

Exploratory versus Explanatory Visual Learning Analytics: Driving Teachers' Attention through Educational Data Storytelling

Vanessa Echeverria¹, Roberto Martinez-Maldonado², Simon Buckingham Shum³, Katherine Chiluiza⁴, Roger Granda⁵, Cristina Conati⁶

ABSTRACT

From a human-centred computing perspective, supporting the interpretation of educational dashboards and visualizations by the people intended to use them exposes critical design challenges that may often be trivialized. Empirical evidence already shows that “usable” visualizations are not necessarily effective from an educational perspective. Since an educator’s interpretation of visualized data is essentially the construction of a narrative about student progress, in this paper, we propose the concept of “Educational Data Storytelling” as an approach for explaining student data by aligning educational visualizations with the intended learning design. We present a pilot study that explores the effectiveness of these data storytelling elements based on educator responses to prototypes by analyzing the kinds of stories they articulate, their eye-tracking behaviour, and their preferences after inspecting a series of student data visualizations. The dual purpose is to understand the contribution of each visual element for data storytelling, as well as the effectiveness of the enhancements when combined.¹

Notes for Practice

- The learning analytics field, to effectively support teaching and learning visually, needs to move on from visual analytics that invite users to explore the data, to visualizations that explain insights.
- In this paper, we propose the concept of “Educational Data Storytelling” as an approach to designing visual learning analytics interfaces that explain student data by aligning educational visualizations with the learning design intended by the teacher.
- We see the potential of learning-design-driven data storytelling elements to support sensemaking by guiding students and teachers to “one learning story per visualization,” given that learning is a complex task.

Keywords

Human-centered design, storytelling, eye tracking

Submitted: 10.06.2018 — **Accepted:** 27.08.2018 — **Published:** 11.12.2018

Corresponding author *Corresponding author* ¹Email: vanessa.i.echeverriabarzola@student.uts.edu.au Address: Connected Intelligence Centre, University of Technology Sydney, 2 Blackfriars St, Building 22, PO Box 123, Brodway, Ultimo, NSW 2007, Australia, ORCID ID 0000-0002-2022-9588

²Email: roberto.martinez-maldonado@uts.edu.au Address: Connected Intelligence Centre, University of Technology Sydney, 2 Blackfriars St, Building 22, PO Box 123, Brodway, Ultimo, NSW 2007, Australia

³Email: simon.buckinghamshum@uts.edu.au Address: Connected Intelligence Centre, University of Technology Sydney, 2 Blackfriars St, Building 22, PO Box 123, Brodway, Ultimo, NSW 2007, Australia

⁴Email: kchilui@espol.edu.ec Address: Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador

⁵Email: roger.granda@cti.espol.edu.ec Address: Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador

⁶Email: conati@cs.ubc.ca Address: Department of Computer Science, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

¹ An earlier, shorter version of this paper (Echeverria et al., 2018) is the foundation for this article, which has been significantly extended in light of feedback and insights from LAK 2018.

1. Introduction and Motivation

Data science is now impacting the education sector, with a growing number of commercial products and research prototypes providing learning dashboards and visualizations (which we will also refer to as “visual learning analytics,” Duval, 2011). These visual representations of student data have gained considerable attention in the field of learning analytics (see recent reviews by Bodily et al., 2018; Bodily & Verbert, 2017; Schwendimann et al., 2017). Borrowing from data and web analytics in other domains, learning analytics dashboards aggregate indicators, from one or more sources, about student activity and/or learning, using one or more visualizations (Schwendimann et al., 2017; Verbert, Duval, Klerkx, Govaerts, & Santos, 2013). Teacher-facing dashboards are commonly intended to help educators gain a better understanding of their whole course or specific tasks, reflect on their teaching strategies, and identify students who require specific attention (Molenaar & Knoop-van Campen, 2017; Verbert et al., 2013). Similarly, student-facing dashboards are intended to help students reflect on aspects of their learning behaviour and potentially assist them in, for example, managing time effectively, accessing key learning resources, or gaining a richer picture of their progress (Bodily & Verbert, 2017; Reimers & Neovesky, 2015).

It may be useful to step back and remember that the most familiar concept of a dashboard that people have in their minds is that of the car dashboard. These interfaces are a minimalistic window to the status of a complex system: the car. They thus include critical data about speed, fuel level, lights, and alarms when something is wrong with the engine. Sometimes they may also include less critical but useful data such as the time or the outside air temperature. In short, car dashboards are typically *uncluttered* interfaces that provide just the information that is critical for the context of use and are aligned with people’s tasks (e.g., driving safely, on-time, or economically; Kim, Dey, Lee, & Forlizzi, 2011). Notably, there are three features (Stevens, Quimby, Board, Kersloot, & Burns, 2002) of these dashboards that are relevant for rethinking the design and use of dashboards in the area of learning analytics:

- **Explanatory purpose:** Car dashboards are *explanatory* interfaces rather than *exploratory*. This means they are designed to prioritize the provision of the information that is essential for the driver — historical data or data not related to the current ride is not shown at first glance.
- **Task alignment:** The usefulness and priority of each feature depends on the task that the driver performs (e.g., the clock or the navigator may be relevant, but only for certain situations).
- **Contextual visibility:** Some critical information is only displayed when relevant/critical actions must be taken (e.g., low-level fuel alarm prompts the driver to refuel), remaining completely invisible at other times.

These basic design features may seem obvious. Unfortunately, the design of most learning analytics dashboards does not necessarily align with these simple principles.

Although the dashboard concept is intuitively attractive, its effectiveness is still unclear when it comes to the very complex set of skills that are part of learning (Corrin & de Barba, 2015; Jivet, Scheffel, Drachsler, & Specht, 2017; Teasley, 2017). Studies report, for instance, that the value of *teacher* dashboards may depend on the degree to which they have been involved in co-designing them (Holstein, McLaren, & Aleven, 2017), and their data literacy (Feng, Krumm, Bowers, & Podkul, 2016). The methodologies now emerging, for learning analytics researchers to co-design analytics with teachers, draw explicit attention to the critical importance of upskilling teachers’ ability to understand what data visualizations mean (Krumm, Means, & Bienkowski, 2018). The value of *student* dashboards could be a function of a complex mix of learner characteristics (Corrin & de Barba, 2015; Teasley, 2017) or their alignment with students’ particular needs (Reimers & Neovesky, 2015). Additional barriers to adoption may include the potential misalignment between the visual analytics and the intended learning/instruction goals set for the learners’ tasks; and the orchestration challenges and time constraints teachers and learners may have to interact with the interfaces (Rodríguez-Triana, Martínez-Monés, Asensio-Pérez, & Dimitriadis, 2015).

This paper proposes an approach to increase the *explanatory* effectiveness of the visualizations contained in learning analytics dashboards. Explanatory visualizations are those whose main goal is the presentation and communication of insights (Iliinsky & Steele, 2011). By contrast, *exploratory* visualizations are commonly targeted at experts in data analysis in search of insights from unfamiliar datasets. (We note, however, the emergence of a hybrid: simpler interactive visualizations designed for untrained audiences, often used in data journalism and media websites for more curious readers to filter the data. This kind of constrained interactivity might augment the work reported in this paper.) While educational researchers and developers are highly motivated, and suitably skilled, to invest time exploring visual analytics, there is a different value proposition to navigate for either students or teachers. Their time is limited, and they typically lack the skills to analyze data, or use more advanced visualizations. Explanatory visualizations may be more helpful, designed carefully to convey specific kinds of insights. Following this argument, the challenge for learning analytics researchers and developers is to discover and communicate insights rather than leaving students and teachers to play the role of data analysts, at the risk of gaining no insight.

The approach proposed in this paper focuses on crafting educational data visualizations that are explanatory by incorporating *data storytelling* elements. The foundation underpinning the proposed approach includes not only sound ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial- 74 NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Information Visualization (InfoVis) and *Data Storytelling* principles, but connects them to the learning context defined by the teacher's intentions, which are commonly explicitly or implicitly stated in their *learning designs*. We refer to this emerging concept of learning design driven by data storytelling as "Educational Data Storytelling." We take a first step toward investigating the feasibility of the approach with an exploratory study analyzing teacher responses to handcrafted prototypes based on the principles. While we hope that this approach will be valuable for designers handcrafting visualizations (the usual context for applying data storytelling), we foresee that they could be *automatically* implemented in dashboards to dynamically enrich visualizations of data that are unknowable in advance. In brief, the contribution of this paper is a Learning Design driven Educational Data Storytelling approach, providing the rationale for combining data storytelling foundations and the teacher's learning design, for developing explanatory learning analytics interfaces, with quantitative and qualitative validation of the concept.

The remainder of the paper is structured as follows. Section 2 discusses the current visual learning analytics challenges reported in the literature and presents the foundations of data storytelling. Section 3 describes the proposed approach and compares it with a typical data-driven dashboard design approach. Section 4 describes an exploratory study with teachers, to illustrate the application of the approach. Section 5 presents analysis and results from the study. The paper concludes with Section 6, with reflections on the feasibility of this work and further directions of data storytelling for learning analytics.

2. Background

2.1. Related Work

Although the concept of "learning dashboards" and other visual learning analytics have been receiving significant coverage in recent years (Bodily & Verbert, 2017; Schwendimann et al., 2017), there have also been numerous reports pointing to the limitations and possible pitfalls of products (Corrin & de Barba, 2015; Jivet et al., 2017; Reimers & Neovesky, 2015; Teasley, 2017). The design of effective dashboards in educational contexts is complex. First, a recent review (Bodily & Verbert, 2017) pointed to the absence of design choice justifications reported in several research outputs presenting educational dashboards usage. A current review finds poor evidence of grounding in the learning sciences (Jivet, Scheffel, Specht, & Drachsler, 2018). Another survey (Reimers & Neovesky, 2015) reports that both commercial and scientific dashboards commonly feature poor interface design and lack of usability testing. Furthermore, the choice of which data is to be visualized does not commonly correspond with what students and educators really want or need, in part, because they are not regularly consulted as part of the design process (Holstein et al., 2017). Marzouk et al. (2016) similarly critique dashboard design as poorly grounded in the learning sciences and demonstrate how designs might be better grounded in a theory of learner self-regulation.

In parallel to the visual and design factors noted above, researchers and designers can easily overlook the learning context and the audience for whom these visualizations have been created (Schwendimann et al., 2017). This may cause a disparity between users' and designers' perspectives. Sometimes, designers and researchers want to communicate multiple insights or dimensions of data about student experience. The conventional approaches adopted by researchers and designers can lead to the design of overly complex visualizations that are often hard to interpret (Duval, 2011) especially "at a glance." Moreover, teachers and students are encouraged to interpret these visualizations in a limited time and, even if the data can be interpreted correctly, they may fail to understand the "call to action," and fail to adapt their behaviour (Greller & Drachsler, 2012). Aligning learning design with the learning analytics tools has been identified as a key issue within the learning analytics community, in order to ensure that automated feedback is integrated coherently into the student activity (Bakharia et al., 2016; Knight, Shibani, & Buckingham Shum, 2018; Lockyer, Heathcote, & Dawson, 2013; Shibani, Knight, Buckingham Shum, & Ryan, 2017).

Our approach seeks a way to address some of the problems highlighted above by developing a learning design-driven approach to bring data storytelling and general visual design principles into the design of visual learning analytics. As will be described in the next section, data storytelling involves designing visualizations that communicate insights effectively. To the best of our knowledge, no prior dashboards proposed in the learning analytics literature use such an approach (including both data storytelling and explanatory visualizations). The closest work to ours has been conducted by Klerkx, Verbert, and Duval (2017) who suggested general considerations regarding the design of visual and interaction properties of learning dashboards.

2.2. Foundations for Educational Data Storytelling

Data storytelling can be described as an "information compression" technique, to help an audience focus on what is most important to communicate via data visualization (Ryan, 2016). Data storytelling builds on classic InfoVis guidelines (Tufte & Schmiegel, 1985) and narrative storytelling foundations (e.g., plots, twists and calls to action; Lee, Riche, Isenberg, & Carpendale, 2015). In the following sections, we introduce a set of data storytelling principles distilled from the literature, and a set of data storytelling visual elements that can be added to the design of visualizations and dashboards.

2.2.1. Data Storytelling Principles

Tufte and Schmieg (1985) argue that visualizations should incorporate “graphical excellence,” meaning that it has “to reveal data with clarity precision and efficiency.” In the same way, Ryan (2016) contends that the goal of using visuals is “to communicate a key message clearly and effectively,” emphasizing the context and meaning through visual enhancements (e.g., changes in size, colour, saturation). While visualizations are intended to either explore or explain insights, data storytelling is focused on the latter. The goal of data storytelling is to communicate insights through the combination of data, visuals, and narrative (Dykes, 2015). Thus, data storytelling is not applicable to exploratory visualization interfaces. Instead, data storytelling aims to explain what is happening in the data and why a feature is important. In this context, data storytelling has been widely used by journalists as a means to engage and inform through “interactive graphics and narratives” (Segel & Heer, 2010). Current reported uses of data storytelling techniques have focused on helping presenters tell more compelling stories through a clear (written or spoken) narrative supported by data visualizations (Knafllic, 2015); communicating scientific data (Ma, Liao, Frazier, Hauser, & Kostis, 2012); and teaching dynamic networks in an educational context (Bach et al., 2016).

Ryan (2016) and Knafllic (2015) have identified a set of “golden principles” that can be applied when crafting a story with data. Although the authors use different languages, they coincide in the following data storytelling principles that we can use to define further what educational data storytelling is:

- **Data storytelling is goal oriented.** The data visualization or dashboard needs to be aligned with a purpose or intention. Once aligned, it provides designers and researchers with clearer boundaries about what needs to be communicated and what does not. Although it may sound simple, many learning data visualizations seem designed to invite students or teachers to explore the data without a clear goal in mind, e.g., it is left entirely up to them which filters to apply, which graph type to use, or which region to focus on. This motivated our learning design-driven approach.
- **Data storytelling drives the audience’s focus of attention.** Visual and narrative elements should be used to create meaning in the visualization. Adding meaning to visualizations can be accomplished using specific elements (e.g., lines, weight, shapes, size, colours, contrast) to emphasize key aspects to direct visual attention. Narrative text is a salient visual element, providing summaries of visual features.
- **Data storytelling relies on choosing an appropriate visual.** Certain visualization techniques work better for certain purposes. For example, line charts can effectively show changes over time (Ryan, 2016). By contrast, Knafllic (2015) dedicates an entire chapter to justifying why pie charts should not be used.
- **Data storytelling relies on core InfoVis design principles.** Clutter in data visualizations adds visual complexity to the graph and makes it harder to find insights (Knafllic, 2015; Ryan, 2016). Tufte and Schmieg (1985) refer to this as maximizing data-ink ratio. *Decluttering* is achieved by removing elements (e.g., unnecessary headers, chart features, borders, grids) that do not add informative value to the graph. Other design aspects are important, for example, Knafllic (2015) proposes the addition of *titles* to drive home the intent of the visual, and *captions* to explain some relevant features of the data. General principles related to *alignment*, *use of colour*, *shape*, and *texture* are information design decisions that can have a substantial impact on sensemaking.

The next section illustrates some of the visual elements that can be added to or removed from data visualizations in order to translate these data storytelling high-level principles into tangible designs.

2.2.2. Data Storytelling Visual Elements

Knafllic (2015) describes how to improve graphs by adding or removing visual elements with the aim of supporting data storytelling. We explain each of these data storytelling elements with an illustrative educational example. Figure 1 (left) presents the typical output of a visualization tool (Google spreadsheets in this case) that we used to plot a fictional dataset regarding the number of visits per module in an online class during the past week, with the next day projection. Figure 1 (right) presents the same data visualization with data storytelling enhancements inspired by Knafllic’s work (with additional annotations A–E for expository purposes). A critical insight that can be emphasized is the increased number of visits for Module 4, which is more than the half of visits that other modules, considering that Module 4 is the most important and complex module of the semester. If we compare both graphs, we can see how this message is better understood in Figure 1 (right), without getting lost in exploring the visualization.

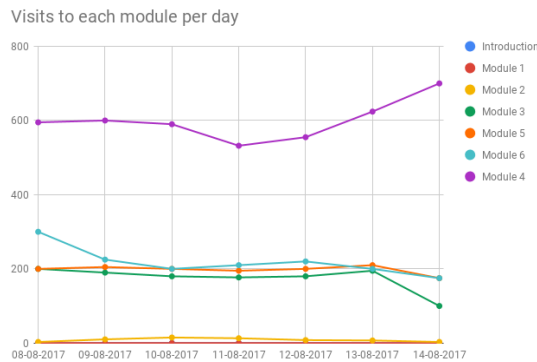
For instance, the following data storytelling design elements contribute to communicating one story in this example:

- A prescriptive title provides a succinct message intended to communicate explicitly the main message of the visualization. Since multiple stories may be told based on the same data, the title may serve to communicate directly the insight that should be sought in the current visualization.

- B. Highlighting specific data series (e.g., by making the line of one data series thicker) can help the audience to focus on the data that is relevant to the current story (in this case, the “Module 4” data series).
- C. Highlighting specific data points and removing irrelevant data points makes explicit which data support the claim in the title.
- D. Decluttering by removing grids, eliminating indirect legends, reducing the number of colours used, and removing non-relevant data point markers result in minimizing distractions. This also involves de-emphasizing data series that are not at the centre of the story (in this case, other data series such as “Module 5,” “Module 6,” “Module 3,” appear in light grey).
- E. Narrative text added directly to the visualization in the form of labels can help to gain a better understanding of or explain interesting changes in the data. In the example, the text next to the vertical line pointing at the day “11-Aug” in the horizontal axis (“Visits were more frequent due to the next quiz deadline...”) explains the increasing trend in the “Module 4” data series.
- F. A shaded area can be used to enclose those data points associated with the same insight. In the example, the shaded area divides the graph into actual data and data projections.

The critical question that follows is thus: given these generic data storytelling principles and design elements, how can these be explicitly contextualized to educational contexts for crafting explanatory visual learning analytics?

A typical **exploratory** chart



The same chart enhanced for **explanatory** purposes

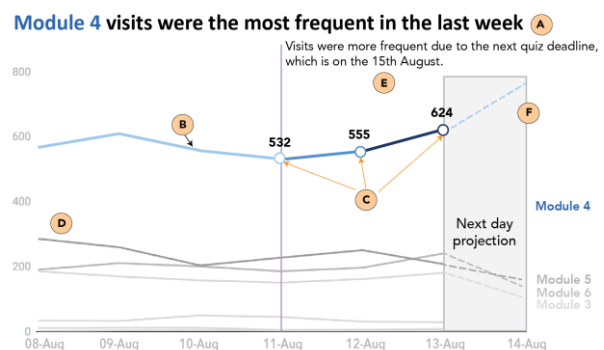


Figure 1. Left: A typical exploratory chart without data storytelling elements. Right: An explanatory chart of the same data with data storytelling elements, including: A) a prescriptive title; B) a thicker line to highlight particular data series; C) only critical data points; D) decluttering by de-emphasizing less relevant data; E) add key contextual data with labelled line; F) a shading area.

3. Approach

In this section, we introduce our Learning Design driven conceptual model. When learners or teachers face a dashboard or visualization, Verbert et al. (2013) proposed that the first stage in the “awareness-sensemaking” process is 1) visualizing and presenting the data to the user. Next, they argue that since “data in themselves are not very useful,” the second stage is where users (learners or teachers) are expected to 2) formulate questions and assess how data is useful or relevant for addressing those questions. The final two stages are concerned with the sensemaking process whereby users 3) respond to those questions (e.g., infer new insights) in order to 4) perform actions accordingly (e.g., educators modify the learning context, or students adapt their behaviours). This process then continues iteratively in what Verbert and colleagues (2013) call the “learning analytics process model.”

Figure 2 (left) summarizes this data-driven approach to visualize learning datasets. For clarity, in our diagram, we draw a distinction between (raw) learner data (presented as Data) and the analytics that are needed to present distilled views of these data for human consumption. The analytics include the *Processing* performed on the raw data (e.g., aggregating functions, machine learning, and prediction models) and the *Visual analytics* that render the data. In short, the process of formulating questions, generating insights from the analytics (or “data” according to Verbert et al., 2013) to respond to those questions, and acting is driven by the user. In this scenario, the designer of the learning analytics has the challenge of generating interfaces for a number of possible questions that learners or educators could have.

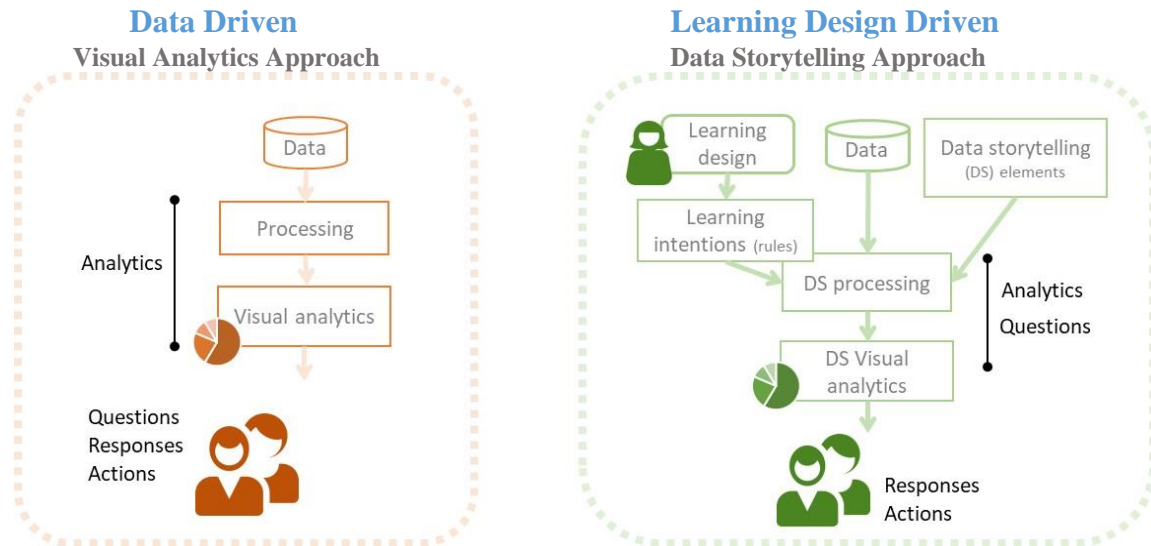


Figure 2. Conceptual model. Left: Conventional data-driven visual analytic approach. Right: Learning design driven, data storytelling approach to support sensemaking.

This is the conventional data visualization approach used in most learning analytics dashboards (see recent reviews by Bodily & Verbert [2017] and Klerkx et al. [2017]). We argue that this is one of the reasons why most visual learning analytics tools invite learners and teachers to *explore* the data rather than serving as tools that can *explain* the data or communicate insights. Teasley (2017) has referred to this exploratory approach as “one-size-fits-all” dashboard design. Notable exceptions are visualizations that incorporate mechanisms that allow user input such as customizable dashboards (e.g., see initial explorations in Roberts, Howell, & Seaman, 2017), and personalized visualizations that automatically adapt according to an individual’s traits (Lallé, Conati, & Carenini, 2016). These initial attempts are still in the research lab and have not yet been adopted by the learning analytics community. Teasley (2017) has also proposed that personalized visualizations may help address some of the already identified negative effects of dashboards and visual analytics that have been reported.

Our approach sits in between “one-size-fits-all” dashboards (simple to produce) and adaptive personalization (complex to produce and train). Figure 2 (right) depicts our proposed conceptual model. We propose to scaffold the formulation of user questions by enabling the dynamic modification of critical features of the visualizations. The goal is to translate the learning intentions of the educator into data storytelling elements that highlight particular aspects of the data. Figure 2 (right) shows three boxes aligned at the top as the foundations of our Learning Design-driven data storytelling approach. These are the teacher’s learning design, the learner’s data, and the data storytelling elements. Our model considers that a teacher’s pedagogical intentions, explicitly or implicitly specified in the learning design, can be translated into *rules* which can be read by a data processing system. Once the student data has been collected, these rules can be used to reveal interesting behaviours or indicators in student activity to which we wish to draw the educator’s or learner’s attention (e.g., when a student has low performance in an activity because he/she could not complete a correct step, or when a certain performance threshold is reached). Thus, the rules codify the teacher’s pedagogical intentions. Given an unknown number of possible interpretations users might draw from the visualization if they are simply left to explore, the rules apply a layer of filters and amplifiers to craft a specific data story.

Variations in data can, to some extent, be detected automatically with no domain knowledge. Progress in this area has recently been made in InfoVis to automatically annotate data points of interest in line charts, stacked graphs, and alluvial diagrams (Bryan, Ma, & Woodring, 2017); and in indoor maps (Metoyer, Zhi, Janczuk, & Scheirer, 2018). However, although specific properties of lines/curves (e.g., peaks and troughs, intersections, gradients), and time-series data (e.g., three consecutive low-performance episodes in a particular session) can be used by the system to highlight a multitude of changes in a graph, *from an educational point of view*, these may also be irrelevant.

The addition of *learning design* information for foregrounding elements of the data visualizations provides the basis for more powerful filters/amplifiers, since the educator can define what the *meaningful* features are likely to be. For example, perhaps the first two of the students’ three “low-performance episodes” are entirely expected by the educator and accounted for in the learning design, so the dashboard should only highlight the third event as a salient indicator. In the terms of Verbert and colleagues (2013), directing the user’s attention to identify interesting features at a glance should help them focus on questions that are relevant for learning/teaching since they have been identified and curated in the learning design. The goal is a more productive sensemaking process for users to respond to those questions and perform actions.

In summary, while the data storytelling elements play a critical role for emphasizing or de-emphasizing visual elements, the learning design (materialized as rules) could in principle provide the contextual information to decide *which* patterns or events should be foregrounded. This way, the proposed learning design-driven educational data storytelling approach aims to keep humans “in-the-loop,” not just as consumers of the learning analytics, but also able to shape what is being shown or highlighted in the interface.

4. Illustrative Study

This section presents an exploratory study that illustrates the feasibility of our approach. We organize this as follows:

- 4.1 A description of the collaborative learning situation from which a dataset was captured and student-facing dashboards were used.
- 4.2 The design rationale of handcrafted data storytelling-enhanced versions of the dashboards.
- 4.3 An evaluation study with teachers to understand the role played by the data storytelling elements to support sense-making.

4.1. The Learning Situation

Fifteen undergraduate students, ranging from 21 to 26 years old, enrolled in an introductory Database Systems subject from ESPOL University in Ecuador, were asked to use a multi-display interactive tool (Echeverria, Martinez-Maldonado, Chiluiza, & Buckingham Shum, 2017) as part of their regular classes, to practice database design related concepts. This tool includes a collaborative entity-relationship (ER) diagram editor distributed across tablets and an interactive tabletop and a multi-display dashboard setup for post-hoc. The system captures interactions (e.g., insert, delete, move, edit) with objects (entities, attributes, relationships) that students perform when designing ER diagrams. With the data logs and the input solution from the teacher, the tool automatically generates a dashboard containing information about social and epistemic aspects of the activity, aiming at supporting the teacher debrief and student reflection.

Students were grouped in teams of three members (five teams in total) and allocated in 30-minute sessions to design an ER diagram for a given case study. Figure 3 shows two visualizations about their **participation** (top) and **performance** (bottom) automatically generated from a session. Visualizations about **participation** show the number of interactions by each student plotted on a timeline. Each student is identified as P1, P2, and P3 and differentiated by colours (yellow, blue, and red). Visualizations about **performance** show how the solution (the team’s ER diagram) evolved. Each data point indicates the number of objects, i.e., entities/attributes (green) and/or relationships (orange) added to the solution. The data point label (above each point) depicts the current (automatically calculated) score of the whole diagram according to the objects added up to that time. The text label below each point shows the number of entities/attributes (e.g., 6E) or relationships (e.g., 1R) added to the design in a 1-minute window.

These and other visualizations were evaluated with the students (Echeverria et al., 2017). A critical outcome of this study was that students used the dashboard visualizations as an analytical tool. It allowed them to *explore* the data, but not necessarily to focus on specific parts of the charts that could have helped them gain any insights from a pedagogical point of view. These results obtained from an authentic scenario motivated the need for enhancing these visualizations with data storytelling elements to make them more explanatory and prompt specific reflections based on the intended learning design.

4.2. Designing Enhanced Explanatory Visualizations

In this section, we describe the design process to add data storytelling elements to the original visualizations shown to students (depicted in Figure 3). Note that although in this study the visualizations were handcrafted, the purpose is to provide a detailed explanation of how this process can be automated. The process consists of the following steps:

Step 1: Understanding the context. Our context is shaped by the teacher’s learning design. The teacher wanted to provide particular information to students/groups in order to improve their awareness about collaboration skills. Also, she wanted to know the how students’ ER diagrams evolved, determining missing concepts and relationships throughout the design process.

Step 2: Mapping teacher intentions with rules. Our second step consisted in extracting the teacher’s **intentions** to be represented through data storytelling elements. For instance, the teacher explicitly stated that she had two critical learning intentions, which were also the drivers of the design of the two visualizations generated for the students:

- *Learning intention 1 (LI1 – Participation):* All students within a group should ideally participate equally.
- *Learning intention 2 (LI2 – Performance):* Some specific database elements are critical and should be included early in the database design. This can directly impact the quality of the final design.

Original Exploratory Visualizations (OVs) About Student Participation

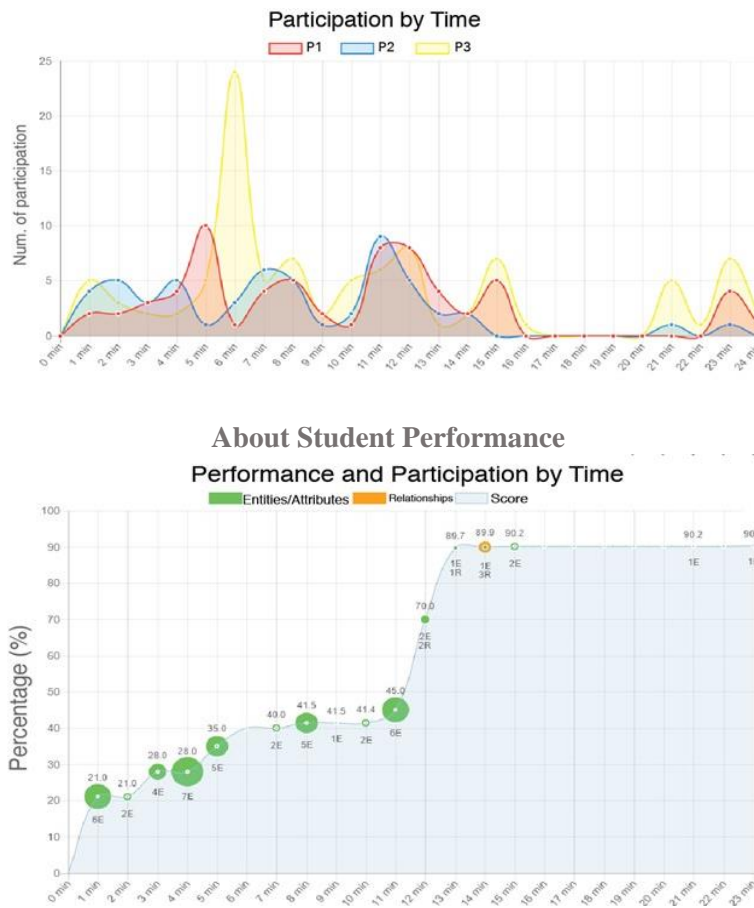


Figure 3. Original exploratory visualizations rolled to students. Top: participation of each team member. Bottom: Evolution of the team’s performance, from a random team.

From these intentions, we built rules that can be implemented in order to automate the generation of new visualizations by translating the teacher’s learning intentions into data storytelling elements. For instance, for LI1 we can create an informally expressed rule (R1): “If a student is dominating the activity in a pre-defined time slot (e.g., for 5 minutes), show [personalized message].” For the case of LI2, we can create a rule (R2): “If [name of a key entity or relationship] is added into the solution, add a marker indicating the name of the entity (or relationship).”

Step 3: Translating rules into data storytelling elements. From the set of data storytelling elements introduced earlier, we designed a set of rules to encode the learning design intentions with visual elements, producing a new set of “data storytelling-enhanced” visualizations. Table 1 lists how some rules were used to encode each learning intention, focused on participation (Rules 1 and 2) and performance (Rules 3 and 4) respectively.

- **Rule 1:** If the teacher’s intention is about knowing equal group participation, the rule will search for patterns in the graph where this behaviour is explicit. Translating this into a rule, would be as follows: if a student was dominating (i.e., participation of one student is higher than the other two students) by x minutes (the teacher defines x), a visual indicator could be generated on the visualization itself highlighting this behaviour (see *if-then rule* in Table 1, row 1). For example, visual elements can be added to the graph such as a label with narrative text or triggering a visual alarm.
- **Rule 2:** If the teacher’s intention is to know the overall team’s participation, the rule will determine if all students featured balanced participation (e.g., using an entropy function). A message could be set to explain this behaviour (see the *if-then rule* in Table 1, row 2). Then, adding the resulting message as a big prescriptive title in the graph

could be a way to depict this behaviour.

- **Rule 3:** If the teacher’s intention is to identify when a database element that makes the task performance increase is added (this element also is defined by the teacher), then the rule can insert a key data point when a student performed that action (see *if-then rule* in Table 1, row 3). Visual elements that support this rule include the following: adding an icon or varying the colour of a marker to differentiate types of data points; adding labels with a numerical value to emphasize current score or text labels to describe the name of the entity added.
- **Rule 4:** If the teacher’s intention is to visualize the overall team’s performance and assess if this score corresponds to low or high performance, a rule can be formulated to determine this depending on the team’s final score and an x threshold (the teacher defines x , see *if-then rule* in Table 1, row 4). We can place the resulting message in the title of the graph or as a caption on the graph.

Table 1. Cases Illustrating How to Map Intentions, Rules and Data Storytelling Elements

Visualization	Intentions	Rules	Data storytelling elements	Example
LI1: Participation	Highlight whether the group is contributing equally or not	Rule 1: if dominating (P1,P2,P3) == “P1” in slot_time(x): show_message = “P1 dominated the activity during x minutes” highlight_series (P1)	narrative text; highlighting data series	Fig. 1 -left (B)
	Summarize the overall group contribution	Rule 2: if balanced (P1,P2,P3) == 1: show_message = “All participants had a balanced participation through the activity”	prescriptive title; narrative text	Fig. 1 -left (A)
LI2: Performance	Highlight elements that increase the design’s quality	Rule 3: if entity == “Scholarship”: highlight_point (entity) show_score (line) message = “Entity Scholarship”	narrative text; highlighting data point; shaded area	Fig. 1 -left (D)
	Summarize the overall group performance	Rule 4: if overall_performance (team) >= 75: show_message = “High-performing team”	prescriptive title; narrative text	Fig. 1 -left (A)

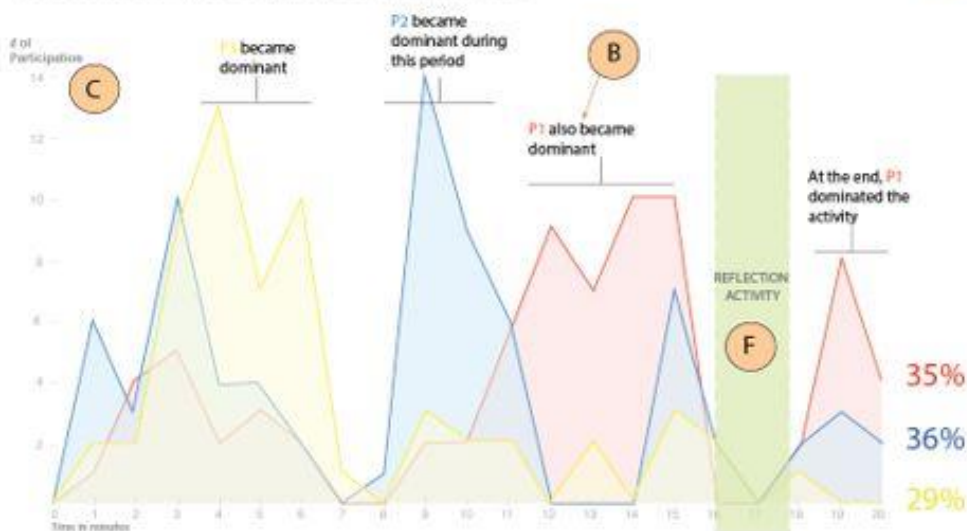
Step 4: Crafting the visualizations. The final step corresponds to generating the explanatory visualizations with data storytelling elements, by applying the rules defined previously. A set of visualizations was prototyped using graphic design software (Figure 4). In addition to the mapping of intention-rule elements, we completed three additional steps regarding the design of the visualizations:

1. Removing elements that do not add value to the graph. This step is known as *decluttering* and, as explained above, it involves removing grids, eliminating indirect legends, reducing the number of colours used, and removing or de-emphasizing non-relevant data point markers or data series. Therefore, we proceeded to make everything into a grey scale to then highlight and/or emphasize only important aspects according to the rules (see the data storytelling elements column in Table 1). This step can be observed in Figure 4C, where we remove grids, legends, and data points. Also, axis labels were rotated, as suggested by Knafllic (2015) so text could be easily read horizontally.
2. Emphasizing lines to show trending in the performance visualization (Figure 4E, bottom).
3. Adding context through a shaded area (shapes and colour) concerning the learning design (Figure 4F, corresponding to the reflection stage).

Visualisations with Data Storytelling Elements (VDSs) About Student Participation

Students had unbalanced participation across the team activity — A

The 3 students dominated the participation at different times



About Student Performance

High-performing Team — A

The team was able to determine all main entities and all main relationships

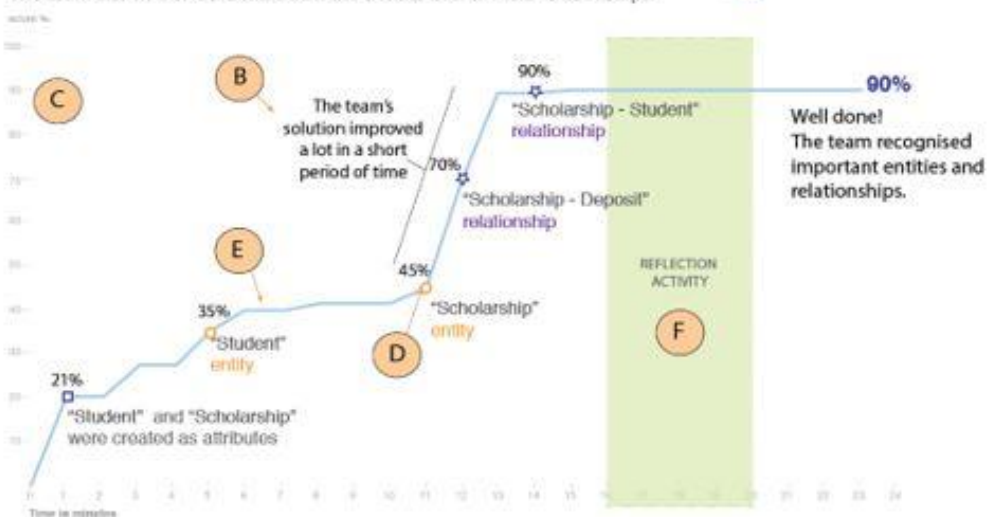


Figure 4. Explanatory visualizations after applying data storytelling principles and elements (A-F). Top: Participation of three students. Bottom: Evolution of the team’s performance.

4.3. Teacher Evaluation Study Design

This section describes an exploratory study conducted, with the aim of generating an understanding of the role of the data storytelling elements to support sensemaking. Through this study, we illustrate how teachers explored and interpreted exploratory and explanatory visualizations.

4.3.1. Study Setup

A study was carefully designed to expose teachers to a number of exploratory visualizations with data storytelling elements (VDS) and original exploratory visualizations (OV) shown previously to the students (Echeverria et al., 2017). Figure 5 presents an overview of the study setup. The intention of this study is not to experimentally test two conditions (e.g., exploratory vs. explanatory visualizations) but rather to focus on generating an understanding of the potential role that data storytelling elements can play to support interpretations of visual learning analytics and dashboards. With this study, we tried to obtain evidence about how teachers reacted to the different types of visualizations, both exploratory or explanatory, and visual design elements. The analysis is focused on three themes:

- **Teacher reactions.** We first investigate what kind of stories or comments could be formulated by teachers while looking at the visualizations. We hypothesize that although both exploratory (OV) and explanatory (VDS) visualizations may be useful for teachers to gain insights into learner activity, the data storytelling elements would play a role in helping teachers aligning those insights with those intentions materialized in the learning design.
- **Teacher behaviour and focus of attention.** We explore whether explanatory visualizations drive teachers’ focus of attention to specific elements (such as text labels, the title, and key data points) as intended.
- **Teacher preferences.** We examine the usability of visualizations, especially those with data storytelling (VDS) and their power in supporting teachers’ monitoring for orchestration of the collaborative activity (Rodríguez-Triana et al., 2015). We also explore which data storytelling elements are most helpful to support storytelling through learning visualizations.

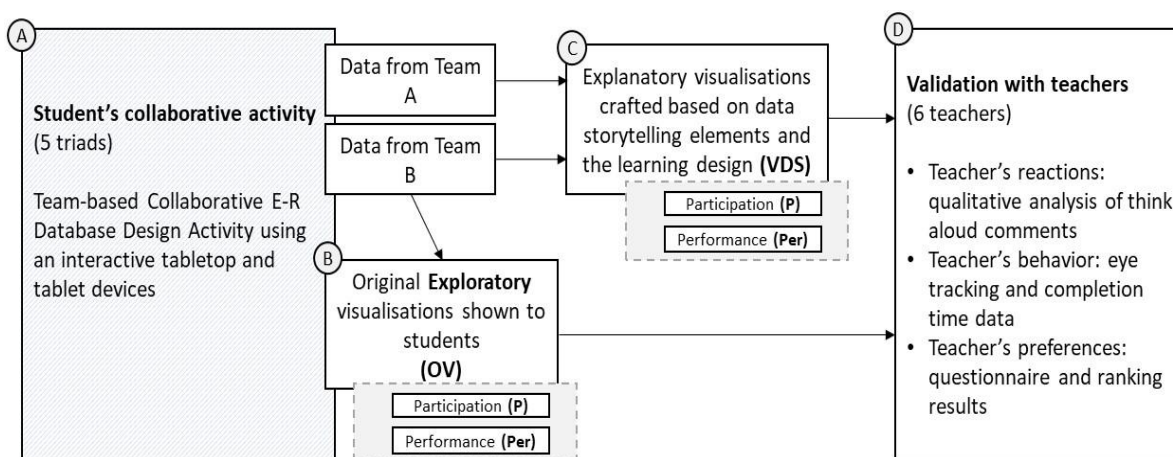


Figure 5. Validation process from an authentic student collaborative activity (A). A series of visualizations (OVs) were shown to those students in a dashboard as formative feedback (B). These visualizations (VDSs) were enhanced based on data storytelling elements (C). Both types of visualizations were then shown to teachers to analyze their reactions (D).

4.3.2. Participants

Six assistant professors (T1–6: one female, five males; age range 26–39 years) from an Ecuadorian university were asked to participate in this study. All teachers had experience in teaching (avg.: 2.33 years), five of them had experience teaching the database systems course, and all of them had experience teaching software design.

4.3.3. Materials

We prepared a total of eight visualizations using data from the study described in section 4.1 (see Figure 5A). These visualizations were generated from two randomly selected student teams (team A and team B, henceforth). We picked four original visualizations (OVs) as shown to the students during the reflection task (see Figure 5B). Then, we applied data storytelling principles and elements to each of the OVs (as explained in Section 4.2), resulting in the second set of four prototypes (see Figure 5C). Figure 4 depicts one example generated from team A’s participation and performance data respectively. In short, each teacher would inspect data from two teams, regarding participation and performance, through visualization with (explanatory — VDSs) and without (exploratory — OVs) data storytelling elements.

4.3.4. Method

The validation with teachers consisted of four tasks (Figure 5D). A think-aloud protocol was used during the whole pilot study.

For the *first task*, each teacher explored both the original (OVs) and the prototyped visualizations (VDSs). The researcher asked each teacher to inspect the visualizations and tell a story about what happened with the team based on what she could interpret from the data (**teacher reactions**). From now on, we will refer to the *inspection episode*, as the task of exploring/verbalizing insights from one visualization.

In order to control for a familiarity effect, we set up two cases of inspection episodes:

- *Case I* included two OVs from team A (participation and performance) and two VDSs from team B (participation and performance) — four in total.
- *Case II* included two VDSs from team A (participation and performance) and two OVs from team B (participation and performance) — four in total.

Therefore, half of the teachers inspected Case I and the other half Case II. In summary, we captured 24 inspection episodes from the six teachers. During the whole first task, data from an eye tracker (Tobii X-60 Pro) was recorded, with the aim of understanding how teachers visually explored the visualizations (**teacher behaviour and focus of attention**).

The *second task* involved the exploration of visualizations regarding their perceived usefulness for supporting monitoring and orchestration of the learning activity. Here, teachers explored OVs and VDSs about participation and performance for teams A and B (e.g., Figure 3, top, representing team A and team B participation, on the same screen). Thus, each teacher had the following sequence of inspection episodes: OVs about participation for team A and B, OVs about performance for team A and B, VDSs about participation for teams A and B, and VDSs about performance for teams A and B (four in total). Teachers were asked to provide commentary on the ease of interpretation and support for orchestration and evaluation (**teacher preferences**).

In a *third task*, teachers compared both versions: OV and VDS on participation for a random team (A or B) and OV and VDS on performance for a random team (A or B; e.g., Figure 3, top, and Figure 4, top, on the same screen). Teachers were asked to verbalize their thoughts about which of the two versions helped them better interpret the story and insights behind the data. Also, teachers were asked to suggest any change or addition to the VDSs (**teacher preferences**).

The *fourth task* involved the evaluation of specific data storytelling elements. For this task, teachers were asked to rank each of the data storytelling elements (A–F) regarding their perceived helpfulness for storytelling. A 5-point ranking scale was created to allow teachers to rank the role that each data storytelling played in the VDSs on participation and performance (**teacher preferences**).

Video-recordings, eye-tracking data, screenshots, and responses to the questionnaires were gathered and organized for its subsequent analysis. A post-hoc semi-structured interview was conducted with each teacher after each session for them to explain any additional aspect of their experience.

5. Analysis and Results

A mixed methods analysis approach was followed, aimed at triangulating quantitative evidence collected via the transcripts of the video-recordings of the think-aloud processes and interviews, the ranking instrument, and eye-tracking data with teachers' qualitative reactions. In this section we summarize the analysis of teachers' data stories and comments, gaze behaviour/focus of attention and preferences regarding the visualization purposes, exploratory (OV) or explanatory (VDS), and the type of data each represented (data about a team's participation and performance).

5.1. Teacher Reactions

As an initial exploration of the evidence we gathered, we looked at the kind of reflections that teachers articulated while inspecting each of the visualizations.

For *exploratory visualizations* (OVs) we noticed that teachers tended to focus on describing the data and then raise questions as they generate hypotheses to explain the data. For example, for the case of OVs about **participation**, teachers recognized that the visualization was presenting all the information about individual student participation. This behaviour was affirmed by teacher T1 as follows: "With this visualization, I can see an overall behaviour of a team's participation." However, teachers' comments related primarily to counting the number of actions performed by each student during particular periods of time, and guessing why sometimes a student performed more (or fewer) actions, instead of gaining specific insights from the data. For instance, this is illustrated by one of the comments expressed by teacher T2 as follows: "At some point, students in this team stopped the activity, but it is not clear why, maybe they were thinking, or only talking." Teachers also tried to explain two or more students' participation, pointing at the equality (or imbalance) in the team's participation. One of the teachers said "It seems that the participation was imbalanced. The major number of actions were performed by participant 1 and participant 2. Participant 3 was only active during the first seven minutes. I have the impression that only two participants worked actively throughout the time" (T3). Finally, some teachers attempted to identify potential leaders by counting the number of actions and guessing turn-takings among participants during the activity. However, as expressed by one of the teachers (T1) this proved not always to be an easy task: "I can easily see who has dominated the activity in periods of time.

However, I cannot tell which participant dominated the participation during the whole activity quantitatively.” Similarly, regarding the OVs about team **performance**, comments were related to the performance of the final product, how it evolved as students added entities and created relationships and how each action made a (positive or negative) impact on the performance. One of the teachers (T3) said: “Some entities were identified, some were right, others wrong, but the final performance was not any better.”

In short, exploratory visualizations (OVs) prompted teachers to construct stories, but they struggled to know if they were accurate: they inspected the data, described it, and raised questions. Some of their questions could be answered from the data, while others could not be from the limited evidence provided. This behaviour is in line with what Verbert et al. (2013) called the “learning analytics process model”: 1) data was presented to teachers; 2) they immediately started to formulate questions, and 3) assessed how data is useful or relevant for addressing those questions. The critical problem is that teachers were not able to elaborate on statements that could sound like interesting insights about student participation or performance. In this way, teachers were performing the analysis of the data without being trained to do so. This behaviour was precisely what we expected to happen when teachers face exploratory visualizations.

For *explanatory visualizations* (VDSs), we hypothesized that teachers would elaborate on insights gained from the visualizations, rather than merely describing the data. In the case of the VDSs about **participation** (e.g., see Figure 4, top), comments were related to a participant’s dominance during different periods of time and equality of participation. Moreover, teachers were more confident about dominance. This was stated by teacher T6 as follows: “It can be clearly seen when a team member is dominating the participation in different time periods.” For the VDSs about team **performance** (e.g., see Figure 4, bottom), teachers articulated their thoughts regarding how key actions made an impact on the final product. Teachers were able to direct their attention to critical moments in a team’s process rather than analyzing all the data points. Teacher T6 described this behaviour as follows: “This [VDS] is more understandable when showing what each object means [referring to the text of each entity appearing in the visualization]. For example, this visualization shows me that during the first minutes, students identified two objects as attributes, but then, these changed to entities. It was interesting to see how the product evolved, and how the changes made to critical objects affected the performance.” Also, another comment from teacher T2 reinforced this idea as follows: “It gives me more detail about the number of entities that were considered for the final score.”

Teacher T6 explicitly appreciated that the explanatory visualization reduced the time they needed to spend processing the data: “This [VDS] visualization is explicit, ... the information is auto-explained.” Another teacher (T5) supported this suggestion: “This graph [VDS] was more detailed, informative. It does not make me overthink about the objects created. In the other visualization [OV] I have to start guessing what 6E, 7E was. Then, after a while, I realized that these were the Entities or Attributes.”

As we intended, teachers were able to construct more detailed stories about each visualization. One excerpt of a teacher’s story (T2) during the visual inspection of VDS-performance was as follows:

At the beginning, the team created two entities, Students and Scholarship [reads the text of the data points]. The team did not make much progress during several minutes, from minute three to minute seven the team did not have significant activity [reads the text label]. Then, they were adding some objects, but the performance just raised a little bit. It is a low-performance score for fifteen minutes of activity. They have a short reflection time [reads the text in the shaded area], and then, they added one relationship, but it was not enough. In the end, the performance was very low [reads the text].

Another teacher also did use the text labels to explain the story (VDS participation) as follows:

At the beginning of the activity, the three participants did not have any big differences in their participation. However, from minute five to minute seven, during two minutes [reads the text labels], participant three (P3) had more participation, he dominated the participation [reads the text labels]. Then, P3 also was dominating the participation between minutes fourteen and fifteen. Then, there was a reflection time [reads the text in the shaded area], and no participation was recorded there. Finally, from minute 20 to minute 23 P3 became dominant [reads the text label] (T4).

One question that emerged during this analysis was whether teachers used any of the data storytelling elements as a basis for their comments. Thus, we coded each phrase when a teacher grounded their comment on the evidence provided by a particular data storytelling element. Seventy-seven comments were coded as *text labels*, *axis*, *data points*, *reflection*, or *title*. Summarizing the coding results, we found that **text** narratives were the most referenced elements with 48%, followed by highlighted **data points** with 22%, the **shaded area** (that depicted a student’s reflection period) 16%, and the **title** (12%).

To conclude this section, the results from this initial qualitative analysis are encouraging. They suggest that teachers reacted as intended to the data storytelling elements. The exploratory visualizations (OVs) invited teachers to explore all the data but

leaving the responsibility on them to discern insights. In contrast, the explanatory visualizations (VDSs) were more directive about the meaning of key points, which the teachers found plausible, and seemed to appreciate.

Self-report is a rich source of evidence, but what did the teachers, in fact, attend to visually? The next step in the analysis consisted of an analysis of teachers’ visual behaviour when presented with these dashboards, in order to provide answers to questions such as the following: *Where did teachers look in a given visualization? What was the role played by each data storytelling element? Did the explanatory visualization help them explore the data more efficiently?* In the next subsection, we address these questions based on a quantitative analysis of gaze behaviour.

5.2. Teachers’ Gaze Behaviour

In this section, we describe how eye-tracking data shed more light on whether the addition of data storytelling elements appeared to shape teacher engagement with the visualizations.

5.2.1. Time Spent Inspecting the Visualizations

As a starting point to understanding teachers’ gaze behaviour, we calculated the time they took for each inspection episode. Table 2 summarizes the minimum (min.), maximum (max.), average (avg.), and standard deviation (std. dev.) of the time spent in inspection episodes of the six teachers (n=24). Given the higher standard deviation and the small sample size, we did not find significant differences regarding the average time spent inspecting exploratory and explanatory visualizations (U=51, Z=-1.212, p=0.242). However, we can observe differences in standard deviations according to the purpose of visualization, being larger in the explanatory visualizations. This suggests that teachers inspected the explanatory visualizations for similar amounts of time. In contrast, some teachers spent more time inspecting the exploratory visualizations.

Table 2. Min, Max, Average Time Spent, and Standard Deviation for Each Inspection Episode (n=24)

Purpose of visualization	Type of visualization	min. (secs)	max. (secs)	avg. (secs)	std. dev.
Exploratory (OV)	participation	115.11	513.00	220.86	148.71
	performance	91.78	727.07	266.69	223.91
Explanatory (VDS)	participation	57.86	228.69	148.14	57.74
	performance	128.57	266.49	183.09	61.71

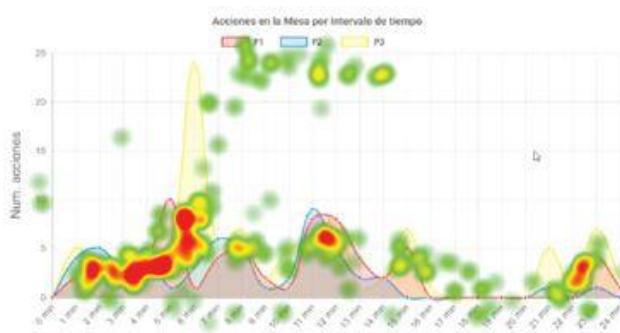
5.2.2. Heatmaps of Teachers’ Gaze Behaviour

The first step in the analysis of teachers’ gaze behaviour consisted of visually inspecting which areas of the visualizations attracted most visual attention. To address this, we generated heatmaps based on the gaze data for each of the teacher’s inspection episodes. First, we generated heatmaps using all the data points captured by the eye tracker. However, after observing the generated heatmaps, we realized that the teacher’s gaze covered almost the whole visualization by the end of the inspection episode, making it difficult to determine which elements in a visualization appeared to be critical in driving attention. Therefore, in order to enable a normalized visual comparison of the gaze behaviour, we generated heatmaps based on the first 60 seconds of each inspection, to try and capture the differences that the design has when first engaged. We generated 24 heatmaps corresponding to the four different teachers’ inspection episodes.

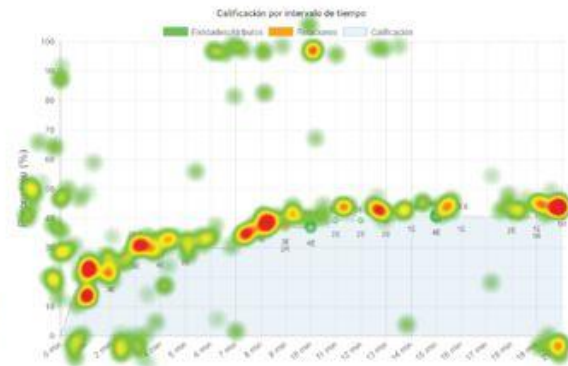
Figure 6 shows one example for each visualization type (OVs and VDSs about student participation and performance each). The heatmaps were scaled so that one second’s fixation corresponds to a brighter red colour. It can be seen that for the exploratory visualizations (Figure 6, top) many hotspots are spread through the data points and legends (located at the top and centre of the OVs), suggesting that the teachers dedicated considerable time to looking at each data point in the time series. These examples are representative of what was similarly observed in the other teachers’ heatmaps (other examples of heatmaps can be found at <https://goo.gl/QQuumyz>). In the interviews, most teachers explained that they struggled with interpreting the small numbers and letters in each data point and had to consult the legend to remember the meaning of each colour. Although this is a simple design issue, quite easily solved during the decluttering of the graph, it is a subtle data storytelling consideration that can have a critical impact when teachers inspect a graph. By contrast, concerning the explanatory visualizations, Figure 6, bottom, depicts how the text labels and key data points drew most visual attention. In these two examples, teachers only focused on the narratives and the highlighted data points rather than the axis or the grid. From a visual comparison of the heatmaps we observed that, in the first 60 seconds of seeing the display, visualizations that featured a clean design, and only highlighted selected data points, and text labels (entity names and percentage of the score) seemed to be effective in reducing visual clutter and helped teachers focus their attention. While the heatmaps assisted us in performing a rapid visual comparison, the next step in the gaze analysis was to test whether we could quantify these differences.

Original Exploratory Visualizations (OVs)

About Student Participation

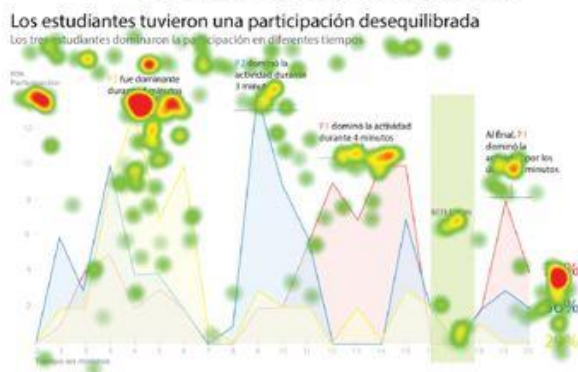


About Student Performance



Explanatory Visualisations with Data Storytelling Elements (VDSs)

About Student Participation



About Student Performance



Figure 6. Example heatmaps of gaze behaviour of exploratory (top) and an explanatory (bottom) visualizations from four different teachers.

5.2.3. Areas of Interest (AOI)

To provide meaning to the data points captured by the eye tracking system, we identified the most relevant Areas of Interest (AOI) in each visualization. This approach is a common practice in eye-tracking analysis methodologies to associate pixels in the screen with higher order areas (Toker, Conati, Steichen, & Carenini, 2013), in our case, data visualization elements. Table 3 presents the list of AOIs regarding the purpose of the visualization, either exploratory or explanatory, the latter being related to data storytelling elements. Each visualization features some shared and some distinct design elements. For example, AOIs that were common across all the visualizations tested in the study included the axis (x and y) and the data points in the time series. The exploratory visualizations were the only ones containing legends. The explanatory visualizations contained three unique AOIs: text narratives, a shaded area, and a prescriptive title. The final AOI was a “not applicable (NA)” area in a visualization, corresponding either to a blank space or another irrelevant element of the screen.

Table 3. Areas of Interest (AOI) Related with Each Visualization Purpose

AOIs in exploratory visualizations	AOIs in explanatory visualizations	Acronym
Axis (x)	Axis (x)	Ax
Axis (y)	Axis (y)	Ay
Data points	Data points	DP
Legends	–	Legd
–	Text narratives	Txt
–	Shaded area	ShA
–	Title	Title
Non-applicable	Non-applicable	NA

5.2.4. Areas of Interest and Gaze Behaviour: Clustering Analysis

We performed a cluster analysis for each of the two types of visualization (OVs and VDSs) with the aim of giving meaning to the eye tracking data points according to the AOIs introduced above. A total number of 147,045 data points (33,184, 43,251, for participation and performance OVs, and 35,988, 34,622 for participation and performance VDSs, respectively) were analyzed at this point. We applied the Elbow method (Kodinariya & Makwana 2013) to identify the appropriate number of clusters in the dataset. We ran this method by grouping all the data points of the inspection episodes of each of the eight different visualizations (two OVs and two VDSs for two teams, A and B, about two types of visualizations, participation and performance). This way, we found that k=15 would be a reasonable number of clusters that would provide enough information to differentiate AOIs across visualizations. Then, we ran the k-means (k=15) cluster algorithm from the scikit-learn Python package (Pedregosa et al., 2011) for each visualization, grouping the data points of the three teachers’ inspections for Case I and the other three teachers’ inspections for Case II (as explained in section 4.3.4). Figure 7 and Figure 8 show two examples of the resulting clustering for an OV and a VDS (respectively) on participation for Team A (Figure 3, top, and Figure 4, top, show original visualizations without clustering points). The centroid of each cluster is marked with a circle.

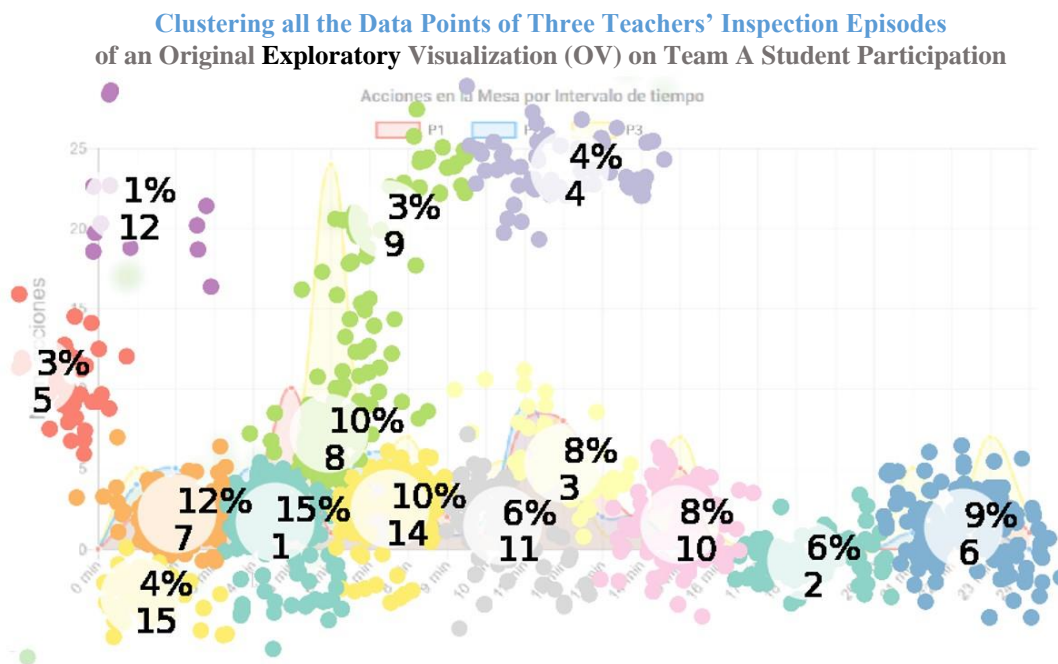


Figure 7. Results of clustering the eye tracking data points of the three inspections episodes recorded for one OV about the team’s participation (k=15). Circles correspond to the centroids of each cluster. Each is associated with the percentage of data points of the cluster (at the top of the centroid) and the id number of the cluster (at the bottom of the centroid).

A visual inspection of these clusters gives an idea of the parts of the chart that the eye tracking data points were associated. To associate clusters with meaningful AOIs, we manually assigned one or two AOIs (from Table 3) to each cluster. When a cluster did not belong to a relevant AOI, we tagged it as a non-applicable AOI. For example, in Figure 7, eye tracking data points in cluster 4 are mostly associated with the *legend* at the top of the chart, cluster 5 with the vertical axis (*y-axis*), cluster 15 with the horizontal axis (*x-axis*), cluster 6 with the horizontal axis and student data of the time series (*x-axis/data points*), etc. Similarly, clusters in Figure 8 were associated with AOIs of the visualizations with data storytelling elements. For example, clusters 1, 2, 5, and 15 were associated with text labels (*text*). Table 4 summarizes the proportion of attention placed on each AOI regarding each visualization purpose (exploratory, explanatory) and type (participation, performance).

Clustering all the Data Points of Three Teachers' Inspection Episodes of an Explanatory Visualization (VDS) on Team A Student Participation

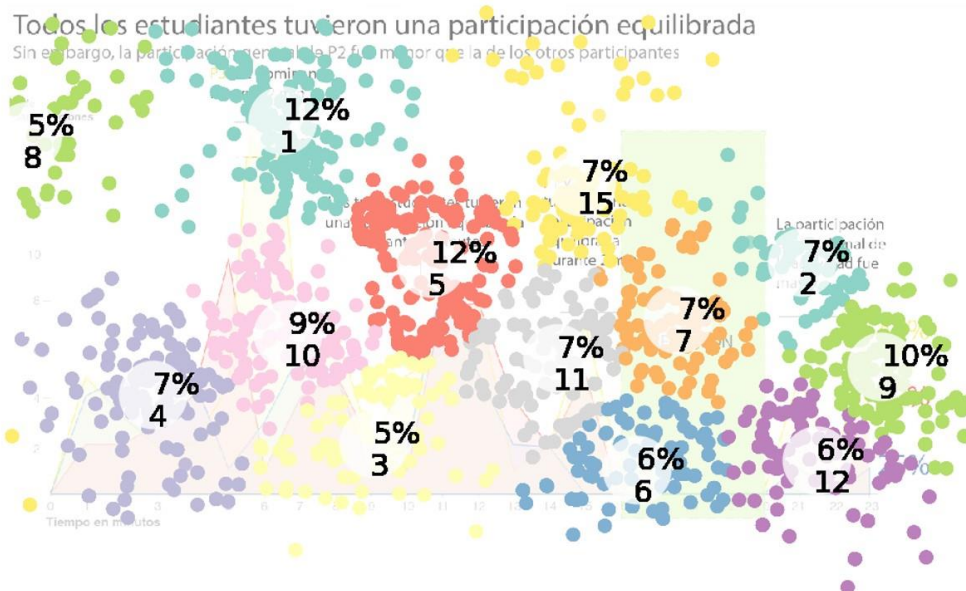


Figure 8. Results of clustering the eye tracking data points of the three inspection episodes recorded for one VDS about the team’s participation (k=15). Circles correspond to the centroids of each cluster.

Table 4. Percentage of Occurrence of Each AOI by Visualization Type from the Six Teachers

Exploratory visualizations (OVs)				Explanatory visualizations (VDSs)			
Participation		Performance		Participation		Performance	
AOI	%	AOI	%	AOI	%	AOI	%
x-axis/data points	48.6	data points	73	data points	32.4	text	41.4
data points	30.2	x-axis/data points	10.7	text	26.7	text/data points	29.4
legends/data points	10.1	legend	7.9	text/data points shaded	13.5	shaded area	4.5
y-axis	4.8	y-axis	5.1	area/data points	8.4	x-axis	4.3
non-applicable	2.4	x-axis	2.9	text/title	6.6	data points	3.8
legend	2	non-applicable	0.4	shaded area	3	y-axis/title	3.6
x-axis	1.9			y-axis	2.7	x-axis/data points	3.6
				x-axis/data points	2.5	y-axis/data points	3.5
				y-axis/title	1.7	x-axis/text	2.9
				title	1	title	1.3
				x-axis	0.92	x-axis/shaded area	1.14
				Non-applicable	0.61	Non-applicable	0.46

For the exploratory visualizations on participation and performance (columns 1 and 2 respectively) teachers spent more time inspecting the **data points** and the **horizontal axis** — 78.8% and 83.7% respectively from *data points*, *x-axis*, and *x-axis/data points* AOIs. The third most inspected AOI was the **legend**. This confirms the qualitative results presented in the previous section. Teachers focused on exploring the data points and making sense of them by looking at the legends and the

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial- 99 NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

axis. In contrast, for the explanatory visualizations (columns 3 and 4) teachers read the **text narratives** (text) along with the **data points** — 72.6% and 70.8% for visualizations on participation and performance respectively, from *data points*, *text*, and *text/data points* AOIs. This suggests that the addition of brief text narratives attract the attention of teachers. If these narratives are brief and communicate insights, they can become effective design elements to be added to learning analytics visualizations. Moreover, in this case, teachers looked at the axis to a lesser extent compared with the exploratory visualizations. This suggests that decluttering the visualizations allowed teachers to focus better on key information.

5.2.5. Scanning Trajectories

To complete the teacher gaze behaviour analysis, we wanted to know which visualization elements attracted the most attention in the first seconds of each inspection episode. To achieve this, we focused on the gaze data during the first 10 seconds of each episode. By associating eye tracking fixation points with the AOIs (through the clustering analysis explained in the previous subsection), we generated visual representations of sequential trajectories that each teacher followed to make sense of the visualization when facing it for the first time. Given the small size of our sample ($n=24$) and the limitations of the space, in this subsection, we illustrate some of these trajectories, shown in Figure 9 and Figure 10.

For the trajectories related to **exploratory visualizations** on student participation, we can see in the example in Figure 9, top, that this teacher (T2) first explored the vertical axis to then start looking at the data points (see transitions $t1-t2$ between nodes Ay and DP at the bottom-left of the figure). Then, the teacher looked at the horizontal axis ($t4$) before inspecting the data points in the middle of the graph ($t5$). Then, the teacher looked up at the region where the legend is and looked at the axis and data points coming back and forth to the legends ($t6-t14$). For the exploratory visualization about student performance (see Figure 9, bottom), this second teacher (T3) started inspecting the data points at the middle of the visualization (see $t1$ and $t2$) and then his fixation was focused on the legend ($t3$). After this, the teacher gaze moved to the vertical axis ($t4$), the horizontal axis, and various data points ($t5-t7$) to then move back to the legend ($t8$). The next eye movements scanned more data points, but the teacher had to return to the axis and the legend while inspecting the data points ($t9-t28$). Interestingly, triangulating the data from these trajectories and the interviews, we found that at the beginning of exploring these visualizations, teachers commonly verbalized the name of the elements they were looking at and started to ask clarification questions. For example, teacher (T2) inspecting the first figure said: “What is the meaning of P1, P2, and P3 in the legend? Are they the participants in the activities, right?” The teacher inspecting the second figure (T3) commented: “[after being quiet for a period] the percentage [the label in the vertical axis] is the score obtained, right? So, the labels related to 5E, 3E? Also, the orange colour in the data points is because students added a relationship object, right?” These examples further suggest that teachers needed some time or some training before they could gain any insights from the visualizations at a glance. As illustrated by these exemplars and confirming the summaries of gaze analysis presented in the previous section, **data points, axes, and legends are essential to make sense of the information for exploratory visualizations.**

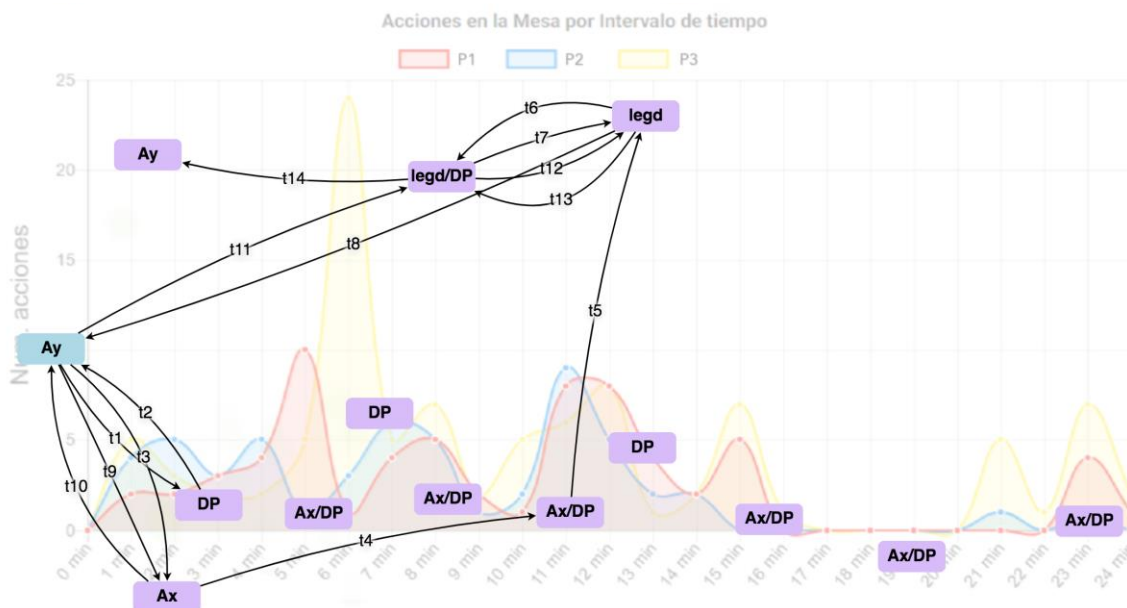
By contrast, regarding the trajectories related to **explanatory visualization** about student participation, we can see in the example depicted in Figure 10, top, that the teacher (T5) started to inspect the shaded area first (reflection) to then move to the first text narrative next to it ($t1$) and some data points below that text (see $t2$). After that, the teacher’s gaze moved to the other pieces of text (see $t5$, $t8$, $t10$). Indeed, this teacher did not spend much time looking at the data points in the first seconds but instead paid more attention to the narratives. The teacher (T6) in the second example (Figure 10, bottom) started inspecting a text narrative at the centre of the visualization and then moved up to the title ($t1-t3$). After, this teacher also looked at a second narrative ($t4-t5$), and then started to scan the highlighted data points from left to right (see $t6-t10$).

These two examples show how the data storytelling elements — the text narratives, the title, and the highlighted data points — drove the attention of these teachers from the beginning. These examples are representative of what was similarly observed in the other teachers’ heatmaps. This suggests that teachers may have dedicated more time to gain insights from the text rather than expending time in trying to make sense of the graph. From the interviews, one teacher (T5) described this critical difference as follows: “This is a visualization about who dominated the participation [referring to an explanatory visualization] and not just the number of actions per participant [referring to an exploratory visualization].” Besides, data storytelling elements supported the teacher’s stories and comments. One teacher (T6) told a story utilizing the same information illustrated in the text narratives, data points, title, and shaded area. Moreover, this teacher followed the same trajectory as the one presented in Figure 10, bottom, when verbalizing the story and comments: “Initially, students added two entities, but as attributes ($t6$). Then, the students added some entities. An object that was added previously as an attribute, Student, now became an entity ($t7$). Then, it [referring to the visualization] depicts a period of inactivity. The team did not make any progress for four minutes ($t8$). Then, the team made progress on seven percent of their actual performance ($t9$). After this, the team’s performance remained the same for seven minutes. They realized that the object Scholarship was an entity instead of an attribute” ($t10$).

In summary, comparing these results with the results from the qualitative analysis, data storytelling elements drive the teacher’s attention and support the externalization of visualization insights effectively.

Two Examples of Teachers' Trajectories in the First 10 Seconds of Inspection.

Inspection of an Exploratory Visualization (OV) of Student Participation



Inspection of an Exploratory Visualization (OV) of Student Performance

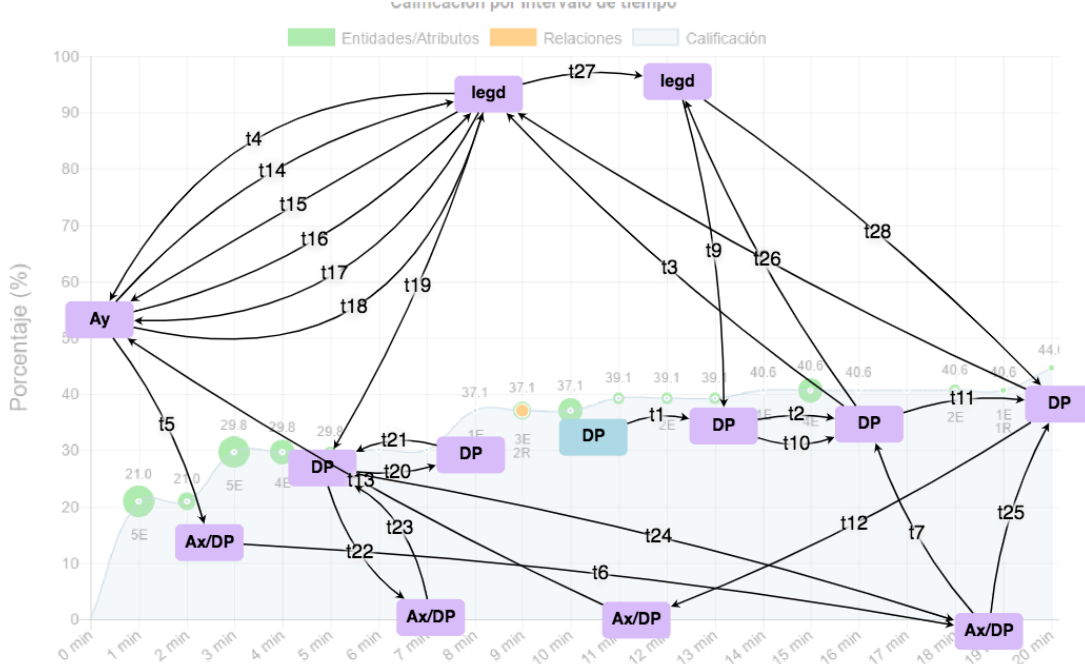
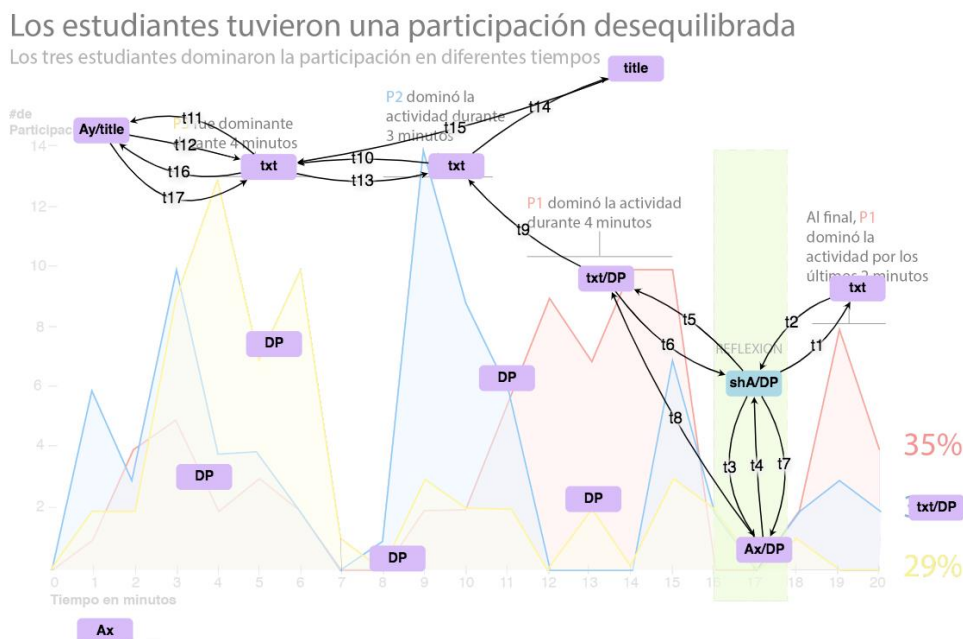


Figure 9. Example trajectories followed during the first 10 seconds of inspection of the OV on student participation (top) and student performance (bottom) by T2 and T3, respectively.

Two Examples of Teachers' Trajectories in the First 10 Seconds of Inspection

Inspection of an Explanatory Visualization (VDS) on Student Participation



Inspection of an Explanatory Visualization (VDS) on Student Performance

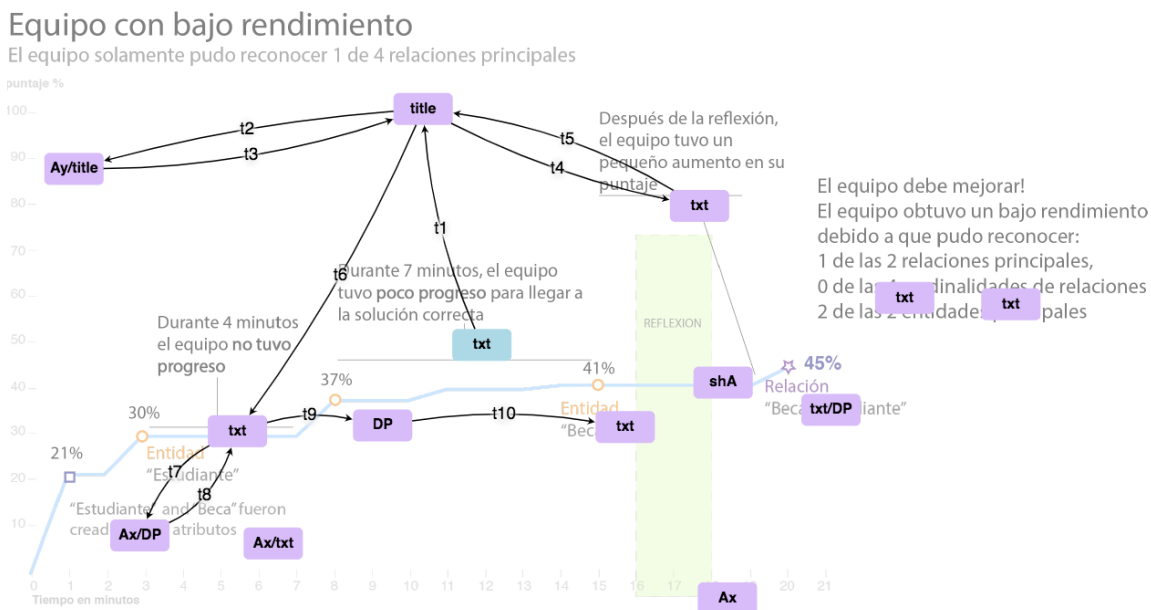


Figure 10. Two other example trajectories followed during the first 10 seconds of inspection of an OV (top) and a VDS (bottom) by two teachers (T5 and T6, respectively).

5.3. Usefulness of Visualizations and Teacher Preferences

In this section, we examine the usability of visualizations, especially those for explanatory purposes and their power in supporting teachers' monitoring for orchestration of the collaborative activity. We also explore which data storytelling elements are most helpful to support storytelling through learning visualizations.

5.3.1. Ease of Interpretation and Orchestration.

In the final interviews, the majority of teachers converged regarding the ease of interpretation that explanatory visualizations provided. Most also agreed that the VDSs helped them to articulate more compelling stories about what happened during student activity due to the context added through the data storytelling elements. This point of view was echoed by teacher T1 who stated the following: “These visualizations [with data storytelling elements] show me exactly what I need. I do not have to guess what happened in the activity.” Teacher T3 added that “this [with data storytelling elements] gives a summary of the whole graph. With the original visualization, I would have had to explore the whole visualization in order to understand the team’s performance”; and teacher T4 expressed the following: “the [data storytelling] visualization shows me what elements students have created very clearly, compared to the other [OV] visualization.”

Regarding orchestration, teachers also mentioned that adding contextual information to the explanatory visualizations helped them understand the behaviour of the group. In the visualizations about student performance, for example, associating text labels to key data points (i.e., the names of entities and relationships added by students) helped teachers understand the causes of potential unintended mistakes, so they could provide students with feedback if this visualization would be generated in real-time or to guide a data-informed post-hoc reflection. For the visualization about student participation, teacher T4 expressed that “the shaded area in the graph made [her] aware that groups took very different amounts of time to reflect and so, with this information, [she] would encourage students to reflect for a long time before submitting a final version of the task.” Another teacher (T2) mentioned that “the title showed the right message of the graph regarding the team’s participation, which made [him] think that this group needed to collaborate more equally.”

5.3.2. Teacher Preferences

During the interviews, we also asked teachers to select which visualization they would prefer and why, in order to explore design issues in detail. When analyzing the interviews, we found that teachers referred to two specific reasons when picking a visualization: the *functionality* and the *design* of the visualization. For the visualization on **participation** data, teachers argued that they would prefer the **VDS visualization**, because of its *functionality*. Teachers reported that VDS visualizations are informative, and the text is legible and easy to follow. Regarding the **performance** visualization, the majority of teachers also agreed they would prefer **VDS visualization**. They indicated that colours were helpful to highlight information (e.g., line and data points) and the summary provided facilitated the interpretation of the data very quickly. However, some critiques were made concerning the *design* of the visualizations with data storytelling elements. For the VDS about **participation**, teachers suggested that instead of having straight lines, they would prefer smoother lines (e.g., curves) because they are visually more appealing. Another criticism about the design was related to the balance of additional text labels, as the visualization could become complex. In fact, two teachers (T1 and T6) mentioned that text labels “may add noise and graph complexity” and that the text label information may be, on occasion, “redundant with no extra information.”

5.3.3. Data Storytelling Elements Helpfulness

Figure 11 shows the final rankings by the six teachers after inspecting VDS. According to teacher perceptions, they ranked the prescriptive **title** as a critical element to facilitate the sensemaking of both types of visualizations. Five of them agreed that the title served to indicate the main message from the visualizations before going deeply into inspecting the details from the visualizations. This was stated by two teachers, as follows: “[the title] shows an overall description of the graph, which helps to interpret faster the visualization” (T4) and that “[title] was important to summarize and understand the story” (T1).

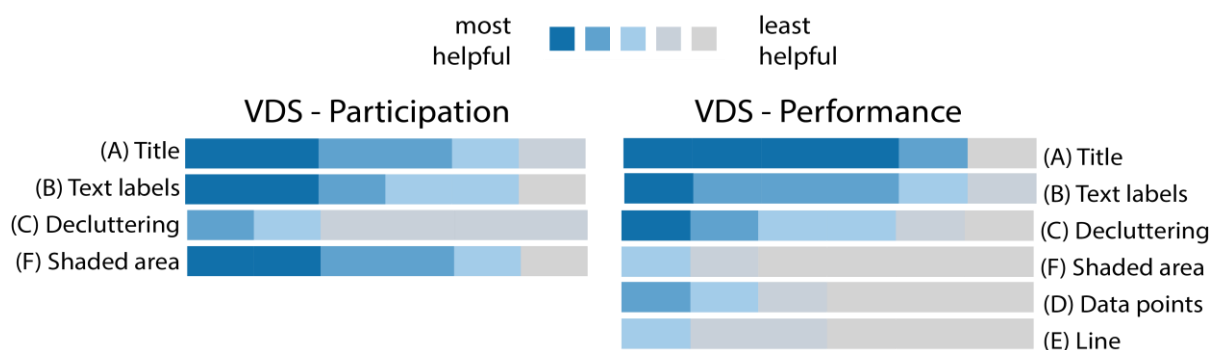


Figure 11. Ranking of data storytelling elements to support storytelling in visualizations.

Next in the ranking of data storytelling elements is the **shaded area**. Teachers made an explicit comment about how the shaded area helped them to interpret better the sudden drop of activity in the visualization: “[the shaded area] explains what students were doing. Without this, I would think that students were doing nothing related to the activity” (T2); “it adds context to the visualization, to know what happened during the activity” (T3); and “if I had to review this information after the class,

this would help me to remember what happened in that point” (T4). However, this element was ranked as the least helpful in the VDS about performance. Differences in both visualization results can be explained by the fact that to understand the extent to which students participated, the teacher needs to be more aware of the behaviour of the team (e.g., *why* there was a drop in the activity), whereas, for task performance, teachers did not perceive that it would help to understand the context.

Ranked in second and third place, for VDSs about performance and participation respectively, were **text labels** adding narrative to the visualizations. Teachers pointed out that text labels facilitated the narrative explanation of what occurred in different moments and ease of reading by describing what happened during the activity: “[the text labels] helped me to see what happened when the three students were collaborating (for visualization about participation). The first thing I read in the graph were the labels, which pointed me to explore in detail the data presented” (T2).

Decluttering was ranked in third place for the VDSs on **performance** (Figure 11, right). Teachers expressed that the information was presented more clearly than in the OV: “Removing the green background, grids and big circles, made easier to explore the information.” Teachers also indicated that because the original visualization (OV) was visually overloaded, it was convenient to declutter the graph. As a result, the visualization was easier to navigate. For example, one teacher (T4) stated that “it was very useful to declutter the graph. Now it is easier to identify relevant objects that were created over the time.” Another teacher (T5) mentioned that “decluttering the graph in this visualization was very relevant. I can now see exactly where the main object is that increased performance significantly.” However, the same element was the lowest in the ranking for VDS on participation (Figure 11, left). From the comments, teachers expressed that the grid was relevant to find data point values. One teacher (T3) suggested that “it [the OV visualization] was better with the grid because it is easier to find values during the inspection.” In the same argument, another teacher (T4) indicated that “you can get confused without the grid when exploring the visualization.” One explanation of this result is that, due to the higher number of data points and lines (one line per participant), backup information is needed to explain the data point values. In our crafted visualizations, we removed the grid and added textual information about equity of participation. We did not add textual information about the number of actions made by each participant, as our intended message was to prescribe the participation behaviour (balanced or unbalanced).

Notably, other elements that in principle are important for data storytelling (Knafllic, 2015) were lower ranked. These included, for example, the highlighted data points and using a thicker line to highlight activity. For instance, teachers mentioned that **data points** are useful for showing only relevant information and that the thicker **line** makes easier to focus the attention on the information. With this, we are not arguing that these elements should not be considered for further explorations in learning dashboards. We agree that teachers could perceive other elements as more helpful for them.

When triangulating this information with gaze behavioural data, we can say that just highlighting specific data points helped teachers to gain understanding about what occurred with the teams without consulting all the data points available. From the previous subsection, heatmap and trajectory analysis suggested that text labels and data points were the most visited elements. Interestingly, according to teacher perceptions, the title, text labels, and shaded areas were catalogued as the most helpful ones.

5.3.4. Study Summary and Limitations

Our study aimed to explore the effectiveness of data storytelling elements as a way to leverage the communicative power of visualization. To achieve this, we analyzed the following: 1) teachers’ reactions to the visualizations using a think-aloud protocol, 2) teachers’ gaze behaviour and focus of attention using eye-tracking technology to cluster gaze according to areas of interest (AOI); and 3) teachers’ preferences after the inspection of visualizations without (OV) and with (VDS) data storytelling elements. Summing up the analysis and results, we suggest the following:

- Data storytelling elements have the potential to help teachers explore visualizations with less effort. All teachers could interpret both the VDSs and OVs, but they explicitly expressed most affinity with VDS.
- Exploratory visualizations (OVs) prompted teachers to explore the data, describe it, and raise questions, whereas explanatory visualization (VDS) provoked more compelling comments about student activities. This suggests that data storytelling elements may be used as a way of guiding people to gain insights from visualizations without the help of a technician.
- Teacher behaviours point to the critical role that both narratives and visuals play in explaining the data. Text labels and key data points were the most relevant elements to draw attention to an explanatory visualization. We confirmed this result from the eye tracker data and teacher perceptions. However, as also suggested by teachers, visualizations should be carefully designed in order to avoid clutter with too many text labels.
- Other design implications should be taken into consideration when designing the final visualizations. Colours (smoother, straight) lines, grids, and legends are visual elements that may help users to understand the visualizations better. Whichever is the case, designers should be aware of the learning intention of the visualization.

- A limitation of this study is that, due to the low number of teachers, these results cannot be generalized in the way permitted by an experimental lab study with a statistically meaningful number of subjects. However, our methodology helped us to elicit rapid formative feedback from teachers to start understanding the impact that explanatory visualization may have on teacher perceptions and sense-making.
- Evidence from eye tracking data allowed us to gain a better understanding of the elements that drove teacher attention during the first minute of seeing a dashboard. By making use of eye-tracking technology, we could understand teachers' gaze behaviours when exploring these visualizations, with fine-grained detail on fixations, trajectories and most seen areas of interest. Future work in the area of visual learning analytics may consider using eye-tracking technology to understand perception aspects of the visual sensemaking process of teachers and students and also to consider the impact of design on the learning analytics process.

6. Conclusion and Future Directions

We have proposed a conceptual model for bringing data storytelling principles to support the interpretation of questions in teacher- and student-facing dashboards that can become explanatory tools. We aligned the learning design with data storytelling by extracting rules from the teacher's intentions, providing the basis for prioritizing which features to foreground. Moreover, we suggested how these rules can be applied in manually generated prototypes. We intend to add these data storytelling elements automatically, and to this end, a pilot study with teachers helped explore teacher interpretations of learning visualizations. The results of our study are encouraging and provide the basis to investigate further the means for crafting personalized and context-dependent dashboards. We see the potential of learning design-driven data storytelling elements to support sensemaking by guiding the user to "one learning story per visualization," given that learning is a complex task. We hope that both designers and researchers can gain value from this conceptual model as a step towards *dashboards explicitly contextualized to learning*. Future work will seek to deploy explanatory visualizations in authentic settings in order to perform a longitudinal, authentic study with students and teachers. As next steps, we will involve teachers and students in the design process of learning analytics dashboards. Finally, we intend to evaluate how teachers might adjust their learning activities based on the insights they might gain from dashboards whose design they have helped to shape.

Declaration of Conflicting Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) declared no financial support for the research, authorship, and/or publication of this article.

References

- Bach, B., Kerracher, N., Hall, K. W., Carpendale, S., Kennedy, J., & Henry Riche, N. (2016). Telling stories about dynamic networks with graph comics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '16)*, 7–12 May 2016, San Jose, CA, USA (pp. 3670–3682). New York: ACM. <http://dx.doi.org/10.1145/2858036.2858387>
- Bakharia, A., Corrin, L., de Barba, P., Kennedy, G., Gašević, D., Mulder, R., Williams, D., Dawson, S., & Lockyer, L. (2016). A conceptual framework linking learning design with learning analytics. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge (LAK '16)*, 25–29 April 2016, Edinburgh, UK (pp. 329–338). New York: ACM. <http://dx.doi.org/10.1145/2883851.2883944>
- Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). Open learner models and learning analytics dashboards: A systematic review. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18)*, 5–9 March 2018, Sydney, NSW, Australia (pp. 41–50). New York: ACM. <http://dx.doi.org/10.1145/3170358.3170409>
- Bodily, R., & Verbert, K. (2017). Trends and issues in student-facing learning analytics reporting systems research. *Proceedings of the 7th International Conference on Learning Analytics and Knowledge (LAK '17)*, 13–17 March 2017, Vancouver, BC, Canada (pp. 309–318). New York: ACM. <http://dx.doi.org/10.1145/3027385.3027403>
- Bryan, C., Ma, K.-L., & Woodring, J. (2017). Temporal summary images: An approach to narrative visualization via interactive annotation generation and placement. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 511–520. <http://dx.doi.org/10.1109/TVCG.2016.2598876>
- Corrin, L., & de Barba, P. (2015). How do students interpret feedback delivered via dashboards? *Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15)*, 16–20 March 2015, Poughkeepsie, NY,

- USA (pp. 430–431). New York: ACM. <http://dx.doi.org/10.1145/2723576.2723662>
- Duval, E. (2011). Attention please! Learning analytics for visualization and recommendation. In P. Long, G. Siemens, G. Conole, & D. Gašević (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK '11)*, 27 February–1 March 2011, Banff, AB, Canada (pp. 9–17). New York: ACM. <http://dx.doi.org/10.1145/2090116.2090118>
- Dykes, B. (2015). Data storytelling: What it is and how it can be used to effectively communicate analysis results. *Applied Marketing Analytics*, 1(4), 299–313.
- Echeverria, V., Martinez-Maldonado, R., Chiluiza, K., & Buckingham Shum, S. (2017). DBCollab: Automated feedback for face-to-face group database design. In W. Chen et al. (Eds.), *Proceedings of the 25th International Conference on Computers in Education (ICCE 2017)*, 4–8 December 2017, Christchurch, New Zealand (pp. 156–165). Asia-Pacific Society for Computers in Education.
- Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiluiza, K., Conati, C., & Buckingham Shum, S. (2018). Driving data storytelling from learning design. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18)*, 5–9 March 2018, Sydney, NSW, Australia (pp. 131–140). New York: ACM. <http://dx.doi.org/10.1145/3170358.3170380>
- Feng, M., Krumm, A. E., Bowers, A. J., & Podkul, T. (2016). Elaborating data intensive research methods through researcher–practitioner partnerships. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge (LAK '16)*, 25–29 April 2016, Edinburgh, UK (pp. 540–541). New York: ACM. <http://dx.doi.org/10.1145/2883851.2883908>
- Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, 15(3), 42–57.
- Holstein, K., McLaren, B. M., & Aleven, V. (2017). Intelligent tutors as teachers' aides: Exploring teacher needs for real-time analytics in blended classrooms. *Proceedings of the 7th International Conference on Learning Analytics and Knowledge (LAK '17)*, 13–17 March 2017, Vancouver, BC, Canada (pp. 257–266). New York: ACM. <http://dx.doi.org/10.1145/3027385.3027451>
- Iliinsky, N., & Steele, J. (2011). *Designing data visualizations: Representing informational relationships*. Sebastopol, CA: O'Reilly Media.
- Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. *Proceedings of the 12th European Conference on Technology Enhanced Learning (ECTEL 2017)*, 12–15 September 2017, Tallinn, Estonia (pp. 82–96). Lecture Notes in Computer Science, Springer. http://dx.doi.org/10.1007/978-3-319-66610-5_7
- Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). License to evaluate: Preparing learning analytics dashboards for educational practice. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18)*, 5–9 March 2018, Sydney, NSW, Australia (pp. 31–40). New York: ACM. <http://dx.doi.org/10.1145/3170358.3170421>
- Kim, S., Dey, A. K., Lee, J., & Forlizzi, J. (2011). Usability of car dashboard displays for elder drivers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 7–12 May 2011, Vancouver, BC, Canada (pp. 493–502). New York: ACM. <http://dx.doi.org/10.1145/1978942.1979011>
- Klerkx, J., Verbert, K., & Duval, E. (2017). Learning analytics dashboards. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *The handbook of learning analytics* (pp. 143–150). Society for Learning Analytics Research.
- Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. Hoboken, NJ: John Wiley & Sons.
- Knight, S., Shibani, A., & Buckingham Shum, S. (2018). Augmenting formative writing assessment with learning analytics: A design abstraction approach. In J. Kay & R. Luckin (Eds.), *Rethinking learning in the digital age: Making the Learning Sciences count. Proceedings of the 13th International Conference of the Learning Sciences (ICLS '18)*, 23–27 June 2018, London, UK (Vol. 3, pp. 1783–1790). International Society of the Learning Sciences.
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, 1(6), 90–95.
- Krumm, A., Means, B., & Bienkowski, M. (2018). *Learning analytics goes to school: A collaborative approach to improving education*. London, UK: Routledge.
- Lallé, S., Conati, C., & Carenini, G. (2016). Prediction of individual learning curves across information visualizations. *User Modeling and User-Adapted Interaction*, 26(4), 307–345. <http://dx.doi.org/10.1007/s11257-016-9179-5>
- Lee, B., Riche, N. H., Isenberg, P., & Carpendale, S. (2015). More than telling a story: Transforming data into visually shared stories. *IEEE Computer Graphics and Applications*, 35(5), 84–90. <http://dx.doi.org/10.1109/MCG.2015.99>
- Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning

- design. *American Behavioral Scientist*, 57(10), 1439–1459. <http://dx.doi.org/10.1177/0002764213479367>
- Ma, K. L., Liao, I., Frazier, J., Hauser, H., & Kostis, H. N. (2012). Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1), 12–19. <http://dx.doi.org/10.1109/MCG.2012.24>
- Marzouk, Z., Rakovic, M., Liaqat, A., Vytasek, J., Samadi, D., Stewart-Alonso, J., Ram, I., Woloshen, S., Winne, P. H., & Nesbit, J. C. (2016). What if learning analytics were based on learning science? *Australasian Journal of Educational Technology*, 32(6). <https://ajet.org.au/index.php/AJET/article/viewFile/3058/1433>
- Metoyer, R., Zhi, Q., Janczuk, B., & Scheirer, W. (2018). Coupling story to visualization: Using textual analysis as a bridge between data and interpretation. *Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI '18)*, 7–11 March 2018, Tokyo, Japan (pp. 503–507). New York: ACM. <http://dx.doi.org/10.1145/3172944.3173007>
- Molenaar, I., & Knoop-van Campen, C. (2017). Teacher dashboards in practice: Usage and impact. *Proceedings of the 12th European Conference on Technology Enhanced Learning (EC-TEL 2017)*, 12–15 September 2017, Tallinn, Estonia (pp. 125–138). Lecture Notes in Computer Science, Springer. http://dx.doi.org/10.1007/978-3-319-66610-5_10
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Reimers, G., & Neovesky, A. (2015). Student focused dashboards. *Proceedings of the 7th International Conference on Computer Supported Education (CSEDU 2015)*, 23–25 May 2015, Lisbon, Portugal (pp. 399–404). http://www.qu.tu-berlin.de/fileadmin/fg41/publications/reimers_2015_student-focused-dashboards--an-analysis-of-current-student-dashboards-and-what-students-really-want.pdf
- Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give me a customizable dashboard: Personalized learning analytics dashboards in higher education. *Technology, Knowledge and Learning*, 22(3), 317–333. <http://dx.doi.org/10.1007/s10758-017-9316-1>
- Rodríguez-Triana, M. J., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology*, 46(2), 330–343. <http://dx.doi.org/10.1111/bjet.12198>
- Ryan, L. (2016). *The visual imperative: Creating a visual culture of data discovery*. Amsterdam, Netherlands: Elsevier Science.
- Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, 10(1), 30–41. <http://dx.doi.org/10.1109/TLT.2016.2599522>
- Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), 1139–1148. <http://dx.doi.org/10.1109/TVCG.2010.179>
- Shibani, A., Knight, S., Buckingham Shum, S., & Ryan, P. (2017). Design and implementation of a pedagogic intervention using writing analytics. In W. Chen et al. (Eds.), *Proceedings of the 25th International Conference on Computers in Education (ICCE 2017)*, 4–8 December 2017, Christchurch, New Zealand (pp. 306–315). Asia-Pacific Society for Computers in Education.
- Stevens, A., Quimby, A., Board, A., Kersloot, T., & Burns, P. (2002). Design guidelines for safety of in-vehicle information systems. TRL Limited. <https://trl.co.uk/reports/PA3721/01>
- Teasley, S. D. (2017). Student facing dashboards: One size fits all? *Technology, Knowledge and Learning*, 22(3), 377–384. <http://dx.doi.org/10.1007/s10758-017-9314-3>
- Toker, D., Conati, C., Steichen, B., & Carenini, G. (2013). Individual user characteristics and information visualization: Connecting the dots through eye tracking. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 27 April–2 May 2013, Paris, France (pp. 295–304). New York: ACM. <http://dx.doi.org/10.1145/2470654.2470696>
- Tufte, E. R., & Schmiege, G. M. (1985). *The visual display of quantitative information* (Vol. 53). American Association of Physics Teachers.
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500–1509. <http://dx.doi.org/10.1177/0002764213479363>