visualizations can provide, one of them being the *process-oriented* feedback, in which the focus is on the "procedural aspects" of learning. This type of feedback is aligned with the concept of *process analytics*, defined by Lockyer, Heathcote, and Dawson (2013) as the investigation of the actual processes behind the "tasks students complete as part of a learning design" — a counterpoint to a more common type of learning analytics called *checkpoint analytics*, which is mostly concerned with the learning products/outcomes and quantitative engagement indicators (ibid.).

As reported by several authors (LOCKYER; HEATHCOTE; DAWSON, 2013; FARRELL, 2018; SEDRAKYAN et al., 2018; VIEIRA; PARSONS; BYRD, 2018; WISE, 2019), more research is needed on the design and evaluation of visual representations for process-oriented feedback. Vieira, Parsons & Byrd (2018), in a comprehensive literature review, found that only 21% of the reviewed works focus on process analytics, and none of them filled all the quality requirements proposed by the authors. We confirmed this scenario in our own literature review on the use of visualizations for learning process visualization (DOURADO et al., 2018a): few works deal with this topic, and even those have many limitations, such as lack of empirical validation — a recurrent problem in the field (FARRELL, 2018; VIEIRA; PARSONS; BYRD, 2018). Finally, Sedrakyan, Mannens & Verbert (2019) propose recommendations for choosing visual representations for each type of feedback, but also recognize the need for empirically validated research on this topic.

Considering the research gaps mentioned above, the goal of this work is to devise empirically validated recommendations for building process-oriented feedback visualizations for distance learning teachers. The remainder of this paper describes the method (Section 2), results (Section 3), and conclusions (Section 4) of our work.

2. METHOD

This work is framed as a *design study, that is,* the development of visualization guidelines for a given real-world problem through an iterative design process (SEDLMAIR; MEYER; MUNZNER, 2012). Our method is organized following the Design Activity Framework for Visualization Design (MCKENNA et al., 2014), which divides the visualization design process into four *design activities* — understand, ideate, make, and deploy — and defines for each one a goal, a set of suggested instruments, and the expected outcomes. However, the framework does not require that a project must conduct all four activities; only the order has to be maintained. Therefore, we conducted the first two design activities defined in the Design Activity Framework — understand and ideate — to devise design requirements and visual encodings for the domain under investigation. Figure 1 summarizes our method, describing the used instruments and expected outcomes for each design activity.

2.1 Understand activity

In the *understand* activity, the goal is to understand the problem domain, the target users and their needs (MCKENNA et al., 2014). As depicted in Figure 1, in this activity we used two instruments: literature review and ethnographic interviews. In the literature review, we analyzed related works on the use of process-oriented feedback visualizations (published as DOURADO et al. (2018a)), the learning theories that could inform our design process (published as DOURADO et al. (2018b)), and the Information Visualization techniques that could support the visualization of learning processes.

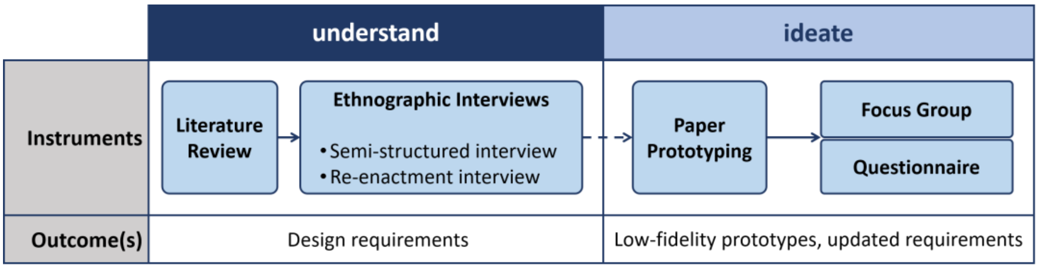


Figure 1 - Summary of our method, including instruments and outcomes.

To complement the literature reviews, we conducted two types of ethnographic interviews with a group of distance learning teachers: semi-structured interview (KUMAR, 2013) and re-enactment interview — a digital ethnography method in which users are asked to explain and demonstrate (re-enact) how they perform a certain task (PINK et al., 2016). These interviews were conducted by the first author of this paper with each teacher, individually, in this order: first the semi-structure interview and then the re-enactment. In the semi-structure interviews, teachers were asked to discuss three topics: i) their current practices regarding process feedback (resources, information, and strategies used, as well as challenges and problems); ii) the challenges of giving process feedback on distance education compared to face-to-face courses (all but one teacher had previous experience on both contexts); and iii) expectations regarding the use of VLE learning traces to get learning process information from students. In the re-enactment interviews, teachers were asked to demonstrate, using their own workstation, how they managed to give/obtain processual feedback to/from students in their daily practice.

Participated in these interviews 10 teachers (8F, 2M; experience with distance learning varying from 2 months to 7 years: average 2.69 years, sd 2.48 years) from the ETEPAC state school, a public institution in Pernambuco, Brazil, which offers free online vocational courses to students from all over the state through a Moodle-based virtual learning environment. The school's courses are grouped into five areas: Business & Management, Social and Educational Development, IT, Design, and Labor Safety. We interviewed teachers from all areas and, therefore, varied backgrounds[1]. The interviews were conducted between March 26th and April 8th, 2019, in the teacher's natural workplace: the school building where they work on 4-hour shifts every day. Both interviews were audio-recorded, and the re-enactment interviews were also photo- and video-recorded, as recommended by Pink et al. (2016). We transcribed the recordings and analyzed them using Ethnographic Content Analysis (ALTHEIDE, 1987) with a deductive open coding approach (CORBIN; STRAUSS, 2015) — using the categories defined by McKenna et al. (2014) — and the help of a QDA software[2].

2.2 Ideate activity

In the *ideate* activity, the goal is to generate and evaluate a set of ideas that addresses the design requirements identified in the *former* activity (*understand*). As depicted in Figure 1, we used two instruments in this activity: paper prototyping (MAGUIRE, 2001) and focus groups (MARTIN; HANINGTON, 2012). Through paper prototyping, we produced six low-fidelity prototypes, which will be presented and discussed in Section 3.2.

[1] The detailed profile of all teachers and other supplementary materials are available at http://osf.io/s6ybc/
[2] QDA Miner: https://provalisresearch.com/products/qualitative-data-analysis-software/

To evolve the prototypes and refine the design requirements, we conducted two focus groups[3] involving a total of 9 teachers (experience with distance learning varying from 7 months to 6 years: average 2.84 years, sd 1.7 years) from the same public school described in Subsection 2.1. In the first focus group, 5 teachers (2F, 3M) participated during an 1h20min session; in the second one, 4 teachers (3F, 1M) participated during an 1h session. Among the 9 participants, only 3 had also participated in the previous interviews conducted during the *understand* activity. The two sessions took place in August 22th, 2019, in one of the school's office rooms. They were conducted by the first author of this paper and audio- and photo-recorded by a colleague from the same research group. We provided the following materials to all participants: a) six colored paper prototypes; b) a perceived usefulness questionnaire with the question "This visualization can help me to better follow my students' learning process", to be answered for all six prototypes using a scale ranging from "1-Completely disagree" to "5 - Completely agree"; and c) creativity toolkits (MARTIN; HANINGTON, 2012) containing pens, black and colored pencils, paper, scissors, glue, and markers, so teachers could modify the prototypes to better fulfill their needs. We started each focus group with a brief explanation about Visual Learning Analytics, then explained the perceived usefulness questionnaire and, for each prototype, we followed this procedure: 1) explained the prototype's goals and features; 2) asked trigger questions to foster the discussion ("How do you think it could be improved?", "How would you use it in your work routine?", "What are the advantages/disadvantages?"); 3) after the discussion, we invited teachers to modify the prototypes, using the creativity toolkits, to incorporate new features or improvements; and finally 5) asked teachers to answer the questionnaire for the prototype under discussion.

We transcribed the audio recordings of the two focus groups and analyzed the transcriptions using thematic analysis (BRAUN; CLARKE, 2012) with deductive open coding — as in the *understand* activity — and the aid of the QDA Miner software. We integrated the interventions made by teachers on the prototypes into the transcriptions.

3. RESULTS

We present our results according to the outcomes expected from each design activity, as defined by McKenna et al. (2014) and illustrated in Figure 1.

3.1. Design Requirements

As shown in Figure 1, both the *understand* and *ideate* activities contribute to the identification of design requirements. Therefore, the results described in this subsection were derived from the analysis of the ethnographic interviews and the focus groups. We characterize the design requirements in the next subsections through five classes: opportunities, constraints, considerations, data abstraction, and task abstraction, as proposed by McKenna et al. (2014).

3.1.1. Opportunities, constraints, and considerations

Opportunities. The main complaint among the interviewed teachers — and the thing they missed the most from face-to-face classes — was the low level of feedback they can get *from* students on the VLE, as put by T1[4]: "We don't have feedback from

[3] By request of the school, we split the 9 participants in two groups to minimize the impact on their work.
[4] We use the notation T*n* for citing teachers. Please refer to the supplementary materials for more details.

them […] we need some feedback from them so we can give them feedback and, in this way, walk together”. This issue limits the interaction between teacher and students — confirming the cyclic nature of feedback described by Sedrakyann et al. (2018) — and forces teachers to play a “reactive role” in the course, as put by T4: “if the student has some difficulty and does not manifest it in some way, we will never know […] the teacher does not have an active role in the virtual environment”. As a result of the lack of process feedback features in the VLE, teachers improvise (PINK et al., 2016) by establishing a *checking routine* — sometimes organized by the program coordinators, sometimes by the teachers themselves — where they use several instruments to get cues about the learning process of their students. The most cited instruments used to this purpose were, in this order: discussion boards, messages sent/received through the LMS messaging subsystem, assignments, and some simple reports offered by the VLE — a chart describing access per day (for the whole class or by student), a textual report with the log of activities for each student, and a student grades report.

Constraints. The biggest constraint is the large number of students per class: some courses have up to 3,000 students, and each teacher in the school is responsible for up to 800 students, making it hard to follow the students’ progress in a detailed way. Given that all assignments are graded by teachers (MOOC-like strategies like peer-review are not used) and teachers perform multiple functions besides grading (production of learning materials such as texts and video lectures, etc.), they have little time left for extra activities. Also — and in part because of the lack of time — some teachers perceived a tool designed for process analytics as more suited for management-level people than to themselves.

Considerations. When expressing expectations from the use of VLE logs for getting process-oriented feedback, the words “investigate” and “synthetize” were the most frequent ones in teachers’ speeches.

3.1.2. Data abstraction

As originally proposed by Munzner (2009) and reframed by McKenna et al. (2014) as one of the dimensions in the *opportunities* class, *data abstraction* means translating domain-specific data into abstract data types, using the computer science vocabulary. In Table 1, we list the most frequent data types — or, in the Learning Analytics vocabulary, *learning proxies* (WISE, 2019) — identified in the ethnographic interviews and map them to the correspondent abstract data types. The column “Mentions” represents how many times the proxy was mentioned during the interviews.

Generally speaking, all proxies listed in Table 1 fit into the abstract category of time-oriented data (AIGNER et al., 2011) or, more specifically, Temporal Event Sequence data (PLAISANT; SHNEIDERMAN, 2016; DU et al., 2017). To map the proxies to abstract data types, we used two taxonomies: Plaisant and Shneiderman’s (2016) taxonomy for describing event data characteristics (“Data characteristics” column) and Munzner’s (2014) taxonomy for describing data structures in information visualization (“Data structure” column[5]). Regarding the data characteristics, all data types except #4 and #5 are both point-based (e.g. “the student downloaded a file”) and interval-based events (e.g. “the student spent 15 minutes on the forum”); types #4 and #5 are only point-based. This information is not shown in Table 1 to save space.

[5] Although not shown on the table to save space, within the “Data structure” taxonomy the dataset type for all items is “table”, except #2-B, which is “tree”.

Table 1 – Data mapping: from domain types to abstract types. For the sake of brevity, we used a set of codes, which are explained on a bottom line in the table.

#	Domain data type	Mentions	Abstract data type: Data characteristics: Time*	Outcome†	Uncertainty	Rep. within record‡	Events per Record§	Data structure
1	*Assiduity*: access to the VLE over time or estimated time spent on it	8	A	-	Y	H	H	**Attributes****: timestamp/session duration [O-Q], device [C], location [C]
2	*A - Access pattern to learning materials* **(single student)**	7	R	G C	N	H	H	**Attributes**‡‡: access order [O-O], material context information (type, title, etc.) [C]
	B - Access pattern to learning materials **(group/class)**		R	G C	Y	H	H	**Nodes**: material type [C] **Links**: access order [O-O] **Attributes****: support (number of students that followed the path) [O-Q]
3	*Assignments:* visualization, attempts, and handouts over time	7	A	G C	Y	M	M	**Attributes****: timestamp [O-Q], event type (vis., access, handout) [C], grade [O-Q]
4	*Questions asked* (usually in discussion boards, but also through direct messages)	6	A	-	N	M	M	**Attributes****: timestamp [O-Q], medium (forum, direct message) [C], context (assignment, lecture, forum, etc.) [C]
5	*Progress in relation to course schedule or milestones*	4	R	A M	Y	N	L	**Attributes****: timestamp [O-Q], progress measure [O-Q]
6	*Overall student trajectory in the course*	3	A	G C	N	H	H	**Attributes****: timestamp [O-Q], event type [C], context (forum, assignment, etc.) [C]

* **Time representation strategy:** [A]bsolute [R]elative
† **Related outcome types:** [G]rades, [C]ourse completion, [A]ssignment handout, [M]ilestone completed
‡ **Repetition within record**: [N]o, [H]igh, [M]edium, [L]ow
§ **Events per record**: [H]igh, [M]edium, [L]ow
** **Attribute types**: [O-O] ordered-ordinal [O-Q] ordered-quantitative [C] categorical

As shown in Table 1, some data types may involve uncertainty. On types #1, #2-B, and #3 uncertainty arises when session duration or resource usage periods are estimated from point-based data — which is the case in the logs of VLEs like Moodle. In these cases, it is not always easy to identify "idle" time, such as when the student has opened a VLE page and, as put by T11, "left for a coffee" or is "playing a game on another tab [in the browser]" (T5). In data type #2, the uncertainty comes from the granularity transformation (student to class/group), and in #5 from the fact that VLEs only capture part of the students' learning experience, which can be mitigated by the use of additional data sources (social media, multimodal data, etc.).

3.1.3. Task abstraction

In the *task abstraction* class, Munzner (2009) and McKenna et al. (2014) propose to map user domain tasks to abstract visualization tasks. In Table 2, we present this mapping by describing the domain tasks and subtasks identified during the *understand*

and *ideate* activities and their corresponding abstract tasks, using Plaisant and Shneiderman's (2016) taxonomy. The column "Mentions" represents how many times the task was mentioned during the *understand* and/or *ideate* interviews.

Table 2 – Mapping of domain tasks to abstract event sequence visualization tasks.

#	Domain tasks		Source*	Mentions	Abstract task(s)
	Task	**Subtasks**			
1	Evaluate instructional design	A. Plan/adapt courses	I	7	T2 - Compile descriptive information about the dataset or a subgroup of records and events
		B. Evaluate the quality of learning materials through interaction pattern analysis	I	5	
		C. Investigate the use of learning materials	U	3	
2	Investigate a single student learning process/trajectory in detail	A. Assessment: review/reconsider grades or assess effort	U/I	6/7	T1 - Review in detail a few records
		B. Visualize/Explore the trajectory	U/I	4/4	
		C. Verify student complaints about technical difficulties in the VLE	U/I	1/3	
		D. Get to know better the student	I	2	
		E. Better answer student's questions	I	1	
		F. Identify improvement needs	U	1	
3	Identify patterns	A. Advise students based on "successful" or "unsuccessful" learning strategies	I	6	T7 – Study antecedents or sequelae of an event of interest T2 - Compile descriptive information about the dataset or a subgroup of records and events
		B. Get insights on how to improve the activities' order in the course	I	1	
		C. Develop strategies to prevent evasion.	U/I	1/1	
4	Compare the progress of a student/group of students against expected goals/milestones		U	4	T3 - Find and describe deviations from required or expected patterns
5	Find struggling (or "idle") students		I	1	T5 - Identify a set of records of interest

* ***Source***: *[U]nderstand activity, [I]deate activity*

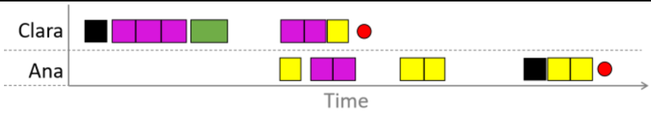
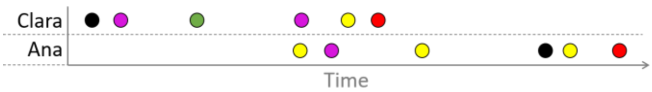
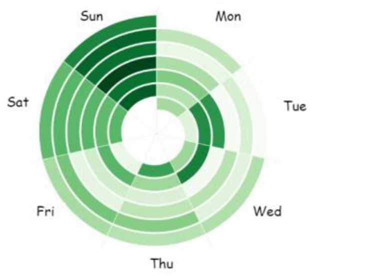
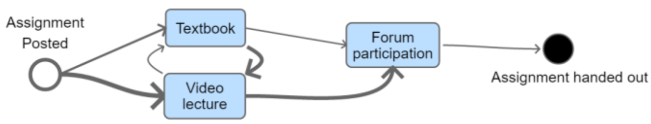
Tasks #1 and #3 corroborate with the work of Lockyer, Heathcote, and Dawson (2013), where the authors propose the use of process analytics for analyzing the effectiveness of a course's learning design. Teacher T15 exemplifies task #1-A saying that process-oriented visualizations could "help plan the course as taught in the teacher training programs […] by using the best sequence of activities". Regarding task #3-A, some teachers envisioned using snapshots of the visualizations as evidence to advise students on successful or unsuccessful learning strategies, as put by T12: "Then you can say [to students] like 'look, the students who are getting good grades are doing this…'". Tasks #2, #4, and #5 are in line with the idea of "class orchestration" (WISE, 2019), that is, when teachers use Learning Analytics tools to provide regular feedback to students and identify struggling learners or groups.

The design requirements described in this subsection informed the development of visualizations for process-oriented feedback, presented in the next subsection.

3.2. Visual encoding: prototypes design and evaluation

The six prototypes designed during the *ideate* activity are presented in Table 3 alongside the metrics and tasks each of them relates to. Note that not all tasks described in Table 2 are covered by the prototypes, which were built based only on the analysis of the ethnographic interviews conducted during the *understand* activity. The task list presented in Table 2 is the final, refined list, compiled using the results from the focus groups conducted during the *ideate* activity.

Table 3 – Visual encoding metaphors and simplified prototypes.

#	Metaphor	Metrics	Tasks	Simplified prototype[6]
P1	Timeline – Interval-based	All	1,2, 4,5	
P2	Timeline – Point-based	All*	1,2, 4,5	
P3	Timeline – Gantt chart	All	2,4	
P4	Hierarquical – Solar Plot	2b, 6	3	P4  P6 
P6	Spiral Graph	1, 2	3	
P5	State diagram	2, 3, 4, 6	1,2, 3	

* *Except interval-based events*

Prototypes P1, P2, and P3 use the timeline metaphor (AIGNER et al., 2011, p. 166) to represent learning events over time. P1 shows both point and interval data types, P2 is a variation where no interval (time spent) assumptions are made, and P3 a variation where, instead of multiple students, only one student is shown — a Gantt diagram (AIGNER et al., 2011, p. 167). Prototype P4 uses a solar plot (AIGNER et al., 2011, p. 182) to show an aggregate view of the most common learning paths followed by students. Prototype P6 uses a spiral graph (AIGNER et al., 2011, p. 185) to help teachers identify seasonal patterns on the use of learning materials; the example shown on Table 3 uses the *week* granularity. Finally, prototype P5 uses a simple state diagram to represent either student or class learning trajectories, signaling the most common transitions by the link thickness. All prototypes were developed using synthetic data.

Figure 2 shows the results of the perceived usefulness questionnaire applied in the *ideate* activity to evaluate the prototypes. Prototypes P1, P3, P5, and P6 received generally higher scores. Teacher T14 considered P1 "visually easy, you look and you

[6] Due to space restrictions, we show simplified prototypes. See the original ones at http://osf.io/s6ybc/

understand it quickly". Also, teachers were able to extract insights from the prototypes. While discussing P5, T11 concluded that "the student followed the pattern 'consult the textbook after watching the video lecture' more times than the opposite path" and T15 said that it would be useful for checking whether "the student is following the course the way we planned". Such insights can help teachers in re-evaluating which topics should be addressed on each learning resource and also better plan the order of activities in the learning design, as discussed by Lockyer, Heathcote, and Dawson (2013). While discussing P6, T14 noted that the visualization could help detect a situation where "only on the day before the exam they watched the video lecture"; this information could help teachers to timely alert students about bad learning strategies and, as a result, promote the co-regulation of learning, as suggested by Sedrakyan, Mannens & Verbert (2019).

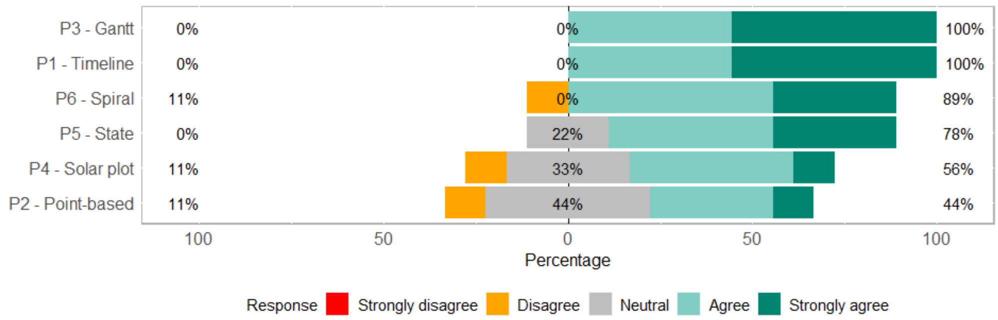


Figure 2 – Results of the perceived usefulness questionnaire for the prototypes.

Finally, teachers suggested several improvements to the prototypes during the focus groups: 1) correlate learning paths to learning outcomes, to make the visualizations actionable (especially P4); 2) use a combination of icons and colors to help differentiate event types on P1, P2, P3, and P5; 3) allow the filtering of learning paths by performance level to help identify successful/unsuccessful paths; and 4) provide a "playback" feature to make changes over time visible, especially on P5, in which users missed the ability to visualize time in a linear way.

4. CONCLUSION

In this paper, we propose design requirements and visual encodings for teacher-targeted process-oriented feedback in distance learning, based on empirical evidence collected through a design-based method. According to our results, the following design guidelines can be derived: i) provide granularity level control and filters, especially when the number of students per class is high; ii) present process-oriented data in context (e.g. relate to course week, activity, etc.); iii) provide checkpoint analytics visualizations as entry-point for process analytics visualizations; iv) inform teachers how the variables were chosen and pre-processed, including the assumptions made during the process and its limitations, to foster teachers' trust on the visualizations; and v) correlate process data with outcome data, such as assessment grades, to facilitate pattern detection and make visualizations actionable. As future work, we envision the development and evaluation of high-fidelity prototypes to generate more comprehensive guidelines, including aspects such as interaction and usage patterns.

ACKNOWLEDGEMENT

Raphael A. Dourado holds a CNPq PhD scholarship (#140973/2017-6). Alex Sandro Gomes is a CNPq DTI-2 researcher. This work was partly supported by CAPES-PrInt grant #88887.363990/2019-00. We thank the state of Pernambuco

Professional Education Secretariat (SEIP/SEE-PE), especially Mr. George Bento, Mrs. Renata Otero, and Mrs. Deisiane Bazante for their invaluable support in the field experiments, and also the 16 teachers that voluntary and enthusiastically collaborated to this work.

REFERENCES

AIGNER, W. et al. **Visualization of Time-Oriented Data**. London: Springer London, 2011.

ALTHEIDE, D. L. Reflections: Ethnographic Content Analysis. **Qualitative Sociology**, v. 10, n. 1, p. 65–77, 1987.

BRAUN, V.; CLARKE, V. Thematic analysis. In: **APA handbook of research methods in psychology, Vol 2**. APA handbooks in psychology®. Washington, DC, US: APA, 2012. p. 57–71.

CORBIN, J. M.; STRAUSS, A. L. **Basics of qualitative research: techniques and procedures for developing grounded theory**. Fourth edition ed. Los Angeles: SAGE, 2015.

DOURADO, R. et al. Mapeamento Sistemático sobre o uso de Visualização de Dados para Análise Processual da Aprendizagem em Ambientes Virtuais. In: **XXIX SBIE/CBIE**. Fortaleza, Ceará, Brasil: 28 out. 2018a.

DOURADO, R. et al. Analisando a Participação e a Aprendizagem em Comunidades Virtuais a partir da Teoria de Barbara Rogoff. **RENOTE**, v. 16, n. 1, 21 ago. 2018b.

DU, F. et al. Coping with Volume and Variety in Temporal Event Sequences: Strategies for Sharpening Analytic Focus. **IEEE TVCG**, v. 23, n. 6, p. 1636–1649, 2017.

FARRELL, T. **Affordances of Learning Analytics for Mediating Learning**. 2018. The Open University, 2018. Available at: <http://oro.open.ac.uk/57621/>. Accessed in: 23 May 2020.

KUMAR, V. **101 design methods**. Hoboken, N.J: Wiley, 2013.

LOCKYER, L.; HEATHCOTE, E.; DAWSON, S. Informing Pedagogical Action: Aligning Learning Analytics With Learning Design. **American Behavioral Scientist**, v. 57, n. 10, p. 1439–1459, 2013.

MAGUIRE, M. Methods to support human-centred design. **International Journal of Human-Computer Studies**, v. 55, n. 4, p. 587–634, 2001.

MARTIN, B.; HANINGTON, B. M. **Universal methods of design**. Digital ed. Beverly, MA: Rockport, 2012.

MCKENNA, S. et al. Design Activity Framework for Visualization Design. **IEEE Transactions on Visualization and Computer Graphics**, v. 20, n. 12, p. 2191–2200, 2014.

MUNZNER, T. A Nested Model for Visualization Design and Validation. **IEEE Transactions on Visualization and Computer Graphics**, v. 15, n. 6, p. 921–928, 2009.

MUNZNER, T. **Visualization analysis & design**. New York: AK Peters/CRC Press, 2014.

PINK, S. et al. (ed.). **Digital ethnography: principles and practice**. Los Angeles: SAGE, 2016.

PLAISANT, C.; SHNEIDERMAN, B. The diversity of data and tasks in event analytics. In: **Proceedings of the IEEE VIS 2016 Workshop on Temporal & Sequential Event Analysis**, 2016.

SEDLMAIR, M.; MEYER, M.; MUNZNER, T. Design Study Methodology: Reflections from the Trenches and the Stacks. **IEEE TVCG**, v. 18, n. 12, p. 2431–2440, 2012.

SEDRAKYAN, G. et al. Linking Learning Behavior Analytics and Learning Science Concepts: Designing a Learning Analytics Dashboard for Feedback to Support Learning Regulation. **Computers in Human Behavior**, 2018.

SEDRAKYAN, G.; MANNENS, E.; VERBERT, K. Guiding the choice of learning dashboard visualizations: Linking dashboard design and data visualization concepts. **Journal of Computer Languages**, v. 50, p. 19–38, 2019.

VIEIRA, C.; PARSONS, P.; BYRD, V. Visual learning analytics of educational data: A systematic literature review and research agenda. **Computers & Education**, v. 122, p. 119–135, 2018.

WISE, A. F. Learning Analytics: Using Data-Informed Decision-Making to Improve Teaching and Learning. In: ADESOPE, O. O.; RUD, A. G. (Ed.). **Contemporary Technologies in Education**. Springer, p. 119–143, 2019.