry following an intervention, say, feedback, that some learners are working hard enough, but still struggling. As a result, s/he may change his/her perception of learners' responsibility from effort to ability.

6.5 Technical Perspectives of Teachers' Dashboards

A great body of research has developed theoretical models, processes, and methods to inform the practice of design in many academic disciplines, including, education (Gustafson and Branch, 1997; Flórez-Aristizábal et al., 2019; McKenney and Kali, 2017), human-centered design (Bannon, 2011; Muller and Kuhn, 1993; Donald A. Norman, 2006), information visualization (Munzner, 2009; Sedlmair et al., 2012), innovation (Brown, n.d.) to name but a few. They all aim at reducing the complexity of the enterprise of design by abstracting domain-specific problems into "abstract" aspects, such as design rationales (Carroll and Rosson, 2003), data semantics (Munzner, 2014), goals (Lam et al., 2018), tasks (Brehmer and Munzner, 2013; Amar et al., n.d.), visual encoding (Cleveland and McGill, 1984b). In this section, we contribute to this body of research by highlighting the steps, challenges, as well as the risks that we have learned from our experience in prototyping and designing dashboards for teachers.

6.5.1 Situate

Situating the domain space is the main design phase. Although our aim is to produce artifacts, e.g., a dashboard, we have to consider phenomena, such as facts, e.g., learners' progression, tasks, e.g., identify outliers, activities, e.g., class orchestration or even values, e.g., equal progress of learners. These phenomena are *situated* and *dynamic*. They develop and change over time in specific time-places, e.g., classroom, school, home, and so on. When designing, we need to understand the interplay between a teacher and a dashboard through those phenomena, and other related entities, including, learners, parents, staff members, and so forth. Situating the domain space is (1) to capture those phenomena and entities, (2) understand their impact on teachers, and (3) explicitly describe their roles as parts and considerations of the design. Domain space reflects the possible range of motivations, needs, and constraints under which teachers are able and want to do their work in real-world settings. This will lead us to articulate an explicit description or visual story of all the dimensions, phenomena, and stakeholders that impact teachers and learners when using the dashboard, which we can conceptualize in the form of claims that the design must accomplish (Carroll and Rosson, 2003).

Challenges. Designing for "real-world practice" is complex, or in Schön's words: "all but as messy, indeterminate situations" (Schön, 1987, p .4). Characterizing teachers' work practice in the field, although very challenging, is important. Role play, scenarios, and interviews are some of the methods most commonly used to *"know users"* through dialog and empathy (Wright and McCarthy, 2008). However, focusing solely on teachers' needs may not be sufficient. We need to gain a deep understanding of the interplay between teachers' different contexts and situations, and the perceptual, cognitive, pedagogical, and communicative intent of the design of the dashboard. Informed by ethnography, contextual inquiry (Beyer and Holtzblatt, 1999) may prove useful for an immersion in teachers' and learners' routines to observe how they perform their activities, participate with them, empathize, and learn about their day-to-day operations, decisions, workflows, pain-points, constraints to gain insights and inspirations. Once deployed in schools or classrooms, a dashboard inevitably inherits all the complexity of classroom conducts and teachers' work practices, constraints, issues, routines, teaching goals, teaching models, self-concepts, beliefs, images, feelings, behaviors, and more.

Risks. Dashboards by their very nature aim at capturing, summarizing, and presenting a set of readily quantified information. Teachers' reliance on and trust in the information recorded and rendered on a dashboard are important factors, while hard to qualify. This is becoming even more important as black-box AI and machine learning are now increasingly powering classroom learning technologies and their associated analytical tools. Teachers are confronted with the black-box and sophisticated nature of commercial learning platforms. They are interested in understanding the inner working of such platforms and how learners' learning is affected.

6.5.2 Ideate

The aim of this phase is to rapidly generate multiple design ideas as early solutions to the list of the claims to accomplish through the design. Showing teachers multiple and diverse concrete design alternatives is valuable for gathering feedback, validating design choices, and refining design ideas. To begin with, this can be done using rapid and cheap methods, such as paper frames, sketches, powerpoint presentations, and wireframes. Once we have refined the initial design ideas, we can start creating design approximation prototypes (B. Hartmann et al., 2006). We find parallel prototyping effective in creating different design alternatives and validating them with teachers in parallel. This helps to rapidly gather multiple points of view on different design ideas, and to discover unseen constraints and new opportunities (Dow et al., 2011). We also find it helpful to create explicit data abstraction to capture the types and attributes of data (Munzner, 2014) that

we need and can gather from the learning environment. Creating explicit data abstraction helps to consider all the constraints related to data early on in the ideation phase. At some point in this phase, we need to incorporate real data into our digital prototypes, which is valuable for discovering unseen limitations and gathering practical insights.

Challenges. Teachers have different interests in using a dashboard in order to achieve different objectives and outcomes. Homogeneity, in terms of their information needs and uses as well as their data, analytical and visual literacy, would seem to be slight. We were amazed to discover that some teachers rely on their colleagues to manage tools in order to inform their practices, while others recommended sophisticated interactions such as sort, hide, resize from tools like a spreadsheet. This poses challenges of designing tailored representations depending on teachers' needs, contexts, and situations, which are very hard to address through a fixed design. Otherwise, teachers may find the design incomplete for their work practices, their ways of doing things, and even their familiarities with other interfaces.

Risks. Dashboards rely on collecting, storing, and processing users' data. Users' privacy, and ethics are a major concern, and sensitive information needs to be manipulated properly. This requires providing enough contexts to all the stakeholders involved in the design and use of a dashboard, regarding the collection, analysis, use, and design of data in dashboards.

6.5.3 Develop

In this stage, we develop the validated prototypes to build a dashboard. This means putting together all the views required on the dashboard and connecting the dashboard with the data sources to feed the views. We find it useful to iterate in parallel from both top-down and bottom-up. In the former, we abstract design goals, rationales, and needs into interactive visual encoding, views, pages, navigation, and layout to shape the dashboard. In the latter, we connect the dashboard with underlying data sources, algorithms, queries, and transformations to expose the data needed to feed the dashboard views. This helps us to discover unseen constraints and/or new opportunities in the early stages of development. We find it interesting to add *traces* loggers to gather the *needed* evidence on teachers' use of the dashboard, such as time spent, clicks, and so on. We also find it interesting to add *errors and audits* loggers to gather evidence on bugs and issues that may occur. We also find it helpful to use modern Javascript frameworks and visualization libraries, such as, Angular, React, and Vue, D3js, Vega, to design interactive views.

Challenges. Building a dashboard poses a variety of challenges, such as the laborious process of preparing data in usable forms (Kandel et al., 2011), analyzing and building analytical models (Chatti et al., 2012) to gather insights into the state of learning, devising different visual summarization techniques to visualize information in a few views (A. Sarikaya et al., 2018) while avoiding loss of data variation through summarization (Matejka and Fitzmaurice, 2017). Moreover, dashboards are usually constrained by display size. That is, they require different approaches to lay out the information needed on the screen. We might need to apply multiple and coordinated views to combine different, while linked, indicators on the screen (Roberts, 2007) while avoiding visual clutter that may overwhelm viewers (Rosenholtz et al., 2010). We might also need to apply techniques to provide viewers with a glanceable reading and allow them to drill up and down for more detail as needed (Shneiderman, 2003). Similarly, we might apply responsive design so that view content changes based on available space on the screen (Perlin and Fox, 1993), thus allowing the dashboard to fit different viewports, e.g., projector, mobile, desktop. Additionally, dashboards, as interactive and computational tools, might require a learning curve. Teachers, on the other hand, (often) lack formal time-slots for data-informed interventions. They might avoid using the dashboard all together if they find it is not self-explanatory, easy and fun to use.

Risks. Dashboards by their very nature build upon the notion of data selection, sampling, and processing of subsets of information to visualize and inform the audience. Even when this much-desired process is properly conducted to shape correct insights, information on dashboards can, however, be misinterpreted by teachers. There are several reasons for this. One reason relates to teachers' data, visual and analytical literacy (Barbara Wasson, 2015). Another reason relates to the actual process of selecting and processing blocks of information, which can lead to loss of data variation through summarization, reduce the quality of data (Matejka and Fitzmaurice, 2017), or even engender biases by highlighting certain parts and aspects of data and shrinking, hiding, or skewing other parts and aspects. In such cases, the representations and information shown by the dashboard might lead to inappropriate interpretations, decisions, and biases.

6.5.4 Evaluate

At this stage, the aim is to take the dashboard to teachers' contexts in real settings so that they can use it to inform their pedagogical practices. Interestingly, at this stage, we loop back to the situating stage with the *aim* of understanding the interplay between teachers and the newly designed

dashboard, considering all the phenomena, e.g., activities, tasks, motivations, pain-points, and entities, e.g., learners, parents, home, to collect evidence and learn insights into teachers' data-informed practices using the dashboard. Different evaluations can be conducted to assess the impact of the dashboard on teachers and understand what practices/behaviors enable/limit the dashboard. The evaluation may compare and contrast different forms and features of the dashboard. It may be more focused, such as assessing whether teachers are able to perform specific tasks, or more exploratory, such as understanding how teachers appropriate the dashboard over the deployment.

Challenges. Ideally, we want the dashboard to be self-explanatory. However, this can be challenging due to teachers' different data, visual and analytical literacy. This becomes even more challenging as the evaluation often concerns larger-scale deployment, involving more teachers than those who participated in the three first stages. New issues and needs from teachers may emerge. Dealing with all the incoming information sources can be an overwhelming experience. It may be useful to define a protocol of how to iterate on the dashboard during the field deployment period: "what to pay attention to, what to dismiss, and how to explore, extract, recognize, and choose useful information" (Stolterman, 2008).

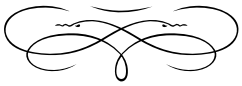
Risks. Appropriation of the dashboard may differ from one teacher to another. Defining appropriate modalities and criteria of using —and perhaps sharing, dashboard information must be defined in advance.

6.6 Summary

We built upon our empirical design studies and longitudinal field deployments of PROGDASH and EMODASH. We articulated and discussed three important dimensions interacting with the design and use of teachers' dashboards. From a social perspective, we engaged with classroom monitoring, awareness, and reflection. From a pedagogical perspective, we engaged with teachers' planning, feedback, and coaching, along with their antecedents and consequences. Finally, from a technical perspective, we engaged with the steps, challenges, and risks that we have learned along the way. We engaged with such dimensions to understand them as concepts and constructs. Together, they underline interesting, challenging, and interdisciplinary areas of research. Furthermore, they provide a conceptual umbrella under which information-driven technologies for classrooms may operate and contribute macro-implications for the design and use of teachers' dashboards.

Chapter 7

Conclusion



> *The major message in this book is that enhancing teacher quality is one of the keys and the way in which to achieve this is through ensuring that every teacher in the school has the mind frame that leads to the greatest positive effect on student learning and achievement [...] What is needed is more space for teachers to interpret the evidence about their effect on each student. This may require some major rethinking of teachers' work [...] to spend such time working together to plan and critique lessons, interpret and deliberate in light of evidence about their impact on each student's learning, in each other's classes observing student learning, and continually evaluating the evidence about how "we as teachers in this school" can optimize worthwhile outcomes for all students and share the errors, the enjoyments, the successes about the impact.*
>
> —Hattie, 2012, p. 457:459

In this dissertation, we studied the design and use of teachers' dashboards to inform pedagogical practices when assisting learners in using online learning platforms. We sought, specifically, to examine what teachers actually *do* with dashboards in their everyday work and in long the run. We argue that this research draws social, pedagogical, and technical ramifications, which are crucial to articulate. We cannot truly elicit reflections on social factors without a critical inquiry into the pedagogical practices and technical challenges that teachers' dashboards underpin. Concurrently, we cannot properly design such technologies without a profound understanding and consideration of the social impetus of users' interactions in the context of classroom situations. We collaborated with teachers, and other stakeholders, in designing teacher-centered dashboards, and conducted two of the first longitudinal studies of dashboards in-the-wild, lasting two and three months, respectively, to examine the impact of dashboards on 34 teachers in online language learning settings.

In the following, we summarize our research contributions. We also discuss the limitations of our work. We conclude with three opportunities for future research into dashboards in education systems.

7.1 Contributions

In this dissertation, we adopted a two-fold approach to study and design teacher-centered dashboards. We first engaged with teachers and designers of two learning platforms to conduct design studies and longitudinal field studies of dashboards in authentic settings in order to assist teachers in online language learning. We then articulated macro-implications of the opportunities and challenges involved in the design and use of teachers' dashboards.

In chapter 3, we presented a **case study** examining a multimodal approach, i.e., applying two theoretical models of emotion: discrete and dimensional, and using cloud APIs for emotion recognition when inferring learners' emotions in online learning. We demonstrated this approach in the context of the SPEAKPLUS platform for video-conferencing language learning, by using audio and video streams when inferring automated emotions along with learners' self-reported emotions, as well as contextual teachers' and learners' interaction traces on the platform. We analyzed 10 learning sessions of 45 minutes each —a total of more than 7 hours' recording. We compared emotional cues from the two modalities, audio and video. We proposed an approach for combining discrete and dimensional emotional cues to capture learners' emotional states.

Findings. In addition to proving the feasibility of using cloud APIs for emotion recognition in online learning environments, our findings highlight the pertinence of the multimodal approach for inferring emotion in distance learning. The results showed that different emotions are detected from different sources, i.e., audio and video, during such settings.

Contributions. The contributions of this work are:

- A case study of using cloud APIs for measuring learners' automated emotions from audio and video, alongside self-reported emotions and contextual learning traces in online and distance learning.
- A method to unify discrete and dimensional models of emotions.
- A qualitative and quantitative exploratory analysis comparing learners' inferred emotions from audio and video modalities.

In chapter 4, we built upon our technical exploration of emotions, and adopted an iterative design process in which we collaborated with a pedagogical manager of SPEAKPLUS, a linguist researcher, and two teachers while designing EMODASH. During the design, we conducted two formal formative think-aloud usability tests leading to the final design of EMODASH. We integrated the dashboard into the feedback report editor of the SPEAKPLUS. We conducted a two-month in-the-wild **field study** with five professional teachers and five learners to evaluate (1) how EMODASH supports teachers' retrospective awareness of learners' emotions in online learning and (2) how EMODASH impacts feedback reports that teachers write to learners after each learning session. We specifically compared teachers' written feedback reports under two conditions *with* and *without* EMODASH.

Findings. The results of this work showed that most teachers reacted positively to the EMODASH. We found that EMODASH led teachers to incorporate more affective elements in their reports, suggesting a stronger awareness of learners' emotions. Teachers also wrote more formative and less summative feedback. Furthermore, our results suggest that glanceable visualizations of learners' emotions may be preferred to foster teachers' awareness of learners' emotions. Finally, the dashboard led teachers to reflect on the way they conduct their lessons, using learners' emotions as a proxy to evaluate their conduct of teaching.

Contributions. The contributions of this work are:

- EMODASH, a dashboard designed based on interviews with five teachers and one learner, and refined via an iterative design process, to facili-

tate teachers' retrospective awareness of learners' emotions, and support teachers' writing of reinforcement feedback reports to learners after video-conferencing language learning sessions.

- An in-the-wild field study of EMODASH. This study extends our understanding of how emotional dashboards support retrospective awareness of learners' emotions in a video-conferencing language learning environment.
- A qualitative and quantitative study comparing teachers' written feedback reports under two conditions *with* and *without* EMODASH.

In chapter 5, to better ground the design of dashboards for teachers in distant and blended learning platforms, we conducted a **design study** based on interviews with seven teachers to understand their needs. We collaboratively iterated on design prototypes with teachers. We implemented and integrated PROGDASH into the PROJET-VOLTAIRE platform for online French grammar and spelling learning. We then conducted a three-month in-the-wild **field study** with 29 teachers to evaluate (1) whether PROGDASH provides useful indicators to teachers about learning progression, and (2) how PROGDASH informs teachers' practices in assessing data to engage in formative interventions.

Findings. Our results show that most teachers reacted positively to the dashboard. They found it actionable to inform several pedagogical practices: monitoring, conducting assessments, planning interventions, sharing in-class, providing debriefing and feedback. Our results further highlight a divide between teachers specialized in the concepts of the learning platform (i.e., French language) and those in disciplines requiring skills (e.g., Marketing), even if both were concerned with achieving the same objectives, mainly helping learners master spelling and grammar. Based on our findings, we provide design implications aimed at improving dashboards to bridge online and in-class learning and foster learners' self-reflection.

Contributions. The contributions of this work are:

- A design study of PROGDASH to support teachers in using a curriculum-integrated online learning platform.
- A three-month in-the-wild field study of PROGDASH, which extends our understanding of how teachers integrate a dashboard in their practices to articulate remote and in-class learning.
- Design implications for dashboards to bridge online and in-class learning and facilitate data-informed pedagogical practices.

In chapter 6, we took a reflexive inquiry building upon the results of our studies of EMODASH and PROGDASH. We articulated the social dimensions – monitoring, awareness, and reflection –, the pedagogical dimensions – planning, feedback, and coaching –, as well as the technical dimensions interacting with the design and use of teachers' dashboards. In each dimension we structured and discussed practical assumptions to inform the design and use of teachers' dashboards. We believe this provides a holistic view of the macro-conceptual implications of the design space of information-driven technologies for classrooms, and helps facilitate a deep consideration of and engagement with important factors underpinned by teachers' dashboards.

In short, we hope to make the following broad contributions:

From a **design perspective**, our research provides important insights into the design of teachers' dashboards. We demonstrate working exemplars laying out an end-to-end design process of teachers-facing dashboards in-the-wild, along with underlying steps, opportunities, and challenges. We illustrate such a process in two different learning environments by situating teachers' needs and contexts of use, distilling design rationales, describing data abstractions, as well as the visual encodings when shaping different views of each dashboard. Moreover, we provide fully functioning implementations and architectures of collecting, analyzing, and visualizing learning data in EMODASH and which may prove useful for more design-focused experiments.

From an **empirical perspective**, our results highlight the usefulness, effectiveness, and more importantly, the impact of dashboards on teachers' everyday practice. We demonstrate different routines surrounding teachers' practices in using dashboards when assisting learners' online learning. We demonstrate teachers' experiences, perceptions, and appropriations of dashboards in the contexts of two authentic learning environments. We provide insights into benefits of dashboards for monitoring, awareness, reflection, planning, feedback, and coaching.

From a **methodological perspective**, our research was theoretically and methodologically informed from different domains. We articulated conceptual assumptions of the social, pedagogical, and technical dimensions of teachers' dashboards. We engaged with such dimensions as concepts and constructs, and drive implications of each dimension, which in turn provide a holistic view of the design and use of teachers' dashboards. Such dimensions serve as a wider conceptual umbrella for the design space of classroom information-driven technologies, and implications for dashboards leveraging teachers' and learners' situations and contexts.

7.2 Limitations

All research work is limited. This thesis is no exception.

From a **design perspective**, in both studies of EMODASH and PROGDASH, we did not tackle the ground truth of the analytical aspects, in terms of accuracy, scalability, generalizability, and such, of the indicators that we have derived, encoded, and represented. Future studies may build upon our work to tackle the underlying technology of the design of dashboards. Furthermore, the results highlight a variability among teachers' practices. More focused design studies seem promising for education dashboards, for instance, focusing on specific practices, such as planning, feedback, and coaching, and accounting for different contexts of uses of an online learning platform by teachers.

From an **empirical perspective**, we conducted our studies of EMODASH and PROGDASH on two commercial learning platforms that specialize in online language learning. As such, we did not tackle generalization of the findings on other online learning platforms such as the open ones, e.g., Moodle, as well as on other platforms that are not language learning centric. Both of our studies were teacher-centric, although the results show the need of engaging with all the stakeholders interacting with teachers' work, and thus possibly affecting the design and use of the dashboard, including learners in the first place. The design of dashboards that connect different stakeholders, including learners, advisors, fellow teachers, parents, and such, is an area yet to be explored.

From a **methodological perspective**, we articulated the social, pedagogical, and technical factors underpinned by teachers' dashboards through a reflexive approach —"posthoc", and by building upon our work. Although such factors provide valuable insights and may prove useful in informing the design space of teachers-facing dashboards, "focused and active" research may be lacking to examine whether the dimensions we have derived would support engaging with teachers to study and design dashboards to leverage, for instance, monitoring through socialization, classroom awareness, and classroom reflection. There is a clear need for more methodologically oriented studies —ideally informed from different settings, to inform the design and use of education dashboards by providing conceptual implications spanning the design space and accounting for the interactions between the various agents.

7.3 Future Directions

This thesis underlines interdisciplinary opportunities for future research into dashboards in education systems.

From a **design perspective**, a promising area of research into learning dashboards could be to examine more focused techniques to serve specific needs and uses of learning information by teachers and learners, such as dashboards that triangulate the social, emotional, and instruction dynamics of teaching and learning; dashboards that foster teachers' and learners' monitoring, awareness, and reflection; dashboards that facilitate informed pedagogical interventions. In the same way, another promising area of research could be to examine ideas from literature, such as information visualization, information organization and communication, and such, to inform the design of learning dashboards. Such literature has developed specific techniques in terms of different design goals and roles of visual interfaces, e.g., exploration, confirmation, explanation Schulz et al., 2013; Lam et al., 2018 communication (Skau et al., 2015; Parsons, 2018), storytelling Segel and Heer, 2010; Echeverria et al., 2018, presentation (Kosara, 2016). Positioning such techniques in education research to leverage teaching and learning needs and uses seems a promising issue. More research is needed in this area, i.e., examining, comparing, and contrasting specific design goals, roles, or rationales of dashboards. More research is also needed in designing reusable analytical architectures and toolkits to collect, process, analyze, and visualize learning data.

From an **empirical perspective**, there is a clear need for *macro*-empirical studies covering all the complexity in the design and use of dashboards in authentic settings. This kind of work can lead to more focused studies on teachers' and learners' information needs and uses. Characterizations identify the dimensions that impact teachers in using dashboards in real-world settings: their needs, social interactions with learners and parents, pain-points, expectations, workflows, and challenges, and such. Data literacy identifies teachers' abilities to analyze, interpret, understand, and make use and sense of data, as well as their associated visual encoding and representations when informing their practices (Barbara Wasson, 2015). Taxonomies identify tasks, such as high-level activities, e.g., monitoring, class orchestration, conducting a debrief, providing feedback; and low-level tasks, e.g., comparing learners, challenging advanced learners, personalizing learning, and design goals, e.g., monitoring, exploration, presentation, communication, storytelling (Alper Sarikaya et al., 2018). Design rationales identify claims about aspects of the design that influence certain behaviors of teachers and/or learners (Carroll and Rosson, 2003), for instance, how certain designs, e.g., descriptive, predictive, and prescriptive designs, affect teachers in developing certain strategies, how dashboards

empower teachers' judgment in contexts, how teachers rely, over-rely, and trust dashboards in the long run.

From a **methodological perspective** there is a clear need for more research to provide guidance and principles in order to inform the enterprise of design. Such work includes thinking, analysis, and process models that inform the way we design. Thinking models, such as the impact model of learning dashboards (Verbert et al., 2014), inform *what* we expect from design. Analysis models, also referred to as theoretical contributions, inform *what* —decisions, and *why* —justifications, we do design. Process models inform *how*, through a practical approach, we design.

Our empirical research suggests a richer characterization of dashboards. A characterization that goes beyond the transmission facet of information via a dashboard. A characterization that brings people and technology together and accounts for their mutual interactions.

When we look at what a teachers' dashboard is all about, we might find that we need to leverage at least three main properties to best stimulate teachers' and learners' needs and uses of information-driven technologies, namely, *reflectional*, *educational*, *recreational*. Taken together, these three facets can help increase the perceived value of teachers' dashboards.

The **reflectional** property continues in the spirit of what we have already discussed in chapter 6. Building upon monitoring and awareness, reflectional design seems very promising to leverage teachers' dashboards as a reflective thinking tool. We may achieve this by rendering informant cues and clues on a dashboard with the ultimate goal of catalyzing some sort of a dialog between the viewers and the interface. The more such a reciprocal dialog is catalyzed, the more likely the dashboard is to lead viewers to embrace a continuous effort of retrospection and introspection. We catalyze such a dialog by enabling viewers to review the flow of their actions and outcomes, reason about the underlying effects, and grasp meanings for effective future acts.

The ultimate intent of the **educational** property is to direct the dashboard to enable teaching and learning, or in other words, to leverage the interface as a cognitive and communicative aid that teachers and learners can use in classrooms and beyond to support online learning. The dashboard can be designed as information portals organizing the information needed to support both teachers' work and learners' construction of knowledge in the setting. That is, to support devising and engaging with informed interventions. To support sharing information with learners, colleagues, parents, and others. To provide opportunities to tailor teaching and learning with respect to learners' *different* needs, styles, abilities, and disabilities.

The **recreational** property is the *fun* and appealing side of the dashboard. A dashboard is hardly ever used in vacuums, but rather in different routines often related, but not directly, to teaching and learning. Such property leverages the design of the dashboard for those situations taking place before, after, and during private moments; and during moments of informal talks, meetings, and discussions. By doing so, the dashboard provides an entertaining dimension which, in turn, improves teaching and learning experiences. As a result, the recreational side supports teachers, learners, colleagues, parents, and others, in finding a passion for information needs and uses, and consequently, stimulating the enjoyment of using and interacting with and via the dashboard.

We hope our work provides a new perspective on the design for teachers and learners, and highlights its value and research. We hope designers and researchers find our research inspiring to inform the design, analysis, and validation of dashboards to best impact teaching and learning. We hope our research helps spark new ideas.

Appendix: EMODASH

1 Online Resources

- Demo and supplementary materials website
- Training video for the field study
- Github repository for the code source
- Teachers' consent form
- Learners' consent form

2 Usability Test Script —Think-Aloud

We would like to gather feedback regarding your experience and perception of a dashboard that we have designed for tutors of SpeakPlus.

We would like you to imagine that you have conducted some learning sessions with a learner. And, we invite you to use the dashboard to explore the emotional state of your learner.

The emotions have been gathered from video-recordings of learning sessions.

Specially we would like to know what do you think as tutor about the dashboard and which part of it works or not.

While doing so, we would like you to think out loud describing your thinking, interaction with the dashboard, frustration, confusion etc.

3 Post Interview Questionnaire Script

First of all, we would like to think you so much for your participation in this interview. So, the main aim is to help us understand how EMODASH can help tutors explore, understand, interpret and reflect on learners' emotions.

We would like to ask you a couple of questions regarding your personal experience as a tutor in your training in using EMODASH in the feedback report editor in SpeakPlus platform. Specially we would like to know what do you think as tutor about the dashboard and which part of it works or not.

Please notice that the goal of our interview is to evaluate EMODASH, but not you? There is nothing that you may have done wrong? And, specially don't worry at all, if you don't have a response of any question?

And, please let us know exactly your feedback about EMODASH, don't be shy and just be honest about everything related to your experience with EMODASH, because this is why we are being here today. Your feedback is really really valuable for us and will for sure help us improving our understanding of the role of emotion data visualization in online learning environments.

I'll start with a couple of questions regarding the usability of EMODASH and the feedback report, okay!

- *What is your overall impression of EMODASH?*
- *Could you please show me the feedback report of the sixth session?*
- *Do you remember that you have used EMODASH, or somehow relied on it in editing this feedback report?*
- *If yes: could please tell me which part of it was the most useful and how?*
- *If not: did you see it? Did you maybe with other feedback report?*
- *Do you find any part difficult to use?*
- *Is there any particular feature of it you find confusing or maybe you don't understand?*
- *Now, I'll give some examples taken from your personal experience with EMODASH, I'm really wondering to get a little bit of context from you regarding these examples! Okay!*
- *Could please tell me know what was your aim when you did ?*
- *I also notice that you didn't ever... is there any specific reason that you didn't try...?*
- *In general, did you find any feature to be more interesting in EMODASH or any feature that is less interesting for you?*

- *Alight,*
- *Is the emotional information has prompted you to reflect on your peer emotions?*
- *Regardless the type of the emotion, do you think that quantifying emotions this way I mean using data visualization useful? Could please tell me how? Or maybe give me a concrete example?*
- *Could please tell me if somehow Emodash has motivated you as tutor to edit the feedback report? Could please give me a concrete example?*
- *Did you use Emodash in another situation not mainly when editing the feedback report?*
- *What could you say if we ask you to explain Emodash to someone else?*
- *For you the first level of Emodash at the top refers to what? Have you come to use it in somehow?*
- *What about the level in the middle? Have you come to use it in somehow?*
- *And the level in the bottom? Have you come to use it in somehow?*
- *Have you used any dashboard similar to Emodash? How do you compare theme to Emodash?*
- *Compared to the feeling of co-located setting, could you tell me what is earned and loss with SpeakPlus platform eventually regarding the user's experienced emotions?*
- *Do you think that information provided by Emodash responds to these points?*
- *Do you have any other comments or questions?*
- *Thank you so much again for your time and participation. Please, feel free to contact us if you have any other questions or ideas coming up later on [...]*

4 Post Usability Test Questionnaire —SUS

- *I think that I would like to use this dashboard frequently.*
- *I found the dashboard unnecessarily complex.*
- *I thought the dashboard was easy to use.*
- *I think that I would need the support of a technical person to be able to use this dashboard.*
- *I found the various functions in this dashboard were well integrated.*
- *I thought there was too much inconsistency in this dashboard.*
- *I would imagine that most people would learn to use this dashboard very quickly.*
- *I found the dashboard very cumbersome to use.*
- *I felt very confident using the dashboard.*
- *I needed to learn a lot of things before I could get going with this dashboard.*

Appendix: Progdash

1 Online Resources

- Demo website
- Training video for the field study
- Github repository for the code source
- Teachers' consent form

2 First Study: Interview Questionnaire Script

Merci d'avoir accepté de participer à cette session. Comme je l'avais expliqué par mail, Woonoz travaille en partenariat avec le laboratoire de recherche LIRIS à Lyon sur une nouvelle version d'un outil de suivi des élèves à destination des enseignants. Notre but est de travailler dans une approche orientée utilisateur, c'est pourquoi on organise ces sessions avec des utilisateurs de l'outil de suivi statistiques afin d'avoir votre retour d'expérience. Notamment, à travers cet entretien, nous souhaitons avoir un retour sur votre propre usage de l'outil, et sur vos besoins en terme d'aide au suivi de vos élèves.

Est ce que vous avez des questions avant que nous commencions ?

Questions introductives:

- *Quelles classes avez vous ?*
- *Depuis combien de temps êtes-vous en poste dans cet établissement ?*
- *Dans quelles classes utilisez vous le projet Voltaire ?*
- *Y a t'il d'autres personnes qui utilisent le Projet Voltaire dans votre établissement ?*
- *Quand est ce que vous avez utilisé le Projet Voltaire pour la dernière fois ?*

- *Combien de temps avez-vous utilisé le projet Voltaire ?*
- *Est ce que c'était en classe ou hors classe ? pour quel type d'activité ? (dirigée / ou non)*
- *Est ce que c'était ponctuel ou tout au long du trimestre/année ?*
- *Pouvez-vous indiquer quel temps vous souhaitiez que vos élèves passent sur Projet Voltaire ?*
- *Avez-vous réalisé un diagnostic de niveau ? Si oui à quel moment (avant de l'utiliser ? après un certain temps ?)*
- *Est-ce que vous évaluez vos élèves à partir du Projet Voltaire ?... Si oui, comment ? Note sur 20 ?*

Outil: Suivi Statistiques:

- *Quand avez-vous commencé à utiliser l'outil de Suivi Statistiques sur l'année scolaire ?*
- *Est ce que vous avez regardé l'outil de suivi statistiques récemment ?*
- *Souvenez-vous pourquoi ?*
- *Est ce qu'on peut le regarder ensemble ?*
- *Est ce que vous pouvez me montrer ce que vous avez regardé ?*
- *Est-ce que vous avez déjà exporté les données ?*
- *Est ce que vous utilisez l'outil de suivi statistiques dans d'autres situations que pour le suivi ? Quand ? lors d'un debrief ? ou un examen ?*
- *Est ce que vous en discutez avec vos collègues ?*
- *Plus généralement, quels types de suivi des élèves faites-vous ?*
- *Est ce que vous avez organisé des évaluations ?*
- *Avez travaillé vers la certification ?*

Généralisation:

- *Est ce qu'il y a un collègue qui gère le projet Voltaire pour l'établissement ?*
- *Est ce que vous avez discutez avec lui des usages que vous avez ou que elle/lui a ?*
- *Est ce que ça s'est passé comme ça les années précédentes, avec d'autres classes ?*

Conclusion:

- *Souhaitez-vous continuer à utiliser le Projet Voltaire ?*

- *Est ce que vous pensez l'utiliser l'année prochaine ?*
- *Est ce que il y a des collègues à vous qui seraient intéressés pour participer à un tel entretien ?*

Autres:

- *Avez-vous déjà testé / essayé les concurrents du Projet Voltaire ? Si oui, lesquels ?*
- *Qu'avez-vous préféré dans le Projet Voltaire par rapport à ces derniers ?*

3 Second Study: Interview Questionnaire Script

Bonjour, [remerciement, context, durée, et objectif de l'entretien.]

Introduction:

- *Où en êtes vous de votre année ? Avez vous terminé les cours ? Avez vous bouclés les conseils de classe / jurys ?*
- *Profil: clarifier les choses qui ne sont pas claire dans le formulaire*
- *Vous souvenez quand est ce que vous avez utilisé le tableau de bord pour la dernière fois ?*
- *Avez-vous pu faire ce que vous vouliez faire ? Si oui, comment ? Si non: Qu'est-ce que vous n'avez pas pu faire ?*

Usage:

- *Vous souvenez vous de quels événements vous ont incité à consulter le tableau de bord ?*
- *Pouvez-vous décrire comment vous avez utilisé le tableau de bord au cours de cette période?*
- *Qu'avez vous regardé dessus en particulier ?*
- *Quel type de suivi ?*
- *Lorsque vous avez utilisé le tableau de bord, l'avez-vous utilisé d'une manière exploratoire ou plutôt pour trouver ou consulter des informations spécifiques ?*
- *Quelles sont les choses que le tableau de bord vous a permis de découvrir sur vos étudiants ?*
- *Quels sont les aspects les plus utiles pour vous dans le tableau de bord ?*

- *En général, dans quelle mesure les données de progression de l'apprentissage des étudiants PV sont-elles prises en compte dans vos pratiques enseignantes ?*
- *Vous y pensez souvent ? Pour vos activités en classe ? Pour vos activités pédagogiques ? Autres ?*
- *Consultez-vous le tableau de bord pour prendre des actions pédagogiques ?*
- *Si c'est le cas, comment ?*
- *Sinon pourquoi ?*
- *Avez-vous généré/utilisé des rapports pdf depuis Progdash ?*
- *Si c'est le cas, comment ?*
- *Sinon pourquoi ?*
- *Avoir accès au tableau de bord a-t-il suscité un intérêt personnel ou une curiosité envers la progression des étudiants ?*
- *Si c'est le cas, comment ?*
- *Sinon pourquoi ?*
- *Quels sont les fonctionnalités, les informations, ou améliorations que vous aimeriez avoir dans le tableau de bord ?*

Wrap up:

- *Est ce que vous pensez qu'il y a quelque chose de spécifique à votre usage qui serait différent chez d'autre enseignants / utilisateurs de Projet Voltaire / gestionnaire de sphère ?*
- *Si le tableau de bord était personnalisable - configurable pour avoir que ce qu'il vous faut comme information - pour vos propres pratiques / utilisations cela serait-il utile ?*
- *Avez-vous eu des conversations avec vos collègues, parents, et / ou étudiants sur le tableau de bord des informations communiquées ?*
- *Autres choses ?*

4 Field Study Diary Questionnaire

- *Combien de fois avez-vous consulté le tableau de bord de suivi la semaine dernière ?*
- *Si vous ne l'avez pas utilisé, pourriez-vous nous dire pourquoi :*
- *Quand avez-vous consulté le tableau de bord ?*
- *Où avez-vous consulté le tableau de bord ?*

- *Quel était votre objectif en consultant le tableau de bord ?*
- *Si des informations vous ont manqué sur le tableau de bord. Lesquelles ?*
- *Quelles vues vous ont été le plus utile ?*
- *Pourriez-vous donner les raisons de cet avis ?*

5 Post Field Study Questionnaire

- *Sur une échelle de 1 à 5, pensez-vous que le tableau de bord proposé répond à vos besoins ?*
- *Au cours de l'année, pour quelle(s) raison(s) avez-vous consulté les informations concernant le Projet Voltaire pour vos élèves ?*
  - *Mener un suivi de ma classe*
  - *Mener un suivi individuel*
  - *Mener un suivi d'un groupe d'élèves*
  - *Préparer et/ou mener une leçon*
  - *Mener un debrief*
  - *Vérifier si les objectifs en termes de progression sont atteints*
  - *Comparer la progression de mes élèves*
- *Quelle(s) tâche(s) n'avez-vous pas pu réaliser ?*
- *Sur une échelle de 1 à 5, les informations suivantes vous intéressent-elles lorsque vous consultez le tableau de bord ?*
  - *Réfléchir sur mes pratiques enseignantes*
  - *Échanger avec mes élèves sur leurs compétences et leurs besoins*
  - *Collaborer avec d'autres collègues pour aider les élèves à progresser*
  - *Échanger avec les parents sur les performances des élèves Etre réactif pour aider mes élèves à progresser*
  - *Autre(s) ?*
- *Sur une échelle de 1 à 5, pensez-vous que le tableau de bord a influencé vos pratiques pédagogiques ?*
  - *Je pense que le tableau de bord a influencé la façon dont je motive et j'encourage mes élèves*
  - *Je pense que le tableau de bord a influencé la façon dont j'explique à mes élèves comment progresser*
  - *Je pense que le tableau de bord a influencé la façon dont je conduis (donne) mes leçons*

  - *Je pense que le tableau de bord a influencé la façon dont j'alerte mes élèves*
  - *Je pense que le tableau de bord a influencé la façon dont j'échange avec des collègues (et/ou parents) par rapport aux données d'apprentissage*
  - *Autre(s) ?*
- *Quels sont les aspects que vous avez le plus appréciés dans le tableau de bord ?*
  - *Le nombre de règles acquises des élèves*
  - *Le temps d'usage des élèves Les élèves ayant le plus besoin d'aide*
  - *Les élèves qui ne pratiquent pas suffisamment*
  - *Les règles de grammaire*
  - *Les résultats des évaluations*
  - *Autre(s) ?*
  - *La vue d'ensemble communique clairement sur l'état de la classe*
  - *La vue de détail communique clairement sur la pratique des élèves*
  - *La vue évaluations communique clairement sur le niveau des élèves*
  - *La vue élève communique clairement sur un élève*
  - *Une vue mode présentiel communique clairement sur l'acquisition des règles*
  - *Autre(s) ?*
- *Quel est selon vous le degré d'utilité des éléments suivants du tableau de bord ?*
  - *Différentes plages horaires pour consulter les données (7 jours ou 30 jours)*
  - *Différents vues pour structurer les informations du tableau de bord (vue globale, détail des élèves, détail pour un élève, vue évaluations) La*
  - *contextualisation d'informations avec des éléments graphiques (courbes, couleurs.. etc.)*
  - *L'intégration du tableau de bord avec le portail de suivi du Projet Voltaire*
  - *La génération des rapports pdf à partir des éléments du tableau de bord*
  - *La génération des exports csv des éléments du tableau de bord Autre(s) ?*
  - *Autre(s) ?*
- *Quels sont les aspects que vous avez le plus appréciés dans le tableau de bord ?*
- *Quels sont les aspects que vous avez le moins appréciés dans le tableau de bord ?*
- *Pensez-vous qu'une durée d'usage de 3 mois est suffisante pour évaluer le tableau de bord ?*
- *Avez-vous des fonctionnalités ou des informations que vous aimeriez avoir dans le tableau de bord ?*
- *Souhaitez-vous continuer à utiliser le tableau de bord ?*

- *Recommanderiez vous le tableau de bord à un collègue utilisant déjà le Projet Voltaire ?*

Appendix: Résumé

1 Motivation

Au cours des deux dernières décennies, la technologie a transformé les modes d'enseignement et d'apprentissage dispensés à des millions d'apprenants à travers le monde (Technavio, 2016) pour l'éducation formelle et non formelle. De nombreux décideurs et chercheurs préconisent la technologie pour améliorer les expériences ainsi que les résultats des systèmes éducatifs (Siemens and Long, 2011; Chasteen et al., 2011; Ferguson, 2012). Cependant, les enseignants sont les *utilisateurs finaux de première ligne* qui triment pour orchestrer les besoins et usages technologiques à l'intérieur (Vermette et al., 2019) comme à l'extérieur (Hillman et al., 2019) de la classe, ainsi que pour exploiter les données d'apprentissage pour améliorer les pratiques (Campaign, 2018).

Dans les environnements d'apprentissage en face à face, les enseignants sont, plus ou moins, experts pour saisir le climat de la classe. Ils observent fréquemment l'état de leurs apprenants. Ils évaluent l'engagement et la participation des apprenants aux activités d'apprentissage. Ils développent une conscience des contingences interactionnelles tant à l'échelle des individus que des groupes. Ils s'engagent dans des interventions éducatives et formatives. Ils (ré)adaptent, (ré)instruisent et (ré)expliquent différentes parties et aspects de leur activités pédagogiques pour atteindre avec succès les objectifs d'apprentissage (Hattie, 2012, p. 81).

L'apprentissage en ligne est complexe de par sa nature même. Les indices vitaux des interactions humaines, tels que, le contact direct, les expressions faciales, les échanges avec les apprenants, entre autres, font largement défaut. Il peut être difficile pour les enseignants de développer une forte conscience émotionnelle, sociale et pédagogique sur les apprenants en ligne (Jennings and Greenberg, 2009). Les apprenants développent aussi des comportements divers et variés en ligne : différents styles d'apprentissage, acquisitions, progressions, régularités, rythmes, etc. (Onah et al., 2014; Y. Lee and Choi, 2010). L'articulation d'activités pédagogiques en classe et à distance (Vermette et al., 2019; Hillman et al., 2019) rend encore plus difficile pour les enseignants l'acquisition un feedback précis sur les interactions des apprenants avec les activités d'apprentissage en ligne.

D'un autre côté, l'utilisation des technologies en éducation peut fournir, au quotidien, un flux de données important. Cela peut fournir des informations précieuses sur les comportements, les

processus, les progressions ainsi que les compétences des apprenants (Siemens, 2013; Siemens and Baker, 2012). L'analyse des données d'apprentissage dans le but d'améliorer des pratiques enseignantes au quotidien est à la fois complexe et laborieuse (Baker and Inventado, 2014). Les enseignants se heurtent à des obstacles pour exploiter les données d'apprentissage dans le cadre de leur de travail. Ils manquent de formations pour analyser, visualiser, comprendre et interpréter les données d'apprentissage (Barbara Wasson, 2015). Ils manquent également de plages horaires officielles pour tirer profit des données d'apprentissage afin d'informer et diriger les pratiques enseignantes (Campaign, 2018; Sandford et al., 2006). Il existe un besoin croissant d'outils pour transformer les données d'apprentissage en informations exploitables pour et par les enseignants. Les *tableaux de bord d'informations* peuvent combler cette lacune.

Les tableaux de bord sont désormais courants dans la société (Few, 2004; Alper Sarikaya et al., 2018). En effet, ils tirent parti de notre familiarité et nos expériences avec des technologies d'information et de communication que nous utilisons au quotidien : téléphones, montres, moniteurs, pour n'en nommer que quelques-uns. Ils présentent les informations d'une manière facilitant la lecture et la compréhension. Ils peuvent présenter l'information de telle sorte à l'adapter à un large éventail de publics. Pour les enseignants et les apprenants, les tableaux de bord semblent des outils prometteurs pour faciliter une variété de pratiques sur une base régulière, telle que le suivi des apprenants (Holstein et al., 2018), l'assistance des apprenants à faible capacités (Molenaar and Campen, 2017), la personnalisation de l'aide aux apprenants en temps réel (Aslan et al., 2019), et la provision du feedback (Roberto Martinez-Maldonado et al., 2015).

Cependant, des recherches récentes soulignent le manque de recherches longitudinales examinant la conception, l'utilisation ainsi que l'impact des tableaux de bord sur les pratiques enseignantes dans des contextes réels (Rodriguez-Triana et al., 2016; Schwendimann et al., 2017; Bodily et al., 2018). D'un point de vue *design*, Shum et al. (2019) soulignent pour un besoin croissant de recherche sur les tableaux de bord centrée sur les besoins des enseignants et des apprenants en conditions réelles afin de mieux servir des pratiques pédagogiques. Verbert et al. (2014) soulignent la nécessité d'explorer des sources "de données plus riches et pertinentes" dans la conception des tableaux de bord en systèmes éducatifs. D'un point de vue *empirique*, des études de plus de 150 tableaux de bord en éducation soulignent la nécessité d'études longitudinales sur le terrain pour mieux comprendre comment les enseignants s'approprient, exploitent et façonnent leurs pratiques pédagogiques autour des tableaux de bord sur le long terme (Verbert et al., 2014; Schwendimann et al., 2017; Bodily et al., 2018). D'un point de vue *méthodologique*, les travaux sur les tableaux de bord en éducation restent disjoints. Shum et al. (2019) soulignent la nécessité d'implications macro-conceptuelles couvrant l'espace de conception et tenant compte des interactions entre les utilisateurs sous-jacents. Ferguson (2012) souligne l'importance de trianguler les ramifications sociales, pédagogiques et techniques dans la conception et l'utilisation des technologies d'analyse de données d'apprentissage.

Cette thèse examine plus en détail le design et l'utilisation des tableaux de bord à destination des enseignants. Plus précisément, il fournit des résultats empiriques, à partir de deux études de design couplées avec deux évaluations longitudinales sur le terrain, sur :

1. Les pratiques des enseignants autour des tableaux de bord ainsi que les avantages pour les enseignants afin d'améliorer les expériences et les résultats d'apprentissage.

2. Les recommandations sociales, pédagogiques et techniques pour guider la conception et l'usage des technologies d'information en éducation.

2 Cadre de recherche

L'utilisation de la technologie d'apprentissage apporte à la fois des opportunités et des challenges. Ces technologies ajoutent souvent plus de fonctionnalités et tendent à ne pas considérer autant les pratiques des enseignants. Du point de vue des apprenants, ils sont souvent conçus autour de ń faites de votre mieux ż. Les tableaux de bord à destination des enseignants peuvent combler le fossé entre la technologie d'apprentissage et la pédagogie. Nous argumentons que cette recherche a des implications sociales, pédagogiques et techniques. D'une part, pour comprendre les facteurs sociaux, nous avons besoin d'une enquête critique des pratiques pédagogiques et des challenges techniques sur lesquels reposent les tableaux de bord des enseignants. D'une autre part, pour bien concevoir ces technologies, nous avons besoin d'une compréhension et considération profondes des interactions sociales des utilisateurs.

3 Approche et objectifs de recherche

Dans cette thèse, nous avons adopté une double approche pour étudier et concevoir des tableaux de bord centrés sur des pratiques enseignantes. Nous avons d'abord engagé avec des enseignants et des concepteurs de deux plateformes d'apprentissage pour mener des études de design ainsi que des études longitudinales sur le terrain pour évaluation l'impact des tableaux de bord dans des environnements authentiques d'apprentissage en ligne. Nous avons ensuite articulé les macro-implications sociales, pédagogiques, et techniques pour mettre en avant les opportunités et challenges du design et usages des tableaux de bord par des enseignants. La figure refappendix:abstract:fig:approach illustre notre approche de recherche sur le plan méthodologique, design et empirique en lien avec les objectifs de la thèse.

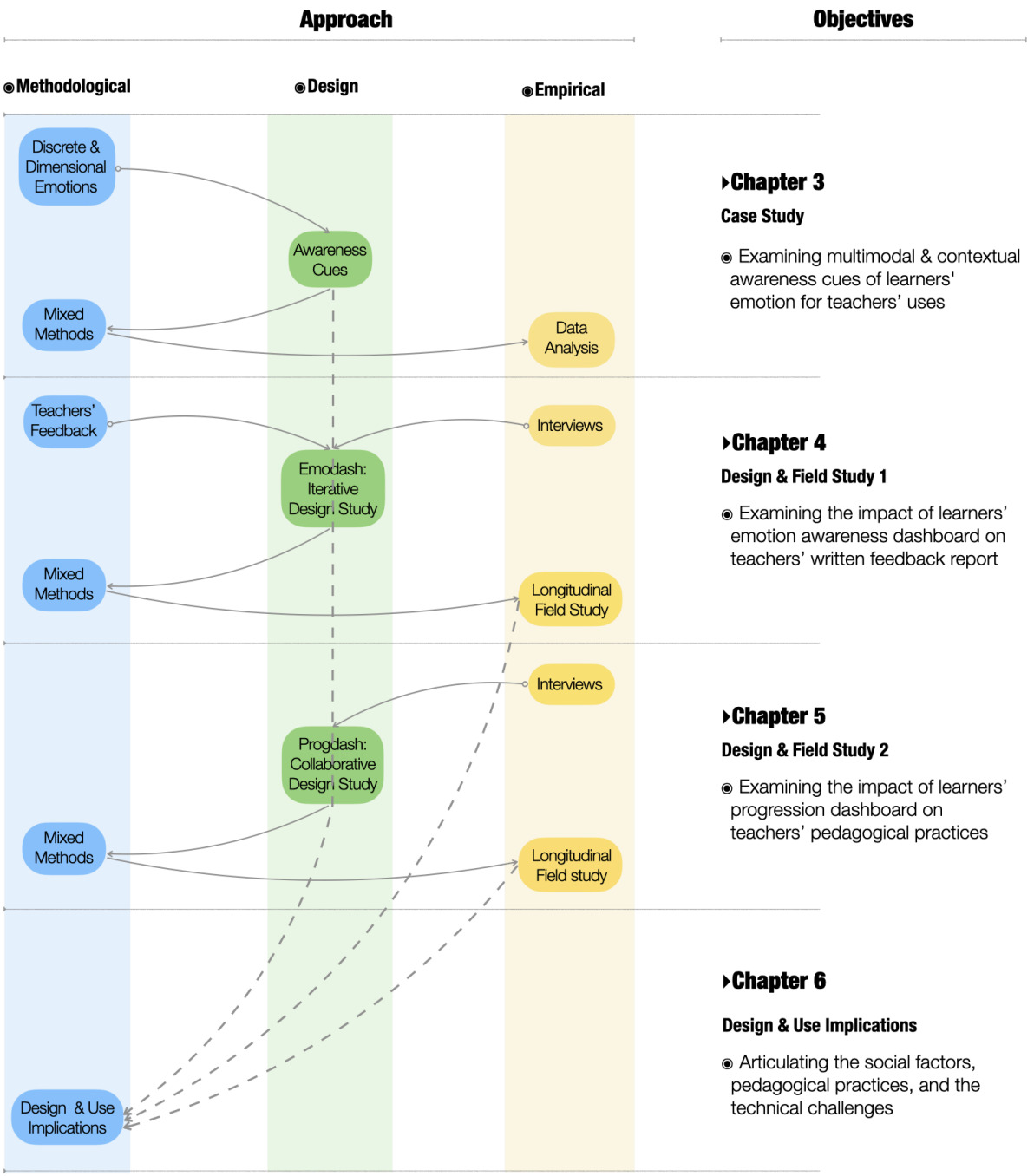


Figure 1: L'approche méthodologique, design et empirique en lien avec les objectifs de thèse.

Nous avons mené cette thèse dans le cadre de deux plateformes d'apprentissage commerciales. Les deux plateformes sont spécialisées dans l'apprentissage de langue en ligne. SPEAKPLUS est une

plateforme de vidéoconférence dédiée à l'amélioration des compétences en communication orale dans une langue étrangère, à savoir l'anglais, le français et l'espagnol. Les enseignants utilisent Speakplus pour donner des cours virtuels aux apprenants. Projet-Voltaire est une plateforme Web dédiée à l'apprentissage du vocabulaire, de la grammaire et de l'orthographe du français. Les enseignants utilisent Projet-Voltaire pour améliorer les capacités d'écriture des apprenants. Les deux plates-formes sont utilisées par les écoles publiques/privées, ainsi que par des enseignants indépendants. Avec Speakplus, nous avons examiné les tableaux de bord des enseignants dans un environnement d'apprentissage entièrement à distance. D'autre part, avec Projet-Voltaire, nous avons examiné les tableaux de bord des enseignants dans un environnement mixte où les enseignants articulent à la fois l'apprentissage en ligne à distance et en face à face. Ensemble, les deux plateformes nous ont fourni des contextes intéressants pour examiner la conception et l'utilisation des tableaux de bord des enseignants dans des contextes d'apprentissage réels.

Nous avons adopté une approche de recherche mixte itérative et nous nous sommes appuyés sur des recherches de différents domaines, notamment les sciences sociales, l'interaction homme-machine, le design centré-utilisateur, la visualisation de l'information, ainsi que l'analyse de données d'apprentissage.

Dans le contexte de Speakplus, les enseignants sont confrontés à un challenge central : le manque d'awareness (conscience) émotionnelle dans l'apprentissage en ligne, qui est dû à des interactions à distantes et médiées par la technologie. Nous nous sommes appuyés sur des modèles d'émotions pour intégrer de manière transparente, dans la plate-forme, des indices permettant de sensibiliser les enseignants aux émotions des apprenants. Nous avons ensuite conçu de manière itérative Emodash, un tableau de bord interactif, pour aider les enseignants à personnaliser des bilans pédagogiques écrits aux apprenants après chaque session d'apprentissage.

Nous avons étudié les tableaux de bord dans le contexte de Projet-Voltaire. Les enseignants sont confrontés à un challenge central à savoir le manque d'un feedback précise sur les interactions des apprenants en ligne afin d'informer des interventions pédagogiques. Nous avons mené une étude de design (Sedlmair et al., 2012) avec des enseignants où nous avons conçu en collaboration Progdash, un tableau de bord aidant les enseignants à informer leur pratiques enseignantes pour assister l'apprentissage en ligne des apprenants.

Nous nous sommes appuyés sur les résultats de nos études sur Emodash et Progdash pour articuler les facteurs sociaux, les pratiques pédagogiques ainsi que les challenges techniques qui s'intersectent dans le design et l'usages des tableaux de bord à destination d'enseignants (Shum et al., 2019; Ferguson, 2012).

Nous espérons apporter une contribution sous trois angles:

Premièrement, du point de vue de design, nous avons conçu, mis en uvre et déployé deux tableaux de bord fonctionnels centrés sur les enseignants dans des contextes réels, pour étudier les tableaux de bord des enseignants in situ:

- EMODASH, un tableau de bord conçu sur une base d'entretiens avec cinq enseignants et affiné grâce à un processus de design itératif, pour faciliter la prise de conscience rétrospective des émotions des apprenants des enseignants, et aider les enseignants à rédiger bilans pédagogiques après les séances d'apprentissage des langues par vidéoconférence.
- PROGDASH, un tableau de bord conçu sur la base d'entretiens avec des enseignants et affiné via une conception collaborative, pour permettre aux enseignants de mener des interventions pédagogiques informées auprès des apprenants et ainsi mieux les assister à utiliser la plateforme d'apprentissage en ligne.

Deuxièmement, d'un point de vue **empirique**, nous avons conçu et réalisé des analyses qualitatives et quantitatives pour articuler des évidences du terrain sur l'impact des tableaux de bord sur les pratiques enseignantes.

- Nous avons mené plusieurs discussions informelles avec les concepteurs de deux plateformes d'apprentissage de langue en ligne pour comprendre les challenges pédagogiques des enseignants en utilisant ces plateformes dans les établissements scolaires.
- Nous avons mené plusieurs entretiens semi-structurés avec des enseignants pour comprendre leurs besoins, leurs challenges et leurs pratiques pédagogiques à travers lesquels ils visent à améliorer l'apprentissage des apprenants.
- Nous avons conçu et réalisé deux études de design (Sedlmair et al., 2012) couplées avec des évaluations longitudinales de deux et trois mois chacune pour étudier comment 34 enseignants utilisent des tableaux de bord pour informer leurs pratiques enseignantes.
- Nous avons mené plusieurs entretiens semi-structurés avec les enseignants pour évaluer l'appropriation et l'impact de nos artefacts sur les pratiques pédagogiques.
- Nous avons effectué une analyse exploratoire qualitative et quantitative comparant les émotions des apprenants inférées à partir de deux modalités différentes, à savoir les flux audio et vidéo.
- Nous avons effectué une analyse qualitative et quantitative pour (1) comparer les bilans pédagogiques écrits par les enseignants *avec* et *sans* utilisation d'EMODASH; et (2) articuler les pratiques pédagogiques des enseignants lors de l'utilisation de PROGDASH.

Enfin, d'un point de vue méthodologique, nous nous sommes appuyés sur de nombreuses théories,

de méthodes et d'études de recherche pour concevoir et évaluer l'appropriation des tableaux de bord des enseignants dans leur contexte :

- Nous nous sommes appuyés sur des théories et des études sur la reconnaissance des émotions (Ekman and Friesen, 1976; Barrett and Russell, 1998) pour capturer, quantifier et intégrer les indicateurs émotionnels (S. K. D'Mello, 2017) en environnements d'apprentissage en ligne.
- Nous avons proposé une méthode d'unification des modèles d'émotions discrètes et dimensionnelles pour faciliter de les combiner et utiliser de manière multimodale (Ez-zaouia et Lavoué, 2017).
- Nous avons effectué une analyse exploratoire des flux vidéo et audio des séances d'apprentissage par vidéoconférence un total de plus de sept heures enregistrements, pour comparer les émotions des apprenants inférées à partir des deux modalités : audio et vidéo.
- Nous avons effectué une analyse exploratoire de plus de 10 heures d'enregistrements d'interactions des enseignants avec Emodash pour examiner l'expérience des enseignants lors de la rédaction des bilans pédagogiques avec la présence d'émotions.
- Nous avons effectué une analyse thématique qualitative (Braun and Clarke, 2006) des transcriptions des entretiens avec les enseignants pour évaluer l'impact d'Emodash et Progdash
- Nous avons effectué une analyse factorielle quantitative (Wobbrock and Kay, 2016) pour comparer les bilans pédagogiques des enseignants avec et sans l'utilisation d'Emodash.
- Nous avons proposé des implications de design et d'usage qui articule les dimensions sociales, pédagogiques et techniques engendrés par les tableaux de bord des enseignants.

4 Travaux, Résultats et contributions

Au Chapitre 3, nous avons présenté une **étude de cas** examinant une approche multimodale, c'est-à-dire en appliquant deux modèles théoriques des émotions à savoir discrets et dimensionnels ; et en utilisant des APIs sur le cloud pour la reconnaissance des émotions d'apprenants en ligne. Nous avons démontré cette approche dans le contexte de la plate-forme Speakplus pour l'apprentissage de langues par vidéoconférence. Nous avons utilisé les flux audio et vidéo pour inférer automatiquement des émotions ainsi que des émotions auto-rapportées par apprenants. Nous avons contextualisé les émotions moyennant les traces d'interaction d'enseignants et d'apprenants sur la plateforme. Nous avons analysé 10 sessions d'apprentissage de 45 minutes chacune —un total de plus de 7 heures d'enregistrement. Nous avons comparé les émotions inférées des deux modalités audio et vidéo. Nous avons proposé une approche pour combiner ces deux modèles d'émotions discrets et dimensionnels pour capturer les états émotionnels des apprenants de manière globale.

Résultats. En plus de prouver la faisabilité de l'utilisation des APIs sur le cloud pour inférer des émotions dans un environnement d'apprentissage en ligne, nos résultats soulignent la pertinence de l'approche multimodale. Les résultats ont montré que différentes émotions sont détectées à partir de différentes sources, notamment les flux audio et vidéo.

Contributions. Les contributions de ce travail sont :

- Une étude de cas sur l'utilisation des APIs sur le cloud pour mesurer les émotions automatisées des apprenants à partir de l'audio et de la vidéo, ainsi que les émotions auto-rapportés et les traces d'apprentissage contextuel dans l'enseignement en ligne et à distance.
- Une méthode pour unifier les modèles discrets et dimensionnels des émotions.
- Une analyse exploratoire qualitative et quantitative comparant les émotions inférées à partir des modalités audio et vidéo.

Dans le Chapitre 4, nous nous sommes appuyés sur notre exploration technique des émotions et avons suivi un processus de conception itératif au cours duquel nous avons collaboré avec un responsable pédagogique de SPEAKPLUS, un chercheur en langue et deux enseignants. Au cours du design, nous avons effectué deux tests d'utilisabilité formels menant au design d'EMODASH. Nous avons intégré le tableau de bord dans l'éditeur de bilan pédagogique de SPEAKPLUS. Nous avons mené une étude de terrain de deux mois avec cinq enseignants professionnels et cinq apprenants pour évaluer (1) comment EMODASH aide rétrospectivement les enseignants à prendre conscience des émotions d'apprenants en ligne et (2) comment EMODASH impact le type du feedback que les enseignants rédigent aux apprenants après chaque session d'apprentissage. Nous avons spécifiquement comparé les bilans pédagogiques écrits par les enseignants sous deux conditions : *avec* et *sans* EMODASH.

Résultats. Les résultats de ce travail ont montré que la plupart des enseignants ont réagi positivement à EMODASH. Les résultats ont montré que EMODASH a conduit les enseignants à rédiger significativement plus de feedback affectif et formatif et moins de feedback sommatif, suggérant une meilleure conscience des émotions d'apprenants. De plus, le tableau de bord a conduit les enseignants à réfléchir sur la façon dont ils mènent leurs séances de cours, en utilisant les émotions d'apprenants comme un proxy pour évaluer leurs pratiques enseignantes.

Contributions. Les contributions de ce travail sont :

- EMODASH, un tableau de bord conçu sur la base d'entretiens avec cinq enseignants et un apprenant, et affiné via un processus de conception itérative, pour faciliter la prise de conscience rétrospective

des enseignants sur les émotions des apprenants, et soutenir la rédaction par les enseignants de rapports de rétroaction de renforcement aux apprenants après les séances d'apprentissage des langues par vidéoconférence.

- Une étude sur le terrain à l'état sauvage d'Emodash. Cette étude étend notre compréhension de la façon dont les tableaux de bord émotionnels soutiennent la conscience rétrospective des émotions des apprenants dans un environnement d'apprentissage des langues par vidéoconférence.
- Études qualitatives et quantitatives comparant les rapports de rétroaction des enseignants sous deux conditions avec et sans Emodash.

Dans le Chapitre 5, pour mieux informer le design de tableaux de bord à destination des enseignants dans des contextes d'apprentissage mixtes où les enseignants articulent à la fois des activités d'apprentissage en ligne et en classe, nous avons mené une **étude de design** basée sur des entretiens avec sept enseignants pour comprendre leurs besoins. Nous avons collaboré avec des enseignants autour des prototypes de design. Nous avons implémenté et intégré Progdash dans la plate-forme Projet-Voltaire pour l'apprentissage du vocabulaire, de la grammaire et de l'orthographe du français. Nous avons ensuite mené une étude longitudinale de trois mois avec 29 enseignants pour évaluer (1) si Progdash fournit aux enseignants des indicateurs utiles sur la progression des apprenants et (2) comment Progdash impact les pratiques enseignantes.

Résultas. Nos résultats montrent que la plupart des enseignants ont réagi positivement au tableau de bord. Ils ont trouvé Progdash utile pour informer plusieurs pratiques pédagogiques : le suivi des apprenants, la conduite des évaluations, la planification des interventions, le partage d'information en classe, la provision du feedback. Nos résultats mettent en évidence un fossé entre les enseignants spécialisés dans les concepts de la plateforme d'apprentissage, c'est-à-dire la langue française, et ceux utilisant la plateforme dans des cours nécessitant des compétences de rédaction, par exemple le marketing, même si les deux étaient soucieux d'atteindre les mêmes objectifs, notamment aidant les apprenants à maîtriser l'orthographe et la grammaire. Sur la base de nos résultats, nous fournissons des implications visant à améliorer le design des tableaux de bord pour mieux faciliter l'apprentissage en ligne et en classe et ainsi que pour favoriser le dialogue et la réflexion des enseignants et des apprenants.

Contributions. Les contributions de ce travail sont :

- Une étude de conception de Progdash pour aider les enseignants à utiliser une plate-forme d'apprentissage en ligne intégrée au curriculum.
- Une étude longitudinale de Progdash de trois mois avec 29 enseignants, qui étend notre compréhension de la façon dont les enseignants intègrent un tableau de bord dans leurs pratiques pour articuler l'apprentissage à distance et en classe.

- des implications de design des tableaux de bord pour faciliter un suivi d'apprentissage en ligne et en classe et informer les pratiques enseignantes.

Dans le Chapitre 5, nous avons mené une enquête réflexive en nous appuyant sur les résultats de nos études sur EMODASH et PROGDASH. Nous avons articulé les dimensions sociales —monitoring (suivi), *awareness* (conscience) et réflexion, les dimensions pédagogiques —planification, feedback et coaching, ainsi que les dimensions techniques engendrés par le design et l'utilisations des tableaux de bord à destination des enseignants. Dans chaque dimension, nous avons structuré et discuté des appropriations pratiques. Nous croyons que cela fournit une vue globale des implications macro-conceptuelles de l'espace de design des technologies d'information en système éducatif, et aide à faciliter la considération de facteurs sous-jacents sur le plan social, pédagogique et technique.

En résumé, nous espérons apporter les contributions globales suivantes :

Du point de vue du **design**, notre recherche fournit des résultats importants sur le design des tableaux de bord des enseignants. Nous présentons deux études illustrant un processus de conception de bout en bout de tableaux de bord destinés aux enseignants dans des conditions réelles, ainsi que les étapes, les opportunités et les challenges sous-jacents. Nous illustrons un tel processus dans deux environnements d'apprentissage différents en situant les besoins et les contextes d'utilisation des enseignants, en distillant les choix de design, en décrivant les abstractions de données, ainsi que les encodages visuels lors de la mise en forme des différentes vues de chaque tableau de bord. De plus, nous fournissons des implémentations et des architectures fonctionnelles de collecte, d'analyse et de visualisation des données d'apprentissage dans EMODASH et PROGDASH qui peuvent s'avérer utiles pour d'autres études.

D'un point de vue **empirique**, nos résultats mettent en évidence l'utilité, l'efficacité et, plus important, l'impact des tableaux de bord sur les pratiques quotidiennes des enseignants. Nous montrons différentes routines des enseignants entourant l'utilisation des tableaux de bord pour assister les apprenants en ligne. Nous démontrons les expériences, les perceptions et les appropriations des enseignants vis-à-vis les tableaux de bord par dans les contextes d'apprentissage authentiques en ligne. Nous fournissons un illustrons des avantages des tableaux de bord pour le monitoring, l'awareness, la réflexion, la planification, le feedback et le coaching.

D'un point de vue **méthodologique**, nos recherches ont été théoriquement et méthodologiquement informées de différents domaines. Nous avons articulé des implications conceptuelles sur les dimensions sociales, pédagogiques et techniques des tableaux de bord pour les enseignants. Pour chaque dimension, nous avons structuré et discuté des pratiques d'appropriation pour informer le design et les usages des tableaux de bord. Ces dimensions articulent un cadre conceptuel global

pour la technologie d'information en éducation et ainsi des macro-implications pour des tableaux de bord adaptés aux besoins et situations d'apprentissage et d'enseignement.