How Do Teachers Use Dashboards Enhanced with Data Storytelling Elements According to their Data Visualisation Literacy Skills?

Stanislav Pozdniakov
Monash University
Australia
stanislav.pozdniakov@monash.edu

Roberto Martinez-Maldonado
Monash University
Australia

Yi-Shan Tsai
Monash University
Australia

Vanessa Echeverria
Monash University
Australia
Escuela Superior Politécnica del Litoral
Ecuador

Namrata Srivastava
Monash University
Australia

Dragan Gasevic
Monash University
Australia

ABSTRACT

There is a proliferation of learning analytics (LA) dashboards aimed at supporting teachers. Yet, teachers still find it challenging to make sense of LA dashboards, thereby making informed decisions. Two main strategies to address this are emerging: i) upskilling teachers' data literacy; ii) improving the *explanatory* design features of current dashboards (e.g., adding visual cues or text) to minimise the skills required by teachers to effectively use dashboards. While each approach has its own trade-offs, no previous work has explored the interplay between the dashboard design and such "data skills". In this paper, we explore how teachers with varying *visualisation literacy* (VL) skills use LA dashboards enhanced with (explanatory) data storytelling elements. We conducted a quasi-experimental study with 23 teachers of varied VL inspecting two versions of an authentic multichannel dashboard enhanced with data storytelling elements. We used an eye-tracking device while teachers inspected the students' data captured from Zoom and Google Docs, followed by interviews. Results suggest that high VL teachers adopted complex exploratory strategies and were more sensitive to subtle inconsistencies in the design; while low VL teachers benefited the most from more explicit data storytelling guidance such as accompanying complex graphs with narrative and semantic colour encoding.

CCS CONCEPTS

• **Applied computing** → *Collaborative learning*; *Computer-assisted instruction*; *Learning management systems.*

KEYWORDS

data storytelling, data literacy, human-centred design, learning analytics, dashboard

ACM Reference Format:
Stanislav Pozdniakov, Roberto Martinez-Maldonado, Yi-Shan Tsai, Vanessa Echeverria, Namrata Srivastava, and Dragan Gasevic. 2023. How Do Teachers Use Dashboards Enhanced with Data Storytelling Elements According to their Data Visualisation Literacy Skills?. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3576050.3576063

1 INTRODUCTION

There is a sustained proliferation of learning analytics (LA) dashboards aimed at supporting teachers [40, 44, 51]. Indeed, dashboards can play a critical role in closing the LA loop by providing teachers with enhanced capabilities to monitor and support learning tasks in contexts such as online learning [14], K12 classrooms [25], game-based learning [20], collaborative learning [2], and adaptive learning [47]. Nonetheless, there is growing evidence in LA literature that teachers can raise wrong conclusions on students' progress [8], and find it challenging to make sense of student activity and make informed decisions [32, 52] based on teacher-facing dashboards. It has also been found that some LA dashboards are either insufficiently accurate [21, 39] or require specialised skills or training to use them [21, 46]. In other cases, the design of such interfaces does not always align with teachers' pedagogical needs [22, 43].

One strand of LA research has suggested that effective use of LA dashboards involves strategies to improve teachers' data literacy skills [53] so they can understand the inner workings of the LA visualisations they interact with. The term *data literacy for teachers* has been coined as their ability to transform information into actionable instructional knowledge and practices by collecting, analysing, and interpreting various types of data [29]. Several authors suggest that teachers lacking data literacy can end up making poor interpretations from such analytics and, thus, ill-informed decisions that can significantly affect students (see review in [33]). However, the

LAK 2023, March 13–17, 2023, Arlington, TX, USA

ACM ISBN 978-1-4503-9865-7/23/03…$15.00
https://doi.org/10.1145/3576050.3576063

impact of data literacy on the ways teachers use dashboards is yet to be empirically investigated. Moreover, it can be argued that the construct of *visualisation literacy* (VL) [27] – as part of broad data literacy (see Section 2.2 for links between data and visual literacy) – more specifically targets the skills required by teachers to interact with LA dashboards.

In contrast, it has also been argued that teachers need to focus on instruction instead of being required to get training experience in data literacy to perform data collection and analysis [50]. Thus, an alternative LA research strand, grounded on information visualisation and human-centred design principles, argues that teachers do not need to become fully proficient in data analysis skills in order to use LA dashboards effectively. Instead, they need to be guided [15]. The *explanatory* design features of LA dashboards could be enhanced to lower the requirements for teachers without formal data analysis to interpret and make sense of the dashboard content [28]. This can be achieved, for example, by adding data storytelling elements to the visualisations (such as visual cues [16], textual advice [49], by linking visualisations meaningfully [10]) to drive teachers' attention to critical insights that can be gained from the data, or by generating fully written reports instead of showing charts in a dashboard [18].

While each of these strands has its own trade-offs, the potential interplay between the dashboard design and the "data skills" that teachers require to use LA dashboards effectively is yet to be explored. We address this gap by exploring how teachers with different visualisation literacy (VL) skills use LA dashboards with data storytelling enhancements. We conducted a quasi-experimental study with 23 teachers with varied VL skills inspecting two versions of a LA dashboard enhanced with data storytelling elements in the context of synchronous online collaborative learning. We applied eye-tracking and qualitative methods to identify potentially distinct strategies followed by teachers with low and high levels of VL to make sense of the dashboards.

2 BACKGROUND

2.1 Visualisation Guidance and Data Storytelling in Education

Visualisation guidance refers to design features aimed at reducing the complexity and cognitive load associated with the interpretation and sense-making of data visualisations [42]. Guidance in this sense involves the provision of supportive elements in the data interface itself (e.g., in the form of direct annotations, signposting or emphasised elements) to orient or steer the interaction with the data representations thus reducing the need for training end-users on the selected visualisation techniques or, in other words, narrowing their knowledge gap [9]. One way to provide such visualisation guidance is by applying the notion of *data storytelling* (DS). DS is an information compression technique that can be applied to help an audience focus on responding to specific questions by combining charts, text and other resources to emphasise the data points and the evidence that are more relevant to such questions [34].

There has been a growing interest in applying DS elements to improve the explanatory features of data visualisation to support teachers' interaction with LA interfaces [24]. For example, Chen et al. [10] proposed a drill-down narrative storytelling approach to guide teachers in navigating various “data slices” generated from MOOC data. Similarly, Pozdniakov et al. [37] suggested to explicitly link each chart in a teacher's dashboard with a specific question to provide context to the visualisations of online discussions. Martinez-Maldonado et al. [30] proposed an approach to emphasise key data points in a multimodal LA visualisation to pace down the order in which teachers and students discuss the insights from the data. Van Leeuwen and Rummel [49] found that enriching a dashboard with advisory text helped teachers to require less effort and gain more confidence when making decisions about identifying groups of students that needed more help. Finally, Fernandez Nieto et al. [18] investigated the extent to which alternatives to the notion of visual techniques in a dashboard (i.e., tabular visualisations or a written report) could also serve to support teachers in interpreting the insights from the data. These works jointly suggest an emerging interest in providing visualisation guidance to help teachers' data interpretations. Notably, only one small-scale DS study with 6 teachers has employed eye-tracking analysis to inspect how teachers' attention can be steered in a dashboard by adding annotations, minimising colour variety and explaining trends with text [16]. Yet, the authors did not investigate the potential role of teachers' visualisation skills in the way they used dashboards. The importance of such skills is that they are specifically required by teachers to interact with LA dashboards.

2.2 Data and Visualisation Literacy

There is an emerging interest in improving the skills required for teachers to work with data [33] and data visualisations [1]. In LA as a field, this interest has been related to the concept of data literacy. Mandinach and Abrams [29] used the term data literacy for teachers to refer to the ability to transform information into actionable instructional knowledge by collecting, analysing, and interpreting all types of data. Ndukwe and Daniel's review [33] demonstrates this increased interest in developing teachers' data literacy skills for the effective use of data-driven innovations. Yet, it has been argued that teachers should not focus on low-level data tasks, such as data collection and analysis, which require quite specialised skills and may distract teachers from their already numerous instructional tasks [50].

The construct of *visualisation literacy* (VL) [27] more specifically targets the skills required by teachers to interact with LA visual interfaces. *Visualisation literacy* refers to the ability to read and interpret visual representations of data; and extract, process and make conclusions based on the information from such visualisations [19]. Notably, instruments such as the visualisation literacy assessment test (VLAT [27]) have been developed and validated to quantitatively measure visualisation literacy. This has been used to investigate the potential impact of novel visualisation techniques (e.g., magazine-style narrative visualisations [26]) on users' comprehension according to their individual VL skills.

To the best of our knowledge, no previous studies have explored the potential role of teachers' visualisation literacy in the ways teachers use LA dashboards. The closest work to ours was conducted by Donohoe and Costello [13], who contrasted teachers' reported and actual visualisation literacy skills measured using VLAT [27]. Donohoe and Costello found that teachers are typically overconfident about their own skills. Another close work tackling this gap was conducted by van Leeuwen et al. [47] who found no effect between teachers' technological self-efficacy (self-reported by teachers using a 5-point likert scale), teachers' characteristics (i.e., age and experience), and the ways they used a LA dashboard. Consequently, the potential interplay between the dashboard design enhancements and teachers' visualisation literacy skills is yet to be explored. Against the literature discussed above, we address this gap by raising the following research questions: **RQ1:** How do teachers with varied visualisation literacy skills use LA dashboards with data storytelling enhancements?; and **RQ2:** To what extent do data storytelling elements support teachers in making sense of students' activity?

3 METHOD

3.1 The Learning Situation

We addressed our research questions in the context of the human-centred design process of a teacher-focused dashboard deployed as a part of the graduate course *IT research methods* at *undisclosed university* to support teachers in monitoring synchronous online learning tasks. In this course, students engaged in weekly 3-hour online classes facilitated by two teachers. Students commonly performed a range of tasks by forming small groups of 4-6 members, using audio-visual and text-based tools (namely, Zoom and Google Docs). The types of tasks ranged from those where students were expected to discuss a topic to those where the main output was a written report. When students went to *breakout rooms* to work on group tasks, teachers often had difficulty knowing what was happening in each room unless they join and spend some time listening to the conversations or check the documents generated by the group. This is a widespread monitoring challenge that has been reported in classroom orchestration literature [48]. Moreover, although Zoom has become one of the staple communication tools to conduct synchronous, online learning activities, it does not offer orchestration features for teachers to "see" what is happening when multiple groups work simultaneously [45].

3.2 Apparatus and Dashboard Design

To overcome the current monitoring limitations in Zoom, a dashboard was iteratively designed with teachers delivering the classes of the aforementioned course. This paper focuses on two iterations of this process which led to two versions of a monitoring dashboard (version 1 and version 2 in Fig. 1). The dashboard uses the data captured by an open-source analytics system called ZoomSense[1] [4]. This deploys virtual agents to each Zoom breakout room to automatically capture which student is speaking at each moment. If configured by the teachers, these agents automatically create a Google Doc associated with each room, providing the link to students. All actions and major revisions performed on the document are also logged. The first design was evaluated with a small number of teachers in an authentic deployment of the dashboard, reported in [38], in terms of usefulness, ease of use and envisaged use. Version 2 is a revised version based on this preliminary feedback from teachers. The use of data storytelling elements has not been previously evaluated. These and the differences between versions 1 and 2 are described below.

3.2.1 Data storytelling elements. Based on [16], we implemented both versions of the LA dashboard with the following data storytelling elements: DSa) *teacher's questions* (defined by teachers in a previous study focused on the design of the dashboard [36]) as the titles of corresponding charts (see Fig. 1, Ch1-5); DSb) *colour emphasis* (or semantic colour encoding), where three contrasting colours are consistently used to emphasise elements in a colour-blind friendly manner (grey and navy blue used for most visual elements, whereas orange and red are used to emphasise cases that may need closer attention); and DSc) *textual narrative* (see Ch1 and Ch4 in Fig. 1) explaining students' behaviours.

3.2.2 Interface design. Both dashboard versions provide the same information using six and five charts, respectively. The first is a timeline chart shown at the top of the interface depicting the activity where the class is (see T on version 1 in Fig. 1 containing three activities and the start and final times of the class). The number of activities can be configured by the teacher at the beginning of each class to reflect their class script. The activity automatically advances every time the students are asked to join breakout rooms. Zoom and Google Docs data is presented for the current activity selected by the teacher (e.g., activity 3 is emphasised in navy blue over grey in Fig. 1).

Other charts respond to the teacher's questions (Ch1-5). Ch1 in version 1 is associated with a narrative-based chart that classifies students based on the Zoom speech data as: active (if the student is talking), inactive (not talking) and dominant/quiet (if one is doing most of the talk or not talking much compared to the others, respectively). Ch2 in both versions is equally addressed with a horizontal state bar, which signals the progress in the Google Docs. To measure this, the system generates text placeholders in the document for students to complete each activity, which are then marked in blue, orange or grey if each has been completed, attempted or not attempted, respectively. Ch3 in both versions is also depicted as a horizontal bar chart, this time indicating the amount of speech detected in each group. If a group has spoken more than the average across all the groups, their bar is automatically emphasised in orange.

Ch4 is addressed with sociogram charts (one for each breakout room), where each node depicts a student or a teacher (shown as "tutor" if they joined the breakout room). The thickness of a link between two nodes represents the amount of time two participants communicated with each other. Some nodes are automatically emphasised if in the last 5 minutes, a student: i) did not communicate (shown in red); or ii) communicated considerably more than the rest of their peers (shown in orange). Finally, Ch5 is addressed in both versions with a bar chart that illustrates the total amount of time (in minutes) a teacher spent in a breakout room, where the bar corresponding to the room in which a teacher spent the most time is emphasised in orange.

3.2.3 Changes between versions. In version 2, sociograms (Ch4) were enhanced with the text shown in version 1 (Ch1), thus the narrative-based chart was removed as shown in Fig. 1 (bottom). This

[1] https://zoomsense.io

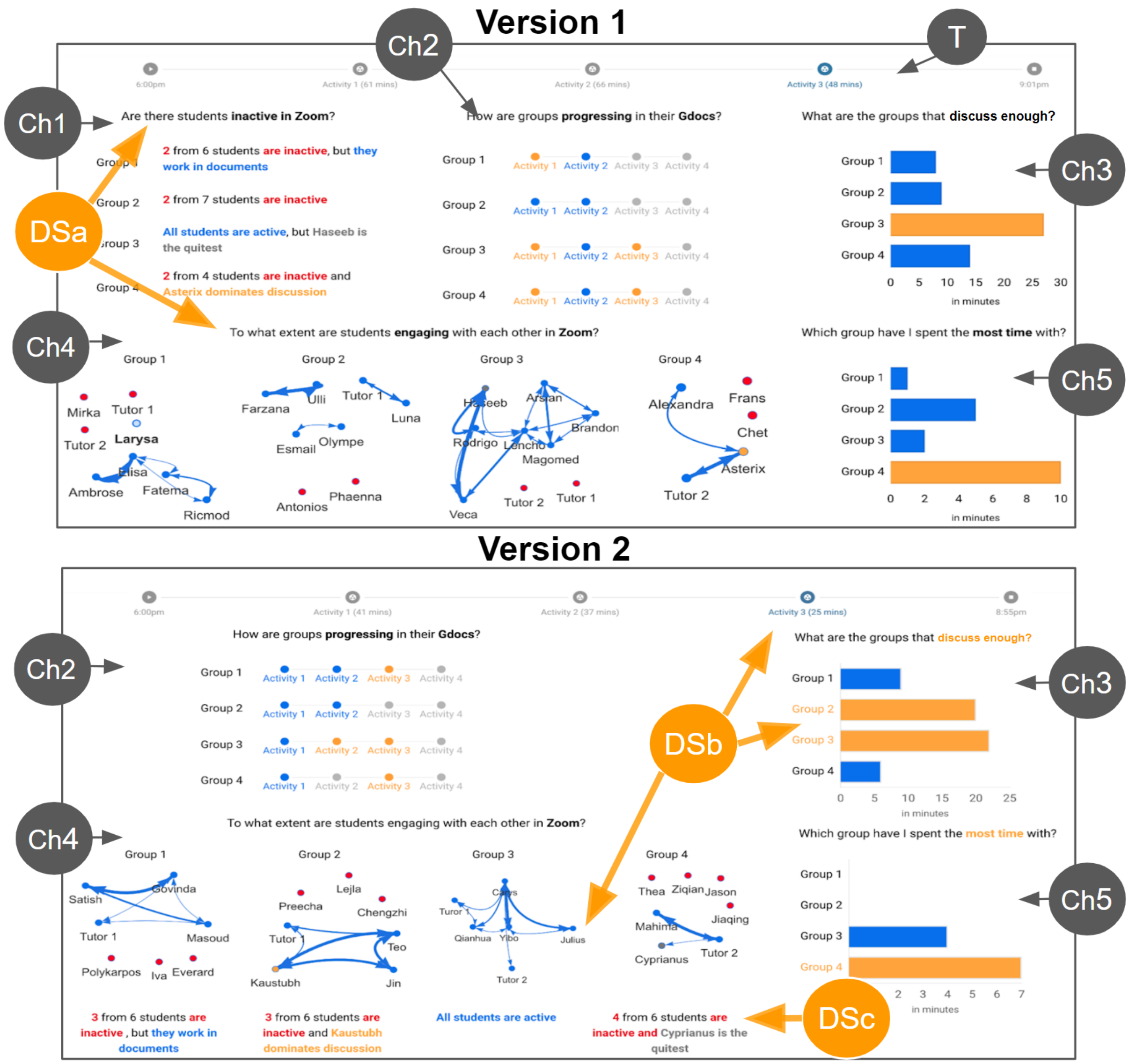

Figure 1: Two versions of the dashboard enhanced with data storytelling elements. Each chart (Ch1-5) is associated with a specific question (QS1-5). Data storytelling elements are instantiated via DSa) questions as chart titles; DSb) emphasised elements such as highlighted bars in charts Ch3 and Ch5, and nodes in charts T, Ch2 and Ch4; and DSc) textual summaries.

was requested by teachers in the design study [38] who explained they needed more time to interpret sociograms or a summary of associated metrics. This is also in line with recent attempts to enhance comprehension of more complex charts by combining them with textual narratives and semantic colour encoding [11, 26]. Additional emphasis was added to the bar chart labels (e.g., Fig. 1 – orange labels in Version 2 – Ch3) aiming to drive attention to areas that teachers should pay attention to. This resulted in, to some extent, version 2 being more decluttered than version 1.

3.3 Participants

A total of 24 teachers (13 females, 10 males) between 22 and 60 years old (*M*=31, *SD*=8) participated in the study. Six had been involved in the design process of version 1. All teachers except one had completed a postgraduate degree (7 masters, 12 PhD students, 2 with a PhD degree), and were either currently employed as teachers in the university or had teaching experience (*M*=3, *SD*=3.4). All teachers held a STEM degree.

3.4 Study procedure

The study consisted of the following three parts outlined in Figure 2: 1) eye-tracking analysis of the dashboard inspection; 2) a post-hoc individual interview; and 3) visualisation literacy assessment. These are described next.

Part 1: Eye-tracking study with dashboard inspection. We conducted an eye-tracking lab study to explore how teachers navigate through different data storytelling elements of the LA dashboard (RQ1). We used Tobii Pro Nano 60Hz eye-tracker hardware and software[2]. The eye-tracker was placed at the bottom of a 24" desktop monitor, with a resolution of 1920x1200. In addition to this, we set up an external video camera to capture audio data. Prior to the study, the eye-tracker was calibrated for each participant. We attempted to ensure that during the study participants did not change their position by providing a stationary chair. This step resulted in the collection of eye-gaze data. These consisted of timestamped x-y gaze datapoints. For every participant, the percentage of identified gaze samples was above 75%, which means that the rest 25% of the samples were removed and not included in the analysis since they were outside the dashboard area. We used Tobii Pro software with the default Velocity-Threshold Identification (I-VT) filter [35] to classify raw data into fixations and saccades. This filter relies on velocity to classify a sample either as a fixation (when teachers' eyes stop scanning the dashboard) or a saccade (rapid gaze movements from one point of interest to another). We used default values such as window size set to 3 samples and velocity threshold set to 30 degrees/per second [35].

This part followed a within-subject design. All participants were exposed to versions 1 and 2 of the LA dashboard. To account for an *order effect*, i.e. participants preferring the first version presented to them, we followed a two-by-two Latin square matrix design to pseudo-randomly assign the order in which the versions are shown to participants. To account for a *carry-over effect*, i.e. participants memorising the data after the exposure to the first LA dashboard and confounding the use of the subsequent dashboard, we had balanced the data from two classrooms to be equally presented to teachers in each dashboard version, such that each time a participant was exposed to a new LA dashboard, data from a different classroom session was presented to the participant. To minimise teachers' getting distracted by the web browser used to display the dashboards, and potential ethical issues, screenshots taken from the live dashboards were presented to teachers with students' names pseudonymised. Before the eye-tracking study, a short training activity was provided. In this, teachers were briefly walked through the charts presented in the version which was shown to them first explaining the meaning of each element, colour and text on the screen. They were also given a training task, where they were asked to read the description of the learning activities of a class session and explore the dashboard to assess the data in light of the provided learning design.

Then, the teachers proceeded with the eye-tracking study. In this, they were asked to assess students' engagement based on the versions of the dashboard. To scaffold this task, a set of prompting questions were used to invite teachers to interpret the visualisations, for example: *Which group or groups have the most inactive students?*, *Which group is the most engaged in the discussion?*, *In your opinion, which group or groups might have experienced issues with progressing in the learning activity?* The participants were asked the same set of questions for both versions.

Part 2: Post-study interviews. To address RQ2, a set of semi-structured interviews with 13 teachers (8 female and 5 male) were conducted given their availability and willingness to participate shortly after the study. Participants were equally distributed across VL groups (6 in the low VL and 7 in the high VL group). The same set of questions was asked for each data storytelling element described in Section 3.2, namely a) teachers' questions, b) colour emphasis, and c) textual narratives) to explore the participants' perspectives. For instance, a set of questions asked to the participants regarding the colour emphasis in one of the charts looked like: *"To what extent would you consider [colour emphasis] being helpful for you to understand what the chart was communicating about students' activity? If so, which colour emphasis was helpful? How did it help? Did it make the interpretation of the chart any easier?"*. Each interview lasted around 15 minutes. The interviews were conducted face-to-face immediately after Part 1 and all were video-recorded.

Part 3: Visualisation literacy assessment. To address RQ1, we applied the validated VLAT test ([27]) to assess teachers' visualisation literacy. The test contains 53 multiple-choice questions that require the interpretation of various data charts (e.g., bar and line charts among several others). A limit of 40 seconds per question is set to make it a standardised test as recommended in [27]. Teachers were able to skip a question if they were unsure about the correct answer. In total, 23 teachers completed the test, which resulted in a final visualisation literacy score per participant, where the maximum possible score was 53 ($Q1$=35, Mdn=39, $Q3$=45, min=17, max=48). Since VLAT could activate or improve VL skills because teachers get exposed to several charts, we administered the test 1-3 days after parts 1 and 2 to minimise bias.

3.5 Analysis

Visualisation literacy test. We used the median value of visualisation literacy scores for all participants (Mdn = 39) as a threshold to attribute participants to low or high visualisation literacy (VL) groups. We ended up with a total of 12 participants in the low VL ($Q1$=30, Mdn=35, $Q3$=37, min=17, max=9) group and 11 participants in the high VL ($Q1$=43.5, Mdn=45, $Q3$=46, min=40, max=48) group.

RQ1 – Eye-tracking study. To analyse teachers' visual attention during the interpretation, we defined *areas of interests (AOIs)*, which are meaningful regions of the interface. We defined 6 AOIs associated with the charts in both versions 1 and 2 (2 AOIs were defined for Ch4 to inspect the sociographs and textual summaries separately in version 2). Additional AOIs were defined for the questions that served as titles of each chart (5 and 4 for versions 1 and 2, respectively).

The analysis tool EMDAT [23] was used to extract two sets of features. The first set included **Non-AOI based** features that described how each interface was generally explored by teachers, such as the number of fixations, mean fixation duration, number of saccades, average saccade distance (*meansaccadedistance* in pixels) and average saccade speed (*meansaccadespeed*, pixel/milliseconds). The second set included **AOI-based** features that provided a more detailed account of teachers' exploration strategies regarding individual charts and data storytelling elements. These features measure, for example, the duration of fixations that are inside an AOI in milliseconds (*<area_of_interest>_totaltimespent*), time before the first fixation inside an AOI (*'<area_of_interest>_timetofirstfixation'*), and the proportion of fixations that are inside an AOI (*'<area_of_interest> _proportionnum'*). This set of features also accounted for transitions

[2]https://www.tobiipro.com/product-listing/nano/

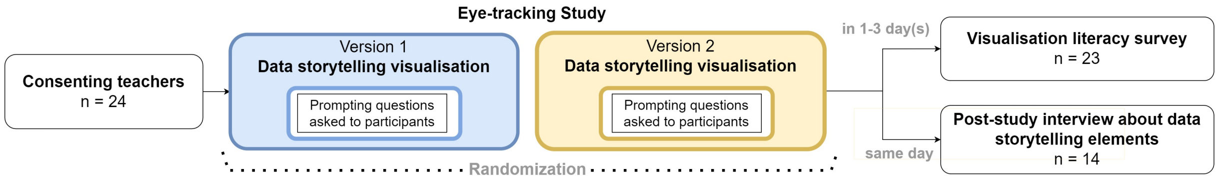

Figure 2: Outline of the study procedure.

between AOIs such as absolute and relative measures of transitions from one AOI to another. A *transition* was a saccade with its first fixation in one AOI and the end fixation in another AOI. Transition-based features were calculated based on saccades from all other active AOIs to <area of interest A>. For each AOI, the following two features were extracted: *<area of interest A>_numtransfrom*, which is the number of transitions from all other AOIs to area A; and *<area of interest A>_proptransfrom*, which is the proportion of transitions from the other AOIs to A overall transitions.

In addition to features provided by EMDAT, we calculated the number and proportions of pairwise transitions, e.g., transitions from '<area of interest A> to <area of interest B>'. Importantly, we disregarded the directionality of such transitions, thus '<area of interest A>_<area of interest B>' and '<area of interest B>_<area of interest A>' were treated as the same, e.g. *Count_Ch4_T* and *Prop_Ch1_Ch4*. We decided to proceed with such simplification to maximise the interpretability of our analysis against the research question. Accordingly, we also calculated the following features: a) number and proportion of transitions from all the questions to the charts (and vice versa), which are *Count_QS_to_other_charts* and *Prop_QS_to_other_charts*; b) number and proportion of transitions from the charts to charts only, which are *Count_chart_to_other_charts* and *Prop_chart_to_other_charts*. The full list of features extracted is available online.

When describing the results, we refer to the AOIs, which contained questions as *QSn*; and to those containing charts as *Chn*, where *n* varies from 1 to 5, and a timeline, *T*. We proceeded with the analysis based on a non-parametric alternative to t-test, the Wilcox test, since the t-test assumptions were not met for all the features. We did not apply any p-value correction method. Instead, we provide effect sizes for the findings, weigh findings against the previous literature [5, 17], and triangulate quantitative results with qualitative findings. Along with the results of the tests, we report the corresponding effect sizes (Pearson's r). In the next section, we report only the results of comparisons where there was a statistically significant difference between low and high VL groups.

RQ2 – Post-study interviews. A thematic analysis of transcribed interviews was conducted separately by two researchers [6]. This was followed up by the discussion of an initial coding technique [41], where a first-level code represents one of the categories of data storytelling elements we asked participants about (i.e., the role of: i) teachers' questions; ii) colour emphasis; and iii) textual narrative). A fourth code emerged from the first code (code: emphasised elements) since several teachers talked comprehensively about the use of colour mapping in the LA dashboard (code: 'colour mapping'). Each first-level code included two second-level codes to contrast the perspectives of teachers in the high and low VL groups. For code 'colour mapping', we applied an affective coding to identify participants' sentiments [41] towards the enhancements. The coding scheme is available online. Once the agreement on the codes was reached, researchers then divided the interviews and coded the rest of them separately. We have not calculated inter-rater reliability and instead ensured reliability by concurrently confirming and reviewing results by two co-authors and triangulating different data [31]. Results are presented in the next section.

4 RESULTS

4.1 RQ1: Quantitative results

4.1.1 Differences in non-AOI based features. When low and high VL groups used version 1, significant differences were observed for *mean saccade distance* and *mean saccade speed*. The mean value of saccade distance for high VL group was higher by 58 pixels compared to low VL group (Mdn_lowVL=270 and Mdn_highVl=328 in Table 1-1). This suggests that high VL participants showed a more pronounced information searching behaviour. They tried to either make comparisons within a single chart or compare evidence from various charts. In addition, the high VL group was characterised by faster gazing strategies (higher median saccade speed by 0.5 px/ms – Mdn_lowVL=3.4 and Mdn_highVL=3.85), which suggests that they may have grasped the visual clues from the visualisation more rapidly compared to low VL group. In contrast, for version 2, there were no statistically significant differences in their high-level gazing behaviour. Additional differences were found in the AOI-based features, which are presented next.

4.1.2 AOIs: Questions as chart titles. There was a significant difference in the ratio of transitions from questions to the other charts between low and high VL participants (Mdn_lowVL=5% and Mdn_highVL=8% respectively, see 'Prop_QS_to_other_charts' in Table 1-2) when using version 2. Secondly, there were at least two questions for which there were statistically significant differences in terms of how each group relied on them. The low VL group made the first visit to question *QS2* after 14 seconds, while the high VL group almost immediately gazed over the same question (Mdn_lowVL=14sec and Mdn_highVL=84ms) when the dashboard was first presented. Similarly, the low VL participants spent around a second gazing on the *QS4* title, as compared to 3 seconds for the high VL group. This suggests that the high VL teachers tended to use the questions as intended, reading them first so they can serve as anchors to understand what each chart is about. This behaviour was captured by the higher proportion of transitions from questions to the charts during their interpretations of high VL teachers. Counter-intuitively, we would expect that the high VL group would disregard the use of questions during their interpretation since they are more familiarised with the charts. However, their eye gaze behaviour showed that the high VL participants explored charts with more scrutiny against the corresponding questions. For instance, they might have been double-checking whether their initial interpretation was aligned with the expected insights they should get from the visualisation.

4.1.3 AOIs: Timeline and emphasised elements. The high VL teachers had the median value of 2.5 and 3 higher for seconds gazing at the timeline of activities (T) compared with the low VL group when using version 1 (Mdn_lowVL=2.6sec and Mdn_highVL=5) and version 2 (Mdn_lowVL=2.5 and Mdn_highVL=5.5sec), respectively (see Table 1-3). Following this, the low VL participants had the median value of 3 transitions from *Ch2* to the timeline, while median value for the high VL was 6 such transitions. Similarly, the high VL group had three times higher median value of transitions between a timeline and visualisation *Ch4* compared with the low VL group (Mdn_highVL=5 and Mdn_lowVL=1.5). This suggests that the high VL participants recognised that an effective interpretation of some highly contextualised charts, such as the chart showing the progress in the google document (*Ch2*) or the sociograms (*Ch4*), should be approached in relation to the current activity students were working on as emphasised in the timeline.

4.1.4 AOIs: Textual summaries added to sociograms. In version 1, high VL teachers had the median value of 9 more transitions from the textual summaries chart (*Ch1*) to the sociograms (*Ch4*) compared to the low VL teachers (Mdn_highVL=18 and Mdn_lowVL=8.5, see Table 1-4). When the participants used version 2 (where such charts got combined into one), the proportion of transitions from sociograms to any other visualisation differed significantly between the low and high VL participants. Precisely, the low VL group had the median value of 35% of the total transitions from or to a sociogram (*Ch4*) while high VL participants did only 26% of the same transitions. Notably, the percentage of such transitions was persistent for the high VL group across the two versions, while for the low VL group it increased from 30% to 35%. Importantly, in contrast to version 1, in version 2 there was an absence of significant difference in transitions, both net sums and proportions, between textual summaries (*Ch1*) and sociograms (*Ch4*) for participants with low and high VL (see Table 1-3). These results suggest that when the low VL group used version 1, they did not recognise that textual summaries charts portray nearly the same information as the sociograms. In contrast, the high VL group seemed to have recognised that textual summaries could help summarise the insights from the sociograms and thus their gaze moved frequently from *Ch1* to *Ch4* and vice versa. Such a pattern changed when the low VL group used version 2, where they might have recognised that *Ch1* served to explain the information in *Ch4*. This might mean that the low VL group benefited more from the version 2 design because such participants were capable to map textual summaries to the sociograms, which resulted in the increased number of transitions not present when inspecting version 1.

Overall, the quantitative results suggest that the high VL participants were better equipped to adopt more complex exploratory strategies, such as identifying which charts should be inspected together and using the questions associated with each chart to guide their interpretations. But the low VL teachers benefited the most from more explicit data storytelling guidance, for example, moving textual explanations closer to the sociogram charts. Yet, the quantitative analysis did not shed light regarding to what extent other data storytelling elements (such as emphasis using colours) shown inside some charts may have been used during the interpretation process. These are addressed in more detail in the next section.

4.2 RQ2: Qualitative results

In this section we present our results from the qualitative analysis regarding the role of i) questions as titles (DSa); ii) emphasised elements (DSb), iii) colour semantics (DSb), iv) textual summaries (DSc).

4.2.1 Questions as chart titles. Participants with both high (P4, P8, P10, P9) and low (P3, P6, P7, P11, P12, P15) VL groups appreciated the presence of questions explicitly indicating what insights might be extracted from the corresponding charts. As illustrated by P4 (high VL): *"I definitely liked how visualisations are linked to the question that best matches it. And that's a good first-order approximation"*; and P12 (low VL): *"The questions made me think more comprehensively. I felt I could actually use all elements to come up with a really good answer."* Yet, only two participants, both in the high VL group (P2 and P5) thought the questions would be principally the same as adding short statements as titles of the charts according to P2: *"It takes me a longer time to read questions compared to statements."*

In fact, some teachers in both groups (P4 and P2 in the high VL and P3 in low VL groups) found that questions could potentially restrict their interpretation. For instance, P4 explained a potential problem with one of the questions as follows: *"'Groups that are discussing enough?' I like the chart as it leaves that very open and I have no issue with that. But I think that any question that puts a rank could be problematic."* The same teacher also explained how single questions could drive teachers to not make comparisons using more than one chart as follows: *"I found that the best insights I could get were from comparing multiple diagrams against each other. Having focused on one might take a tutor away from doing that sort of comparison."* P2 and P3 also added that for some terms used in some questions, such as *"engaging"* and *"discuss enough"*, teachers may need more information about how the information in the chart was captured and what thresholds are used to assess the acceptable level of "engagement" or "discussion".

In sum, the majority of teachers valued the questions as titles, because they were able to know the intention behind each chart. However, some may prefer titles that explain how the data is related to particular educational constructs.

4.2.2 Emphasised elements. The high VL teachers generally appreciated the emphasised elements in the bar charts (P2, P8, P9). As P2 highlighted: *"This bar [chart] is definitely able to help me because I can directly tell which one [group] has the most engagement. It brings my attention to it, such as, and I spend too much time with group 4."* In contrast, only a few low VL participants (P3 and P7) mentioned that the emphasis in the bar charts was helpful. This may be explained by the fact that only a few of the low VL teachers could understand why some bars in such charts were highlighted using colour. P12 (low VL) explained this as follows: *"I don't know if I would actually say that the colour means anything to me because*

Category	Feature	V1	V2
1-Non-AOI	meansaccadedistance	W=113, p=0, Mdn_lowVL=269.93, Mdn_highVL=327.56, ES=0.6 †††	W=91, p=0.13, Mdn_lowVL=297.1, Mdn_high=315.63, ES=0.32 ††
	meansaccadespeed	W=104, p=0.02, Mdn_lowVL=3.43, Mdn_highVL=3.85, ES=0.49 ††	W=79, p=0.45, Mdn_lowVL=3.74, Mdn_high=4.04, ES=0.17 †
2-Questions as titles	QS4_totaltimespent	W=80.5, p=0.39, Mdn_lowVL=3090, Mdn_highVL=3984, ES=0.19 †	W=104, p=0.02, Mdn_lowVL=958, Mdn_high=2980, ES=0.49 ††
	QS2_timetofirstfixation	W=59, p=0.64, Mdn_lowVL=-1, Mdn_highVL=-1, ES=0.11 †	W=30, p=0.03, Mdn_lowVL=14140, Mdn_high=84, ES=0.47 ††
	Prop_QS_to_other_charts	W=52, p=0.41, Mdn_lowVL=0.06, Mdn_highVL=0.07, ES=0.18 †	W=23, p=0.01, Mdn_lowVL=0.05, Mdn_high=0.08, ES=0.55 †††
3-Timeline	T_totaltimespent	W=104, p=0.02, Mdn_lowVL=2666, Mdn_highVL=5198, ES=0.49 ††	W=107, p=0.01, Mdn_lowVL=2424.5, Mdn_high=5697, ES=0.53 †††
	Count_Ch2_T	W=36, p=0.07, Mdn_lowVL=3, Mdn_highVL=8, ES=0.39 ††	W=31.5, p=0.04, Mdn_lowVL=3, Mdn_high=6, ES=0.44 ††
	Count_Ch4_T	W=61.5, p=0.8, Mdn_lowVL=2, Mdn_highVL=3, ES=0.06 †	W=21.5, p=0.01, Mdn_lowVL=1.5, Mdn_high=5, ES=0.58 †††
4-Textual summaries	Ch4_proportionnum	W=59, p=0.7, Mdn_lowVL=0.29, Mdn_highVL=0.26, ES=0.09 †	W=33, p=0.04, Mdn_lowVL=0.35, Mdn_high=0.26, ES=0.42 ††
	Count_Ch1_Ch4	W=29, p=0.02, Mdn_lowVL=8.5, Mdn_highVL=18, ES=0.48 ††	W=65.5, p=1, Mdn_lowVL=26.5, Mdn_high=16, ES=0.01 †
	Prop_Ch1_Ch4	W=24, p=0.01, Mdn_lowVL=0.01, Mdn_highVL=0.03, ES=0.54 †††	W=66, p=1, Mdn_lowVL=0.03, Mdn_high=0.03, ES=0 †

Effect sizes: †: '0.10 - < 0.3' (small effect), ††: '0.30 - < 0.5' (moderate effect) and †††: '>= 0.5' (large effect). [a]
In version 2, sociograms (Ch4) were enhanced with the text shown in version 1 (Ch1). Thus, in version 2, Ch1 refers to {DSc}. [*] [*]

Table 1: Selection of key differences between high and low VL groups per eye-movement feature and dashboard version (1 and 2). Features are grouped in the same way they are described in the results section. Statistically significant results, where p < 0.05, are highlighted in blue, where high opacity indicates a large effect size.

it's actually the longest bar anyhow. So I would just look at the length of the bar." The emphasis in sociograms (Ch4) was perceived to be the most useful by both low VL (P3, P11, P15) and high VL (P2, P5, P10) participants. Teachers stressed the ease of identifying inactive students via sociograms, as P15 (low VL) explained:*"In terms of [students'] activeness I, as a teacher, can say that some students are inactive. It is easy."*. P2 (high VL) mentioned that *"[Ch4] especially allows me to differentiate between the teachers and active or inactive [students] ... it kindly gives me some hint about whether a student dominates a [conversation] or not."*

Two teachers, P7 (low VL) and P5 (high VL), mentioned that emphasising the current activity in the timeline chart was the most helpful element to remind them that the first priority was to interpret the data based on the context of the class script. P5 even questioned whether all the charts are always relevant to their teaching practices. She suggested providing a mechanism to indicate which visualisations are relevant for the current learning activity, as follows: *"The [timeline] indicates the activity which we're supposed to look at. It's a good indication. It would be good for us to also quickly identify the chart [to look at] because if we are having different types of activities I had to go through all the diagrams."*

These results suggest that while emphasised elements may support teachers regardless of their VL, those with low VL may still require a deeper understanding of the mechanisms used to highlight elements using colour, and teachers in general may also benefit from adapting the charts according to the characteristics of the learning activity.

4.2.3 Colour semantics and mapping. In general, low VL participants were more positive regarding the role of colour in the dashboard. In contrast, teachers in the high VL group were more reserved towards the role of colour. Participants P3, P6, P7, P11 and P12 in the low VL group and P10 in the high VL group expressed a **positive sentiment** towards the colour use and semantics, i.e., *"Even though I don't recall what a colour stands for, it motivates me to think"* (P6). Participants P6, P11 and P10 acknowledged the use of colours to focus their attention on notable elements in the charts and helped them to interpret the dashboards, as illustrated by P11: *"The colour is good, you're contrasting students who are not alike."* P6 (in the low VL group) explained that colour helped her to quickly see the differences in students' progress: *"[Colour] drives your focus there, because it's a different colour and you're thinking, why is it a different colour? Especially the red ones, for students who are not talking. That was very good. And also for understanding which amongst all the groups are finishing activities, because the orange and blue are much easier to understand."*

In contrast, participants in the high VL group (P2, P8 and P9) expressed more **neutral sentiment** regarding the colour mapping to emphasise elements, i.e., *"I need to think about the colour; it doesn't represent all I want; colour is not consistent)"* (P9). P2 indicated that they were unsure about the mapping behind the colours at the beginning, but were able to figure this out after reading textual summaries: *"If there's no legend, I forget what does orange mean. I need to read the text below to understand what is orange"*. Only two participants, P15 (low VL) and P4 (high VL) expressed a **negative sentiment** regarding the colour mapping, i.e., *"Even though I don't recall what a colour stands for, it rather discourages me."* P15 mentioned that they have been in situations, where they could not figure out what the colour means and it was rather discouraging. P4 stated that the information encoded in the colour is redundant and only leads to a cluttered view. Moreover, some high VL teachers were more sensitive to subtle inconsistencies in the colour design as it was explained by P9 as follows: *"In general people appreciate the colour to be consistent ... because in a bar chart, bars lower than the [orange one], they are less than an average [so those bars are blue] ... but for [google] documents, the [completed section] is coloured in blue. I can recall that [difference] when I use the [interface]."*

These results suggest that whilst colour mapping can be used to drive the attention of teachers, for teachers with low VL either training may be needed for them to make sense of the underlying meaning of colour, or colour needs to be used sparingly and consistently in the design of the dashboard.

4.2.4 Textual summaries. Some teachers in the low VL group (P3, P7, P15 and P12) reported that they found the textual summaries useful (e.g., *"I found the comments at the bottom to be a very useful tool."* - P3). Teachers in the high VL group (P2, P8, and P10) also acknowledged that textual summaries might be instrumental for the interpretation of charts (e.g., *"I think it's good that we have like a brief summary here"* - P8). Despite teachers appreciating textual summaries, both high (P2, P4 and P8) and low (P6 and P11) VL groups reported that they did not use them, because they found it easier to work with charts directly. They mentioned that textual summaries are rather redundant for their interpretation because such summaries provide only a partial story while the charts contain additional information that high VL teachers would like to decode. As P8 (high VL) explained: *"I think it's good that we have a brief summary here, but I still prefer to look at the graph because it tells you the amount of engagement between students."* Even though some teachers preferred to interpret charts directly, P10 (high VL) mentioned that teachers might use textual summaries to double-check their interpretation, especially in situations where they could not have time to fully inspect the charts because of their teaching duties: *"In the 2nd version, you combine the description with the chart. Now I can quickly look at the two charts at once. I found it easier to look for the information. Otherwise, I need to spend time looking at the diagram and counting them"*. Yet, two teachers in the low VL group (P3 and P7) were not able to make a connection between textual summaries and information in sociograms. Instead, they interpreted those charts separately, as stated by P3: *"The comments at the bottom say things about each group. And in particular, it says all students are active, but [a student] is the quietest. That's important information."* This suggests that while the high VL teachers may have been capable of interpreting complex graphs (such as the sociograms) by themselves, the low VL teachers could benefit from more explicit data storytelling guidance in the form of textual summaries and high VL teachers can still benefit from them under regular class time constraints.

5 DISCUSSION

5.1 Research questions

Regarding RQ1, results from the eye tracking analysis suggest that high VL teachers can show more explicit information searching behaviours by i) identifying and utilising related charts during their interpretations; ii) interpreting charts with more scrutiny against their corresponding explanatory chart titles; and iii) recognising emphasised elements which might be instrumental to aid interpretation. Teachers in the low VL group benefited more from changes in the design, for example, bringing textual summaries closer to the sociograms. This emphasises the importance of designing explainable data interfaces in education [15]. Our quantitative results suggest that adding explainable [24] or information guidance [42] features to LA dashboards may benefit teachers with low VL the most. However, our qualitative analysis (addressing RQ2) provided more details about other challenges that low VL teachers may still face.

Although both low and high VL participants appreciated the enhancement of the charts using data storytelling elements, teachers in the low VL group could not always retell the principles upon which emphasis was applied to some charts such as the meaning of some colour variations. Yet, the sentiment towards the helpfulness of colour was more positive for the low VL teachers compared to those in the high VL group. This suggests that for teachers with low VL literacy emphasising key data points with colour can support interpretation but the design should very explicitly explain how colour is used to convey meaning or some basic training may still need to be provided but focused on VL. This empirically confirms the potentially positive role of dashboards with data storytelling enhancements [10, 18], especially for supporting teachers with low VL. In contrast, teachers in the high VL group showed more complex dashboard inspection strategies both in the eye tracking analysis and in the ways they explained how often they used the charts by combining them to make sense of students' activity. This finding is supported by visualisation literacy literature which explains that VL skills can enable the effective exploration of data visualisations [27].

Moreover, incorporating text into dashboards may benefit teachers regardless of their VL. In our study, even though teachers in the high VL group did not use them as much, some teachers stressed that textual summaries were instrumental to double-check their interpretation, especially in situations of high orchestration workload [48]. However, even though using text as an alternative to communicating LA outputs to teachers has been suggested in previous research [18], more research is still needed to understand in which situations teachers would benefit from specific data storytelling and narrative techniques. For instance, Van Leeuwen and Rummel [49] found that teachers need more time to interpret groups' progress in dashboards where charts are displayed along with the descriptive text. Our study adds another dimension to these previous studies by suggesting that the effectiveness of providing text in dashboards may depend on teachers' VL. For example, while some teachers may benefit from explicit textual explanations, these may also require teachers to have the skills to effectively integrate evidence from two different modalities: text and visual charts. Here is where the effective design of the LA interface may play a key role to scaffold such integration, as we found for the case of the sociograms enhanced with textual descriptions in version 2 of the dashboard explored in our study.

5.2 Implications for research

This work emphasises visualisation literacy as a key notion to investigating LA dashboard use by teachers. Even though previous works have emphasised the role of *data literacy* for successful use of visual learning analytics [29, 33], the fact that teachers primarily interact with the visual interface of a LA system can get inadvertently de-emphasised. Little attention has been paid to the notion of visualisation literacy in the LA community. To the best of our knowledge, this paper is one of the first attempts (besides [13]) to explicitly focus on the notion of visualisation literacy, which can be key in the effective use of several teacher-facing dashboards. Investigating visualisation literacy in connection to data storytelling interfaces is important because it can enable the provision of targeted support for user groups of varying abilities to utilise information from visualisations. Although increasing general levels of VL among teachers might be an unattainable goal [50], providing intelligent interfaces that account for differences in VL can be an attainable aim [12]. Yet, future works can further investigate to what extent teachers need to develop visualisation literacy skills in order to effectively interact with a range of LA end-user interfaces and dashboards.

5.3 Implications for design practice

Some implications for LA design emerged from our study. While our results suggest that enhancing LA visualisations with text may support teachers regardless of their VL levels, this feature should be added in alignment with the context of use for avoiding text to become a distraction [49]. Moreover, as it has been suggested by some authors, teachers could be given the option to disable such a feature according to their personal preferences [1] or specific data points that are described in the text can be highlighted for the text to be contextualised [3]. Nevertheless, more research is required to understand to what extent LA interfaces enhanced with explicit and implicit guidance can support teachers with different VL skills. Since teachers commonly lack formal training in data analysis, a potential way to indirectly improve their understanding of the LA innovation, which can be further explored, is to involve teachers in the design of the LA interfaces or the training materials aimed at improving teachers' visualisation literacy competencies or recommended data practices [7]. This could simplify the design and ease the adoption of LA.

5.4 Limitations and Future Work

Our study has some limitations. First, this study employs the full VLAT test [27] which includes several visualisation techniques that are not included in the dashboards shown to participants. Yet, the rationale to use it *as-is* was that this is already a validated instrument we did not want to change. Future research could focus on measuring visualisation literacy for specific visualisation techniques used in a particular dashboard. Second, the two dashboard versions utilised in the study are part of an ongoing human-centred LA project. These versions should be considered as prototypes that still require improvements. Yet, working with an authentic scenario also brings the advantage of contributing to the design of a LA tool that can be enhanced and deployed to support teachers' monitoring tasks in online classes. Future work could conduct more experimental studies which can test a wider range of visualisation designs.

Third, our analysis is only exploratory and results should be confirmed by future studies. Finally, our study was deployed in the context of a specific university, academic discipline, and a single course. More work would need to be done to understand how the findings are applicable to other contexts.

6 CONCLUSION

Results from our study indicate that a) high VL teachers adopted exploratory strategies when using two dashboard versions; b) low literacy teachers find it problematic to immediately figure out how to use visual enhancements from various visualisations; yet, they benefited the most from more explicit data storytelling guidance. Hence, to make a difference, dashboard designers should very explicitly couple such visual enhancements for low VL teachers to benefit from them. However, this raises the question of how to personalise teacher-facing interfaces to provide equal support for both high and low VL teachers. This paper should be seen as the first of further work needed to understand the potential interplay between visualisation literacy and design enhancements in LA dashboards.

ACKNOWLEDGMENTS

This research was at least in part funded by the Australian Research Council (DP210100060) and Jacobs Foundation (Research Fellowship).

REFERENCES

[1] J. Ahn, H. Nguyen, et al. 2021. From visible to understandable: Designing for teacher agency in education data visualizations. *Contemporary Issues in Technology and Teacher Education* 21, 1 (2021), 155–186.

[2] H. Bao, Y. Li, et al. 2021. The effects of a learning analytics dashboard on teachers' diagnosis and intervention in computer-supported collaborative learning. *Technology, Pedagogy and Education* 30, 2 (2021), 287–303. https://doi.org/10.1080/1475939x.2021.1902383

[3] O. Barral, S. LallÉ, et al. 2021. Effect of Adaptive Guidance and Visualization Literacy on Gaze Attentive Behaviors and Sequential Patterns on Magazine-Style Narrative Visualizations. *ACM Transactions on Interactive Intelligent Systems* 11, 3-4 (Dec. 2021), 1–46. https://doi.org/10.1145/3447992

[4] T. Bartindale, P. Chen, et al. 2021. ZoomSense: A Scalable Infrastructure for Augmenting Zoom. In *Proceedings of the 29th ACM International Conference on Multimedia* (Virtual Event, China) *(MM '21)*. Association for Computing Machinery, New York, NY, USA, 3771–3774. https://doi.org/10.1145/3474085.3478332

[5] R. Bender and S. Lange. 2001. Adjusting for multiple testing—when and how? *Journal of clinical epidemiology* 54, 4 (2001), 343–349.

[6] V. Braun and V. Clarke. 2012. Thematic Analysis. In *APA Handbook of Research Methods in Psychology, Vol. 2.*, H. Cooper, P. Camic, D. Long, A. Panter, D. Rindskopf, and K. Sher (Eds.). American Psychological Association, Washington, USA, 57–71. https://doi.org/10.1037/13620-004

[7] S. Buckingham Shum, R. Ferguson, et al. 2019. Human-Centred Learning Analytics. *Journal of Learning Analytics* 6, 2 (Jul. 2019), 1–9. https://doi.org/10.18608/jla.2019.62.1

[8] F. Campos, J. Ahn, et al. 2021. Making Sense of Sensemaking: Understanding How K–12 Teachers and Coaches React to Visual Analytics. *Journal of Learning Analytics* (July 2021), 1–21. https://doi.org/10.18608/jla.2021.7113

[9] D. Ceneda, T. Gschwandtner, et al. 2017. Characterizing Guidance in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan. 2017), 111–120. https://doi.org/10.1109/TVCG.2016.2598468

[10] Q. Chen, Z. Li, et al. 2019. Designing narrative slideshows for learning analytics. In *2019 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 237–246. https://doi.org/10.1109/pacificvis.2019.00036

[11] A. Choudhry, M. Sharma, et al. 2021. Once Upon A Time In Visualization: Understanding the Use of Textual Narratives for Causality. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (feb 2021), 1332–1342. https://doi.org/10.1109/tvcg.2020.3030358

[12] Y. Dimitriadis, R. Martínez-Maldonado, et al. 2021. Human-Centered Design Principles for Actionable Learning Analytics. In *Research on E-Learning and ICT in Education*, Thrasyvoulos Tsiatsos, Stavros Demetriadis, Anastasios Mikropoulos, and Vasileios Dagdilelis (Eds.). Springer International Publishing, Cham, 277–296. https://doi.org/10.1007/978-3-030-64363-8_15

[13] D. Donohoe and E. Costello. 2020. Data Visualisation Literacy in Higher Education: An Exploratory Study of Understanding of a Learning Dashboard Tool. *International Journal of Emerging Technologies in Learning (iJET)* 15, 17 (Sept. 2020), 115. https://doi.org/10.3991/ijet.v15i17.15041

[14] R. A. Dourado, R. L. Rodrigues, et al. 2021. A teacher-facing learning analytics dashboard for process-oriented feedback in online learning. In *LAK21: 11th International Learning Analytics and Knowledge Conference*. 482–489. https://doi.org/10.1145/3448139.3448187

[15] V. Echeverria, R. Martinez-Maldonado, et al. 2018. Exploratory versus Explanatory Visual Learning Analytics: Driving Teachers' Attention through Educational Data Storytelling. *Journal of Learning Analytics* 5, 3 (Nov. 2018). https://doi.org/10.18608/jla.2018.53.6

[16] V. Echeverria, R. Martinez-Maldonado, et al. 2018. Driving Data Storytelling from Learning Design. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. 131–140. https://doi.org/10.1145/3170358.3170380

[17] R. J. Feise. 2002. Do multiple outcome measures require p-value adjustment? *BMC medical research methodology* 2, 1 (2002), 1–4.

[18] G. M. Fernandez Nieto, K. Kitto, et al. 2022. Beyond the Learning Analytics Dashboard: Alternative Ways to Communicate Student Data Insights Combining Visualisation, Narrative and Storytelling. In *LAK22: 12th International Learning Analytics and Knowledge Conference*. ACM, Online USA, 219–229. https://doi.org/10.1145/3506860.3506895

[19] E. E. Firat, A. Joshi, et al. 2022. Interactive Visualization Literacy: The State-of-the-Art. *Information Visualization* 21, 3 (July 2022), 285–310. https://doi.org/10.1177/14738716221081831

[20] M. J. Gomez, J. A. Ruipérez-Valiente, et al. 2021. Applying learning analytics to detect sequences of actions and common errors in a geometry game. *Sensors* 21, 4 (2021), 1025. https://doi.org/10.3390/s21041025

[21] C. Herodotou, B. Rienties, et al. 2019. A Large-Scale Implementation of Predictive Learning Analytics in Higher Education: The Teachers' Role and Perspective. *Educational Technology Research and Development* 67, 5 (Oct. 2019), 1273–1306. https://doi.org/10.1007/s11423-019-09685-0

[22] R. Kaliisa, A. I. Mørch, et al. 2022. 'My Point of Departure for Analytics Is Extreme Skepticism': Implications Derived from An Investigation of University Teachers' Learning Analytics Perspectives and Design Practices. *Technology, Knowledge and Learning* 27, 2 (June 2022), 505–527. https://doi.org/10.1007/s10758-020-09488-w

[23] S. Kardan, S. Lallé, et al. 2021. EMDAT: Eye Movement Data Analysis Toolkit (1.x). Zenodo. https://doi.org/10.5281/zenodo.4699774

[24] H. Khosravi, S. B. Shum, et al. 2022. Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence* 3 (2022), 100074. https://doi.org/10.1016/j.caeai.2022.100074

[25] C. A. Knoop-van Campen, A. Wise, et al. 2021. The equalizing effect of teacher dashboards on feedback in K-12 classrooms. *Interactive Learning Environments* (2021), 1–17. https://doi.org/10.1080/10494820.2021.1931346

[26] S. Lalle, D. Toker, et al. 2021. Gaze-Driven Adaptive Interventions for Magazine-Style Narrative Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 27, 6 (June 2021), 2941–2952. https://doi.org/10.1109/TVCG.2019.2958540

[27] S. Lee, S.-H. Kim, et al. 2017. VLAT: Development of a Visualization Literacy Assessment Test. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan. 2017), 551–560. https://doi.org/10.1109/TVCG.2016.2598920

[28] M. Liu, S. Han, et al. 2021. The Current Landscape of Research and Practice on Visualizations and Dashboards for Learning Analytics. In *Visualizations and Dashboards for Learning Analytics*. Springer, 23–46. https://doi.org/10.1007/978-3-030-81222-5_2

[29] E. B. Mandinach and L. M. Abrams. [n. d.]. Data Literacy and Learning Analytics. In *The Handbook of Learning Analytics* (2 ed.), Charles Lang, George Siemens, Alyssa Friend Wise, Dragan GaÅ¡eviÄ‡, and Agathe Merceron (Eds.). SoLAR, 196–204. https://www.solaresearch.org/publications/hla-22/hla22-chapter19/ Section: 19.

[30] R. Martinez-Maldonado, V. Echeverria, et al. 2020. From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–15. https://doi.org/10.1145/3313831.3376148

[31] N. McDonald, S. Schoenebeck, et al. 2019. Reliability and Inter-Rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23. https://doi.org/10.1145/3359174

[32] I. Molenaar and C. A. N. Knoop-van Campen. 2019. How Teachers Make Dashboard Information Actionable. *IEEE Transactions on Learning Technologies* 12, 3 (July 2019), 347–355. https://doi.org/10.1109/TLT.2018.2851585

[33] I. G. Ndukwe and B. K. Daniel. 2020. Teaching Analytics, Value and Tools for Teacher Data Literacy: A Systematic and Tripartite Approach. *International Journal of Educational Technology in Higher Education* 17, 1 (2020), 1–31. https://doi.org/10.1186/s41239-020-00201-6

[34] A. Ojo and B. Heravi. 2018. Patterns in Award Winning Data Storytelling: Story Types, Enabling Tools and Competences. *Digital Journalism* 6, 6 (July 2018), 693–718. https://doi.org/10.1080/21670811.2017.1403291

[35] A. Olsen. 2012. The Tobii I-VT fixation filter. *Tobii Technology* 21 (2012), 4–19.

[36] S. Pozdniakov, R. Martinez-Maldonado, et al. 2021. Question-driven learning analytics: Designing a teacher dashboard for online breakout rooms. In *2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 176–178.

[37] S. Pozdniakov, R. Martinez-Maldonado, et al. 2022. The Question-driven Dashboard: How Can We Design Analytics Interfaces Aligned to Teachers' Inquiry?. In *LAK22: 12th International Learning Analytics and Knowledge Conference*. 175–185. https://doi.org/10.1145/3506860.3506885

[38] S. Pozdniakov, R. Martinez-Maldonado, et al. 2022. The Question-Driven Dashboard: How Can We Design Analytics Interfaces Aligned to Teachers' Inquiry?. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (Online, USA) *(LAK22)*. Association for Computing Machinery, New York, NY, USA, 175–185. https://doi.org/10.1145/3506860.3506885

[39] M. J. Rodríguez-Triana, L. P. Prieto, et al. 2018. The Teacher in the Loop: Customizing Multimodal Learning Analytics for Blended Learning. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK '18*. ACM Press, Sydney, New South Wales, Australia, 417–426. https://doi.org/10.1145/3170358.3170364

[40] S. Z. Salas-Pilco, K. Xiao, et al. 2022. Artificial Intelligence and Learning Analytics in Teacher Education: A Systematic Review. *Education Sciences* 12, 8 (2022), 569. https://doi.org/10.3390/educsci12080569

[41] J. Saldaña. 2015. *The Coding Manual for Qualitative Researchers*. SAGE. https://doi.org/10.46743/2160-3715/2009.2856

[42] H.-J. Schulz, M. Streit, et al. 2013. Towards a characterization of guidance in visualization. In *Poster at IEEE Conference on Information Visualization (InfoVis)*.

[43] S. Sergis and D. G. Sampson. 2017. Teaching and Learning Analytics to Support Teacher Inquiry: A Systematic Literature Review. *Learning analytics: Fundaments, applications, and trends* (2017), 25–63. https://doi.org/10.1007/978-3-319-52977-6

[44] E. B. d. Sousa, B. Alexandre, et al. 2021. Applications of learning analytics in high schools: a Systematic Literature review. *Frontiers in Artificial Intelligence* 4 (2021), 737891. https://doi.org/10.3389/frai.2021.737891

[45] A. Stefanile. 2020. The Transition from Classroom to Zoom and How It Has Changed Education. *Journal of Social Science Research* 16 (July 2020), 33–40. https://doi.org/10.24297/jssr.v16i.8789

[46] Y.-S. Tsai and D. Gasevic. 2017. Learning Analytics in Higher Education — Challenges and Policies: A Review of Eight Learning Analytics Policies. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, Vancouver British Columbia Canada, 233–242. https://doi.org/10.1145/3027385.3027400

[47] A. van Leeuwen, C. A. Knoop-van Campen, et al. 2021. How teacher characteristics relate to how teachers use dashboards: Results from two case studies in K-12. *Journal of Learning Analytics* 8, 2 (2021), 6–21. https://doi.org/10.18608/jla.2021.7325

[48] A. van Leeuwen and N. Rummel. 2019. Orchestration Tools to Support the Teacher during Student Collaboration: A Review. *Unterrichtswissenschaft* 47, 2 (April 2019), 143–158. https://doi.org/10.1007/s42010-019-00052-9

[49] A. Van Leeuwen and N. Rummel. 2020. Comparing teachers' use of mirroring and advising dashboards. In *Proceedings of the tenth international conference on learning analytics & knowledge*. 26–34. https://doi.org/10.1145/3375462.3375471

[50] R. Voithofer and A. M. Golan. 2018. Data Sources for Educators: Mining Meaningful Data for Course and Program Decision Making. In *Responsible Analytics and Data Mining in Education*. Routledge, 83–100. https://doi.org/10.4324/9780203728703-7

[51] K. Williamson and R. Kizilcec. 2022. A review of learning analytics dashboard research in higher education: Implications for justice, equity, diversity, and inclusion. In *LAK22: 12th International Learning Analytics and Knowledge Conference*. 260–270. https://doi.org/10.1145/3506860.3506900

[52] A. F. Wise and Y. Jung. 2019. Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics* 6, 2 (2019), 53–69. https://doi.org/10.18608/jla.2019.62.4

[53] A. Wolff, J. Moore, et al. 2016. Data literacy for learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge*. 500–501. https://doi.org/10.1145/2883851.2883864