An Eye-Tracking Study of Notational, Informational, and Emotional Aspects of Learning Analytics Representations

Ravi Vatrapu1, 2, Peter Reimann3, Susan Bull4, and Matthew Johnson4

1Computational Social Science Laboratory (CSSL), ITM, Copenhagen Business School, Denmark
2Norwegian School of Information Technology (NITH), Norway
3MTO Psychologische Forschung und Beratung, Germany
4Electronic, Electrical and Computer Engineering, University of Birmingham, United Kingdom

vatrapu@cbs.dk, preimann.undef@gmail.com, s.bull@bham.ac.uk, johnsmdy@bham.ac.uk

ABSTRACT
This paper presents an eye-tracking study of notational, informational, and emotional aspects of nine different notational systems (Skill Meters, Smilies, Traffic Lights, Topic Boxes, Collective Histograms, Word Clouds, Textual Descriptors, Table, and Matrix) and three different information states (Weak, Average, & Strong) used to represent student’s learning. Findings from the eye-tracking study show that higher emotional activation was observed for the metaphorical notations of traffic lights and smilies and collective representations. Mean view time was higher for representations of the “average” informational learning state. Qualitative data analysis of the think-aloud comments and post-study interview show that student participants reflected on the meaning-making opportunities and action-taking possibilities afforded by the representations. Implications for the design and evaluation of learning analytics representations and discourse environments are discussed.

ACM Classification Keywords

Author Keywords
Learning analytics, teaching analytics, computer supported collaborative learning (CSCL), open learner models representational guidance, affordances

INTRODUCTION
One of the core concerns of research in learning analytics, teaching analytics, and open learner models (OLM) is the design, development, and evaluation of methods and tools for visualizing students’ and teachers’ learning and teaching processes and products. Learning analytics “is the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning.” Teaching analytics [23-25] as a sub-field of learning analytics is concerned with the design, development, and evaluation of visual analytics methods and tools for teachers’ pedagogical decision-making. An "open learner model" is a learner model that can also be externalised to the user [4]. This externalised (open) learner model may be simple or complex in format using, for example: text, skill meters, concept maps, hierarchical structures, animations [3]. Since learners and teachers perceive and act upon representations of their learning and teaching in these systems, conceptual and empirical attention needs to be devoted to notational, emotional, informational and interactive aspects of the underlying representations.

This paper presents the study of notational, emotional, and informational aspects of nine different kinds of learning analytics representations Skill Meters, Smilies, Traffic Lights, Topic Boxes, Collective Histograms, Word Clouds, Textual Descriptors, Table, and Matrix). The representations are referred to as “Static Representations” as they only present snapshot views of three knowledge states (Weak, Average and Strong) without offering any interactive capabilities to the student participants.

The remainder of the paper is organized as follows. The Theoretical Framework section to follow presents and discusses three lines of conceptual and empirical work relevant to the design and evaluation of representations. Methodology section presents details on the experimental study design, participant recruitment, sampling and

1http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics
assignment, materials, task, and the protocol. The section on Results presents the eye-tracking findings and qualitative observations from the exit interviews. Substantive interpretations and implications for the design of learning analytics representations are reported in the Discussion section.

THEORETICAL FRAMEWORK

Representation as a proxy to information plays a crucial role in design in general. The nature of representations, their structures and interactions is one of the central concerns of cognitive science [26]. Philosophically speaking, the function of representation is to “re-present”. Representation, in the philosophy of mind sense of the term, “is something that stands for something else”. Representations employed in learning analytics, teaching analytics, and open learner models “re-present” the ongoing learning processes and artifacts of the individual student and/or group of students. The technological, interactional, social and pedagogical aspects of representations have received significant conceptual and empirical attention in the fields of human computer interaction and learning sciences. Three following lines of conceptual and empirical work are particularly relevant for the design, development, and evaluation of representations in learning analytics, teaching analytics, and open learner models.

- Perception and Appropriation of Socio-Technical Affordances [21, 22]
- Representational Guidance [16, 19, 20]
- Cognitive Dimensions of Notations [1, 10]

Perception and Appropriation of Socio-Technical Affordances

The notion of affordance was introduced by J. J. Gibson [9]. Gibson was primarily concerned with providing an ecologically grounded explanation to visual perception. By drawing upon ecological psychology research on affordances, Vatrapu [21] defined a socio-technical affordance as

“action-taking possibilities and meaning-making opportunities in a socio-technical system relative to actor competencies and system capabilities.”

With regard to learning analytics representations, Perception of Affordances (PoA) refers to the action-taking possibilities and meaning-making opportunities that become available (that is, perceivable) to students in a given situation. Appropriation of Affordances (AoA) refers to the intentional utilization of the action-taking possibilities. AoA refers to the enactment of an interactional practice (generative, creative, or transformative). The eye-tracking experimental study reported here focused on the Perception of Affordances aspect of the phenomena. As such, the key issues of study were to what extent were the nine different kinds of OLM representations meaningful and actionable to the students.

Representational Guidance

The central premise of the representational guidance line of work is articulated by Suthers (2001) as:

The major hypothesis of this work is that variation in features of representational tools used by learners working in small groups can have a significant effect on the learners’ knowledge-building discourse and on learning outcomes. The claim is not merely that learners will talk about features of the software tool being used. Rather, with proper design of representational tools, this effect will be observable in terms of learners’ talk about and use of subject matter concepts and skills.

The above hypothesis follows from two lines of reasoning. First, the guiding ontological dimensions of representations—constraint and salience—prompt a user for what is missing as well for what is present [18]. The ontological dimensions of representations are not intrinsically social. Second, external representations play a role in guiding collaborative learning by amplifying certain kinds of social interactions [19] and knowledge building interactions [20].

Definition of Representational Guidance

“Representational guidance” refers to how these software environments facilitate the expression and inspection of different kinds of information.

[17]

Figure 1: Schematic of Representation Guidance

Figure 1 above, taken from Suthers [18] indicates that representational guidance has tripartite origins in the (a) affordances of a representational notation, (b) in how that notation is realized in a representational tool such as software, and (c) in the actual configuration of representational artifacts created by users of that tool.
How representational notations (such as Smilies, Word Clouds, and Traffic Lights etc.) are realized in the software and the actual configuration of representational artifacts are issues of concern for the design and evaluation of learning analytics systems. The eye-tracking study reported here primarily deals with the affordances (meaning-making opportunities and action-taking possibilities) of the nine different kinds of representational notations for learning analytics systems in three different information states.

Cognitive Dimensions of Notations

The cognitive dimensions framework [1, 10] is relevant to understanding the notational aspects of the learning analytics systems as it provides insights into the ontological characteristics of notations and their potential pedagogical implications. Further, Gibson's ecological optics [9] and Green and Blackwell's cognitive dimensions [1] share conceptual terms such as medium and environment. He next section presents key concepts in the cognitive dimensions framework and discusses their relevance to learning analytics systems. The following definitions are taken from Green and Blackwell [11]:

- **Information Artefacts**: "the tools we use to store, manipulate, and display information" (p.5)

Information artefacts are further classified as “non-interactive artefacts” and “interactive artefacts”. Learning analytics representations are information artefacts and the static representations studied here are an example of “non-interactive artefacts”.

- **Environment**: "The environment contains the operations or tools for manipulating those marks" (p.8). The environments in the learning analytics systems are the various “dashboards” designed for the different stakeholders.

- **Medium**: "The notation is imposed upon a medium, which may be persistent, like paper, or evanescent, like sound" (p.8). In the case of learning analytics systems, the medium is persistent and dynamically changed.
  - advance" (p.10)

Students and teachers engage in all four kinds of user activities listed above with learning analytics systems. The nine different kinds of representations have been selected with *Exploratory Design* in mind with the student participants engaging in all four activities as detailed in the Methodology section.

**Definitions of Cognitive Dimensions**

- **Abstraction**: "An abstraction is a class of entities, or a grouping of elements to be treated as one entity, either to lower the viscosity or to make the notation more like the user’s conceptual structure" (p.24)

- **Closeness of Mapping**: "Closeness of representation to domain" (p.39)

- **Consistency**: "similar semantics are expressed in similar syntactic forms" (p.39)

- **Diffuseness**: "verbosity of language" (p.39)

- **Error-Proneness**: "notation invites mistakes" (p.40)

- **Hard Mental Operations**: "high demand on cognitive resources" (p.40)

- **Hidden Dependencies**: "A hidden dependency is a relationship between two components such that one of them is dependent on the other, but that the dependency is not fully visible" (p.17)

- **Premature Commitment**: "Constraints on the order of doing things force the user to make a decision before the proper information is available" (p.21)

- **Progressive Evaluation**: "work-to-date can be checked at any time" (p.40)

- ** Provisionality**: "degree of commitment to actions or marks" (p.41)

- **Role-Expressiveness**: "the purpose of a component (or an action or a symbol) is readily inferred" (p.41)

- **Secondary Notation**: "Extra information carried by other means than the official syntax" (p.29)

- **Viscosity**: "Resistance to change: the cost of making small changes" (p.12)

- **Visibility**: "ability to view components easily." (p.34)

- **Juxtaposability**: "ability to place any two components side by side" (p.34)

The nine different kinds of learning analytics representations selected and evaluated for this study embody and exemplify at least one of the above listed cognitive dimensions of notations. For example, the Textual Representations exemplify verbosity of language (Diffuseness), Traffic Lights and Smilies are high on Abstraction but also high on Role-Expressiveness, Skillmeters provide Progressive Evaluation, Horizontal and Vertical Tables offer Closeness of Mapping to the domain of formative assessment, Collective Bar Chart representation embodies the Juxtaposition dimension.

**METHODODOLOGY**

**Experimental Design**

Given the central role that motivation plays in learning [6, 7] and the use of representations for formative assessment practices, emotional impacts of representations are of research interest with implications for practice. The exploratory research question is stated below:

RQ1: What, if any, are the relationships between notational, informational, and emotional aspects of
different kinds of representations of learning and teaching processes and products?

To empirically answer this exploratory research question, we designed a controlled laboratory study of a selected set of nine different kinds of learning analytics representations with varying notational aspects but isomorphic informational aspects. It is to be noted that the nine different notational systems selected for the study are not an exhaustive list. The research objectives were to investigate how students perceived different kinds of representations and what, if any, are the differences in emotional arousals between them. The nine different kinds of representations are presented in the materials section below.

**Materials**

As mentioned earlier, nine different kinds of notational systems (Skill Meters, Smilies, Traffic Lights, Topic Boxes, Collective Histograms, Word Clouds, Textual Descriptors, Table, and Matrix) were used to generate three informational states (Weak, Average, & Strong) of the learning analytics representations. For each of the nine notational systems, three isomorphic representations were created to embody the three informational states of the individual’s learning state. Thus, a total of twenty seven (9 notational systems x 3 informational states = 27) different static representations were developed. Further, we developed domain-specific (business and engineering) and domain-generic (nondescript) versions of the 27 representations. The domain-specific representations for Business subject area were used as the study sample consisted of international undergraduate business students at the Copenhagen Business School in Denmark. The 27 representations are presented below.

**Skill Meters (Weak, Average, Strong)**

**Smilies**

**Traffic Lights**

**Topic Boxes**

**Collective Histograms**
Participants
The sampling frame for study was the student population of the International Summer University Program (ISUP) 2011 the Copenhagen Business School, Denmark. Study recruitment was done by an email solicitation sent by the program secretariat to all enrolled students. Participants expressed their interest and indicated their availability on the study registration form. A total of 15 students participated in the study. The gender composition was 6 males and 7 females.

Tasks
Participants were instructed to talk aloud about what sense they can make of the representations displayed on the screen and their subjective preferences of the nine different notational systems while playing one of the four roles (themselves, a close friend taking the same course, a classmate who is not a close friend, and the teacher of the course).

3.5 Procedure

Pre-Investigative Session
- Participant was welcomed and seated in the laboratory. They were reminded that they are about to participate in an eye-tracking study and a short interview will be conducted after the study.
- An informed consent form was given. Participants were explicitly informed that it is the software that is being tested and not them. Further, participants were explicitly informed that they may withdraw consent at any time during the study session. Participants were informed that they would still be compensated with a movie ticket coupon in case they stop participating before completing the study session.
- A copy of the informed consent form was given to the participants after they signed it. An anonymized session code was then assigned.
- This concluded the Pre-Investigative Session

Investigative Session
- The participant was seated in front of the eye-tracker and the position was adjusted so that participant eyes were visible in the eye-finder of SMI iView X eye-tracker device\(^2\) driver software and iMotions Attention Tool 4.1\(^3\) study software
- 9-point eye calibration was then conducted followed by light calibration

\(^3\)http://www.imotionsglobal.com/
• The study session consisted of a randomized presentation of the 27 static representations.

**Post-Investigative Session**

• A brief open-ended interview was conducted about the participants’ study experiences, reflections, and subjective preferences.
• The participant was then given the movie ticket coupon and a signature of receiving the movie ticket coupon obtained. The participant was then thanked and shown out of the laboratory. This concluded the Study Session.

**RESULTS**

Eye-tracking data analysis was conducted at the aggregate level for each of the 27 static representations of open learner models. Three different kinds of analysis of the eye-gaze data were conducted using the iMotions Attention Tool 4.1 software: (a) emotional activation, (b) heatmaps, and (c) area of interest (AOI) analysis. Emotional activation is calculated based on the changes in participants’ pupil diameters [2, 13]. Emotional activation measures the level of arousal and engagement towards the stimulus image. The higher the emotional activation measure, greater the emotional impact of that learning analytic representation.

Heatmap presents the spatial distribution of students’ gaze on a particular representation. Heatmaps are composite images that contain an overlay of a gradient colour layer on the stimulus image (in our case, one of the 27 representations corresponding to the nine different notational systems and three different informational states) with areas of the stimulus image that received a greatest allocation of students’ gaze ranging from red to yellow and with areas that received the least gaze allocation ranging from yellow to green. The heatmaps presented below are static images of the aggregate gaze distribution on a particular image for all respondents. The Area of Interest (AOI) analysis was conducted on regions of the images that were of particular importance from pedagogical and/or user interface design perspectives.

The results section is organized as follows. First, findings about mean view time and mean emotional activation are presented. Second and last, for each of the 27 stimulus images, descriptive statistics, emotional activation, heatmap, and AOI results are presented and important observations are discussed.

**Mean View Time**

As can be seen from Figure 2, mean view time was the highest for the Collective Histogram notation followed by the Skillmeter and the Word Cloud representations. As can be seen from Figure 3, average view time was highest for the average learning representations compared to the Strong and the Weak learning state representations across all nine notational systems.

**Mean Emotional Activation**

As mentioned earlier, emotional activation is calculated based on the changes in participants’ pupil diameters. The higher the emotional activation score, greater the emotional impact of it. Figure 4 presents the average emotional activation for the nine notational systems. The least preferred notation of Word Cloud also received the least emotional activation. The Traffic Lights representations received the highest emotional activation followed by the Collective Histogram and Smilie notations. Figure 5 presents the average emotional activation for the three information states (weak, average, and strong). Unlike mean view time, no differences in emotional activation were found across the three information states.
Heatmap analysis showed that regions with greater information variance receive higher aggregate gaze allocation. In the Average Information State for the Notational System of Skillmeters (Figure 6), the gaze distribution is between the topic names, the skillmeter bars and the legend. In the Strong case (Figure 7), the “hotspot” is at the boundary between the green and the white areas of the topic. In the Weak case (Figure 8), the aggregate gaze distribution is greater around the region representing misconceptions in the individual student’s current knowledge (the red colored bar).

Regions corresponding to the Smilies Notational Systems in the main learning representation and the legend were selected for Area of Interest Analysis. AOI results show that participants found disambiguation of the Excellent and Very Good Smilies to be an issue. This was confirmed by the analysis of talk-aloud and semi-structured interviews data. Disambiguation of the OK and Weak Smilies were also problematic. Figures 9, 10, and 11 present the AOI results for the Average, Strong, and Weak information states of the Smilies representation.
DISCUSSION
Mean view time was the highest for the Collective Histogram notation followed by the Skillmeter and the Word Cloud representations. An analysis of the talk-aloud and post-investigative session interviews shows that participants found the Collective Histogram to be informative but challenging initially. It was informative because participants could perceive the individual’s learning within the context of the whole class and it was challenging as it required additional decoding of the social significance of the relative positioning of the individual student with respect to the collective of the classroom. Further, Word Cloud and Skillmeter were the two representations that participants liked the least and have higher view times. Unlike the Collective Histogram, the primary reason here is the difficulty of making sense of the representations for action-taking. The Traffic Lights and Smilies metaphors received relatively lower view time than the Word Cloud, Skillmeter and Collective Histogram. Average view time was the lowest for the representations that participants described with phrases such as simple, easy and straightforward during the talk-aloud and post-investigative session interviews. Given the prevalence of the Skillmeter notations in gaming, and in the context of current arguments about the “gamification of learning” [5, 8, 12, 14, 15], it is interesting that Skill Meters underperform both on sense-making and satisfaction dimensions.

Mean view time was highest for the average learning representations compared to the Strong and the Weak learning state representations across all nine notational systems. This could be due to the fact that decoding of the average case representations with a combination of both weak and strong learning sub-states is cognitively more demanding than decoding information at the extremes of weak and strong information states.

The least preferred notation of Word Cloud also received the least emotional activation. The Traffic Lights representations received the highest emotional activation followed by the Collective Histogram and Smilies notations. Many participants felt that traffic lights and Smilies were creative and easy to decipher. Traffic lights were also perceived as being simplistic and not depicting the semantic range of learning assessments from very weak to excellent as with other notations. There were little to no differences in emotional activation of static representations across the three information states of Weak, Average, and strong. Implications for the design of the nine notational systems are discussed below.

Implications for Design of Learning Analytics Notations

**Skillmeters**
Participants like the Skillmeters notation as it provided nuanced information. Several suggested adding a numeric value to the proportion (10%, 90% and such). The utility of for representing uncovered curriculum areas is not clear to some participants.

**Smilies**
Many participants thought that the use of Smilies was creative. In the future, we could consider the possibility of implementing Chernoff faces⁴.

**Traffic Lights**
Participants in general thought that the use of Traffic Lights like the Smilies was creative. Negative comments included the decreased range of knowledge level representations (only Strong, Average and Weak). Some participants suggested that the green light should be at the top as it is the best. One participant wanted the topics to be arranged from the strongest to the weakest so that they can know the “bottom” knowledge level. Adding more colors for the other knowledge levels is an option (Very Good and Very Weak). The design challenge here is to extend but at the same time preserve the metaphor of the traffic lights.

**Topic Checkboxes**
Topic Checkboxes are the second most disliked representation after Word Clouds. The design issue is the low color contrast between the different knowledge levels. Design implications are to increase the color contrast or to use multiple colors as in other representations (dark green to dark red).

**Class Histogram**
Class Histogram had the highest mean view time but participants found the absolute information of the individual embedded with the relative information about the whole class to be highly informative. The notational mark “star” needs to be changed to a neutral symbol and the legend should indicate clearly that it is a collective representation.

**Word Clouds**
Word Clouds are the most disliked and the most confusing representation. One design implication is not to repeat the topics between good and weak understanding level and to create just one simple Word Cloud with color coded topics. For example green—red color range indicating positive vs. negative category with topic size also coded for level of understanding.

**Textual Descriptions**
Most participants said that this representation was not that useful. They found the repetition of the text tedious and the overall representation boring. Unlike other representations, the Textual descriptors don’t reveal the full scale of the ratings. One design change to explore is to color code the knowledge level terms.

**Tables**
Many participants liked the Table representation and the mean view time was the lowest. Participants made

---

contradictory suggestions during the debriefing interviews. Some participants would like the vertical scale to range from the negative to the positive (unlike the traffic lights case) while others would keep it as it is.

Matrix
Matrix was by far the representation that most participants find as the easiest to interpret and has the lowest mean view time but second lowest emotional activation. Many participants suggested that the scale should be re-organized from the left to the right being negative to positive.

Other scale related suggestion was to add Excellent and Unacceptable as anchors (Excellent, Very Good, Good, Ok, Weak, Very Weak, and Unacceptable). For formative assessment purposes, “unacceptable” might be too strong a term and could be counter-productive.

Applications to the Design of Representations
A close analysis of participants’ talk-aloud comments during the investigative session and their observations in the post-investigative session interviews shows that learning analytics representations in themselves might be necessary but not sufficient for supporting meta-cognitive reflection and collaborative discourse. Within the context of the NEXT-TELL EU project, the results of this eye-tracking study of representations informed the design of representations for the Open Learner Model, the Communication and Negotiation Tool (CoNeTo), and the Teaching Analytics Dashboards for Repertory Grids. As mentioned earlier, an “open learner model” is a learner model that can also be externalised to the user [4]. This externalised (open) learner model may be simple or complex in format using, for example: text, skill meters, concept maps, hierarchical structures, animations [3]. CoNeTo provides computational support for the socio-cultural process of intersubjective meaning-making between students and teachers centered on their learning analytics representations (in this case, from the open learner models). Teaching analytics dashboards for RGFA allow teachers to use visual analytics techniques to conduct collective analysis of students’ personal constructs and ratings of domain concepts from the Repertory Grids for Formative Assessment application [23]. Figures 12, 13, and 14 present screenshots of different notations from the NEXT-TELL Open Learner Model.

---

5 http://www.next-tell.eu

Figures 12, 13, and 14: NEXT-TELL OLM Representations
In closing, based on the empirical findings from the study reported here, we argue that learning analytics representations are not always already artifacts that can support meaning-making and action-taking. Instead, learning analytics representations require explicit discursive support. Towards this purpose, we suggest the design,
development and evaluation learning analytics systems that facilitate artifact-centered discussion and negotiation.

ACKNOWLEDGEMENTS
This work is supported by the NEXT-TELL - Next Generation Teaching, Education and Learning for Life integrated project co-funded by the European Union under the ICT theme of the 7th Framework Programme for R&D (FP7). This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its content.

REFERENCES