Visualization of Professional Training Data in Academic Medicine

First Author* Second Author† Third Author‡
Institution Institution Institution

ABSTRACT

Visual analytics is increasingly being used to analyze educational data and learners' behavior in recent years. Online learning management systems often integrate educational dashboards that help learners to understand their performances, and educators to analyze learners' progress and needs. Although these tools have become ubiquitous for managing structured term-based courses, they are not suitable in the context of advanced training for qualified learners. Professional training often requires need-based scheduling and observation-based assessment. In this paper, we present a visualization platform designed for managing the training data in a medical education domain, where the learners are resident physicians and the educators are certified doctors. The system was developed through four focus group discussions with the residents and educators, several individual meetings with the educators, and many development iterations. We present how the professionals involved, nature of training, choice of the display devices, and the overall assessment process influenced the design of the visualizations (e.g., visual encodings and responsive design choices). The final system was deployed at the department of emergency medicine, and evaluated by both the residents and medical experts. Our analysis on one month of user logs reveals interesting usage patterns consistent with real-life training events. The users' feedback shows that both educators and residents found interesting use cases for the system.

Index Terms: Information Visualization—Visual Analytics—Visualization of Educational Data—Medical Training;

1 INTRODUCTION

Learning management systems often provide learners with visual cues to help them understand their performances and plan future actions. Learning analytics research examines how such information representation could be used to motivate learners to improve their performances. However, such studies are usually based on educational theories and put less emphasis on sophisticated visualizations. Most innovative visualization dashboards focus on helping educators to understand learners' performances at a glance, and perform comparative analysis of the learners' progresses [9]. Visual analytics research that help examine the educational data coming from online course platforms [3,10,11,24], has started to bridge the gap between learning analytics and educational data visualization.

Although much effort have been invested in understanding the scope of visualization in term-based courses that follow a structured (weekly/biweekly) activity schedules over a fixed academic term, the usefulness or need of visualization in the context of managing professional training is less explored. The scenario is particularly distinct in the following aspects:

- A professional training is often semi-structured and scheduled based on the need of an individual learner.
- A large part of the training may rely on observation-based assessment.
- Sometimes comparison among the learners are not meaningful due to the unique training characteristic for individual residents.

This motivated us to examine whether visual analytics can still be used to assist the learners and educators in such a context, and if so, then how the visualization views should be designed to help users perform their intended tasks.

In this paper we concentrate on a context of academic medicine, where qualified medical physicians (residents) practice their profession under the supervision of a senior clinician. All the aspects of a professional training mentioned above are prevalent in a residency training program. To understand the scopes and needs of visualization, we collaborated with a group of residents and clinicians in the department of emergency medicine. We first identified a set of requirements from the perspective of the residents, reviewers and program evaluators. We then took a 'simplified design' approach and developed three interactive views: *Resident View* (available for residents and reviewers to look at a resident's performance individually), *Reviewer View* (available only for reviewers to compare multiple residents at a glance) and a *Program Evaluator View* (available only for program evaluators to analyze the effectiveness of the training program).

The visualization system was designed through 4 focus group discussions, 6 major development iterations, and many agile scrum meetings during the development process. We evaluated the designed system through an uncontrolled longitudinal study and feedback sessions with the residents and reviewers. The results reveal that the both the residents and reviewers found the system to be useful in assisting them with their learning and training activities.

Our contribution in this paper can be summarized as follows.

(i) Identification of the visualization scopes and needs for learners and educators in a medical training program.
(ii) Design of a visual analytics system that helps residents to understand and track their performances and activities, and educators to make decisions in the training process.
(iii) Evaluation of the system by 6 medical practitioners and 16 residents in an uncontrolled longitudinal study, which shows the potential of the system in assisting in the residency training process.

2 RELATED WORK

In this section, we review the literature on the use of visualization in learning analytics, visual analytics for analyzing data coming from online courses, and learning analytics in the context of medical training.

2.1 Visualization in Learning Analytics

Learning analytics research often use various forms of information representation to motivate learners, whereas those representations are inspired by educational theories. For example, a dashboard can facilitate social comparison with the other learners by showing the learner's position among the peers, or can present a comparison with an earlier self by showing the learner's progresses, or can show the progress towards achieving a goal [16]. This line of research focuses on personalized visualization to motivate learners [19]. In our

*e-mail: first@abc.com
†e-mail: second@abc.com
‡e-mail: third@abc.com

context, the residents have a diverse set of needs and goals, which makes social comparison harder compared to the other representations (self-comparison and progress towards the goal).

In a recent survey on visual learning analytics, Vieira et al. [22] observed that most learner-facing dashboards created for learning analytics are limited to standard charts and plots [23], although there exist sophisticated visualizations (interactive multiple linked-view visualization [4]) and static dashboards [9] for presenting educational data that have been designed by the visualization experts. One reason for this gap seems to be the lack of collaboration among the developers, end users and visualization experts. Since traditional dashboards commonly focus on structured academic courses, they need to be adapted for the specialized educational programs such as residency training based on the need of the residents and educators.

2.2 Learning Management System Dashboards

The popularity of massive open online courses (MOOC) have created many MOOC platforms and generated vast amount of educational data. This has greatly advanced research in learning analytics that investigates learners' behaviors, learner-educator collaboration, and effective teaching methods. Most MOOC systems (e.g., Coursera [6], edX [7], FutureLearn [12]) integrate learner-facing dashboards to help learners understand their position in the learning context.

A rich body of research proposes various ways to use learning dashboards for motivating learners. A number of visual analytics systems have been proposed to analyze MOOC data to understand learners' behaviour [23], video streams viewed by the learners [3,24], communication among the learners on a forum [10, 11]. However, these approaches are not directly applicable to the professional training context, where the learners and educators frequently meet in face-to-face meetings for training and assessment activities.

2.3 Visualization in Medical Training

A number of commercial training management systems exist [5, 8] that focuses on general administrative processes such as session registration, scheduling, and reporting. But to the best of our knowledge, the scope of visualization for professional training is not yet well investigated, and a few studies have appeared only recently in learning analytics [18]. Recently, Boscardin et al. [1] have proposed some guidelines for designing learning analytics dashboard in medical education program, where they emphasized that visualization could play an important role in enhancing learners' experience.

Residents often need heavy faculty support to follow the learning goals and develop skills [17]. Intuitively, a learner facing dashboard for residency training should help analyze how the residents self-reflect, interpret and act on the educators feedback. However, determining how or whether the residents use the performance dashboard to guide learning varies widely among the students [14].

3 Application Background

Our investigation in this paper spans a professional training context in academic medicine. Our goal was to understand the scope of visualization to improve the training and learning process in a residency program (a training program for medical physicians). The investigation was carried out in a collaboration with a group of 16 residents and 6 clinicians of the department of emergency medicine.

The subsequent four sections describe the necessary background for problem formulation (Section 3.1), task abstraction (Section 3.2-3.3), and dataset characterization (Section 3.4).

3.1 Relevant Terminologies

- **Resident**- A medical physician undergoing training in a specialized field in a supervised setting.
- **Competence Committee**- A supervising committee that monitors a resident's progress and takes decisions regarding promoting a resident to the next phase of their learning.
- **Program Director**- A senior physician responsible for the development and management of the residency training program.
- **EPA** (Entrustable Professional Activity)- A task or a responsibility that a resident can be entrusted to perform in a heath-care context to demonstrate competence in that activity (e.g., EPA 2.1-Initiating and assisting in resuscitation of critically ill patients) [21].
- **EPA count**- This is the number of observations that have been made for a particular EPA type.
- **Rotation Schedule**- A year long schedule spanning 13 blocks (rotations) with 4 weeks each, in which a resident works across different departments (Cardiology, Trauma, Neurology, etc.) to gain diverse clinical exposure.
- **Training Phase**- A part of a resident's learning program. Residents are progressively promoted based on their performance through the following 4 phases before they finally graduate: Transition to Discipline, Foundations of Discipline, Core of Discipline, and Transition to Practice.

3.2 Problem Domain

The residents are recorded throughout their learning phase by a faculty or a senior resident in that medical program. The residents' scores depend on how they respond to patients on the tasks assigned to them. To ensure that residents have a broad understanding of their field, they are required to attend to a wide variety of medical scenarios. On each of these medical scenarios, the residents are rated on a five point scale and provided with feedback on how to improve.

Although this information is collected on a case by case basis, it can also be collated across multiple levels from various sources to gain a wealth of information about not just the individual resident, but the reviewers and the entire training program.

3.3 Requirement Analysis

To develop our visualization platform, we first conducted focus group studies with the intended users (residents and users). We then extracted a list of analytical tasks.

3.3.1 Focus Groups

To analyze the requirements of the visual analytic system, we conducted 4 focus group meetings, 3 with residents and 1 with competence committee members. Each meeting took approximately one hour.

Focus group with the competence committee: In the meeting with the competence committee, the participants were asked on what information they would like to know about a resident to review them and how they would want it to be presented.

Focus group with the residents: We used the first two of the three meetings with the residents for requirement gathering, and the third meeting to refine our design choices.

The first focus group meeting with residents involved 10 residents from the emergency medicine program. They were asked to describe all the data they wanted to highlight to their reviewers, and also to see for themselves for personal reflection. They were then asked to make rough sketches on how they wanted this information to be visualized.

We followed this focus group meeting with a second meeting two months later with 5 residents from the same program. The meeting was conducted as a think-aloud session, where participants were shown an initial list of requirements and a basic sketch of how the data would be presented and asked to comment on how they would improve it.

A third focus group meeting was conducted four months from the second focus group meeting. This meeting was held with the residents after our third major design iteration to refine our original requirements and get feedback on the system prototype.

Figure 1: Different Views of our visual platform (on anonymized data): (left) Resident view. (Right) Reviewer and Program Evaluator views.

3.3.2 Decision-Making Tasks

We collated all the requirements into three themes, each to be presented as a module in our visualization platform (Figure 1).

Resident View

- Q_1. Where is a resident currently in his training program, where was he before and where will he be next in a given academic year?
- Q_2. Is a resident meeting the required number of EPA observations needed in a specific rotation as per the resident's schedule?
- Q_3. How is a resident performing over time for a specific EPA? Does a resident meet the minimum number of EPA observations required (for a specific EPA) to demonstrate competence as per standard requirements?
- Q_4. In a given EPA, has the resident achieved the required diversity in terms of patient demographics (child, adult and senior cases) and medical scenarios?

Reviewer View

- Q_5. How is a resident performing compared to all the other residents overall and in a given time period?

Program Evaluation View

- Q_6. For a given academic year, how many EPA observations are being acquired monthly, and how do these numbers change based on the learning phase?
- Q_7. How many EPA observations are being acquired in each of the various departments as the residents cycle through their rotation schedule?
- Q_8. Given an EPA, how the EPA observations are distributed over various rotations? Similarly, given a rotation, how the EPA observations are distributed over various EPA types?

3.4 Data Characterization

To answer the questions Q_1-Q_8, we collected and combined two datasets: One is the residents' EPA assessment data and the other is their rotation schedule.

3.4.1 EPA Assessment Data

Residents are monitored every time they perform an EPA by a faculty or a senior resident. They receive a rating on a 5 point scale, as well as a qualitative feedback. This data is then entered into Mainport, an online website maintained across Canada by the Royal College of Physicians and Surgeons. For the purpose of our visual platform, we exported this data from the Mainport website and uploaded it into our system on a weekly basis. Since the start of this project, we have collected around 2800 EPA assessment records for 15 residents. Every assessment record has the following information: *Resident Name, Reviewer Name, EPA Type, Rating, Feedback, Observation Date, and Situation Context*.

Resident Name and *Reviewer Name* identify the residents and reviewers, respectively.

Observation Date is the day on which the EPA was completed.

EPA Type is a numerical label with a description of the professional activity. Every EPA assessment includes information about the patient demographics (e.g. age, gender) and the type of clinical presentation (e.g., shock, cardio-respiratory arrest, sepsis, etc.), which is provided as a list of values in *Situation Context*. *Feedback* comprises of a description of the actual situation and how the resident responded to it with additional comments on how to improve in future.

Rating is an ordinal value provided using an O-Score Entrustability Scale [13] with **1** being the worst and **5** being the best in the following order: I had to do **(1)**, I had to talk them through **(2)**, I had

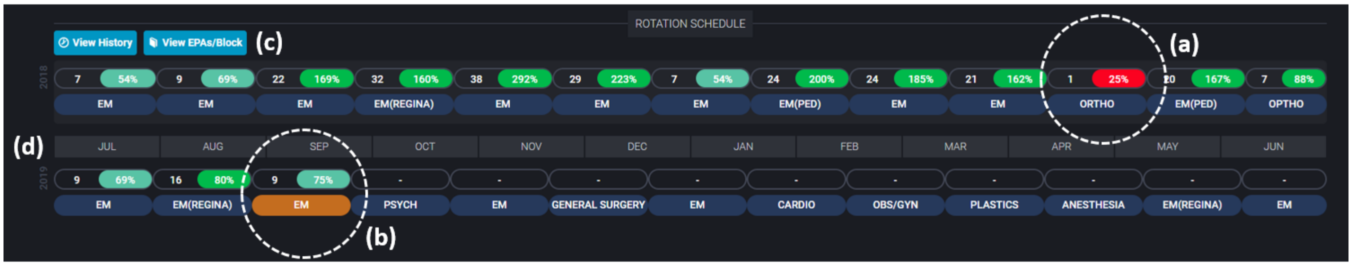

Figure 2: Rotation schedule of a single resident with the following components: **a)** Rotation with very low EPA count visually encoded in red to stand out. **b)** Current rotation with its corresponding rotation above showing that the resident has 9 EPA observations currently with a target achievement rate of 75%. **c)** *View History* and *View EPAs/Block* toggles turned on because of which, historical years and EPA count information layers are visible respectively. **d)** Reference timeline showing the academic year and dividing the current and historical years.

to prompt them from time to time **(3)**, I needed to be in the room just in case **(4)**, and I did not need to be there **(5)**.

3.4.2 Schedule Information

Every academic year, residents go through a 13 block rotation schedule. This **schedule is unique for every resident** and depends on their learning plan.

Along with this information, we also record important dates such as when the resident started the residency program and the different dates on which the resident was promoted to the various training phases. In a data preprocessing stage, we combined the schedule and the training phase promotion dates with the EPA assessment data to tag every record with two additional identifiers: *Rotation Tag* and *Phase Tag*. They were both calculated by placing the *Observation date* on a time scale to identify which rotation and training phase the resident was in when he completed the EPA.

4 Visualization Design

In this section we describe every part of our system and the visual encoding used in each to answer the questions of various user groups (residents, reviewers, and program evaluator).

4.1 Resident View

Resident view focuses on Q_1-Q_4 (discussed in Section 2.2) for an individual resident, i.e., the rotation schedule of the resident, acquirement of various EPA types, EPA-specific performance trend, and quick review of the resident's recent activities.

4.1.1 Rotation Schedule

A Gantt chart is a common type of chart used to illustrate a project schedule, but this takes much vertical space (one row per project activity). We observed that a nice property of a resident's rotation schedule is that no pair of rotations (e.g., Cardiology and Trauma) overlap as seen in our initial design in Figure 3. Hence we created an interval representation for the activity blocks, where the length of an interval represents the length of the rotation and collapsed our design vertically, as illustrated in Figure 2.

The interval chart is a combination of chronological rotational information along with EPA completion rate. To answer question Q_1 (i.e., the resident's current position in the training schedule), we first create a time axis consisting of a series of rotations with month names on them. The rotation schedule start from July and end at June, which represents the academic year of the medicine program. We then use the scale itself as a horizontal separator between the current academic year and the historical years (Figure 2(d)). Specifically, we show the rotations of the current academic year below the time axis, whereas the past years are depicted above.

For every academic year, we create a series of 13 consecutive rotations with the name of the rotation inside the block and arrange

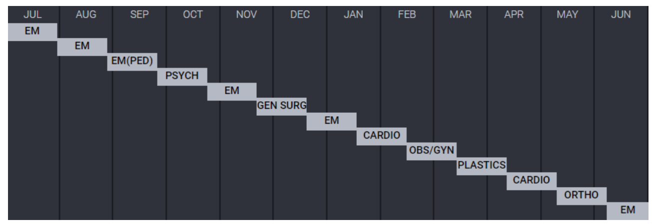

Figure 3: First design iteration showing a regular Gannt chart approach for the rotation schedule.

them chronologically based on the start date of each block. We highlight the current block the resident is in with an orange shade (Figure 2(b)) so that users can easily identify it. This provides an overview of what rotations the resident have completed, where the resident is now, and what rotations the resident will be completing next. Such an overview is not only important for the resident to self-reflect, but also to the reviewers to understand the resident's context. The uniqueness of every resident's schedule further justifies the need for such a concise representation.

4.1.2 EPA Count in a Rotation

To answer question Q_2 on whether a resident is achieving the required EPA count in a given rotation, we add a second block right above every rotation block and fill it with 2 numbers. The first is a number aligned to the left providing the sum of the EPA counts for all EPA types that a resident performed in that rotation. The second is the EPA target achievement, which is defined as the percentage of EPA observations completed by a resident to the target number of EPA observations required in a given rotation. The target is set by the competence committee based on historical performance of residents.

While looking at the target achievement rate, competence committee members are interested in identifying under-performers more than over-achievers, because in rotations with a small target count, residents can often over perform leading to very high percentages. We adopt a capped color scale instead of positional encoding because theoretically, the maximum target achievement rate is unbounded. We fill the target achievement number inside a cylindrical bubble and color the bubble using a linear scale that ranges from dark-red to dark-green for the percentage value of 0% to 100%. The scale is capped at 100% (i.e., any value above 100% remains dark-green). This ensures that rotations, in which a resident under performs, can be easily identified, but the over performing rotations will not stand out. Even though the target achievement rate is visually encoded by a color scale, we still explicitly list the numbers to mitigate any effect which could be associated with our choice to clamp the color scale on the higher end. Thus while under performing blocks visually stand out, the information about other blocks can be gleaned by reading the achievement rate text on every block.

Two buttons to toggle the visibility of the historical years and EPA target achievement rate are provided, and are turned off by default. This design ensures visual clarity by showing only basic information and hiding additional information under a collapsed view, which can always be brought up on demand. This option to toggle visibility of visual blocks and charts has been reused in several other components in our visualization platform, helping us achieve a minimalist design.

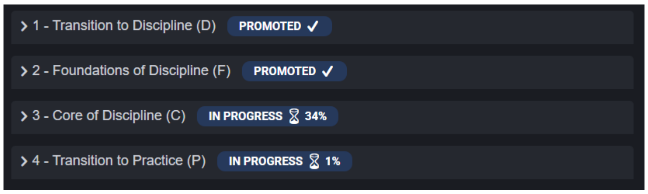

Figure 4: EPA Overview Table for a single resident currently in the training phase *Core of Discipline*.

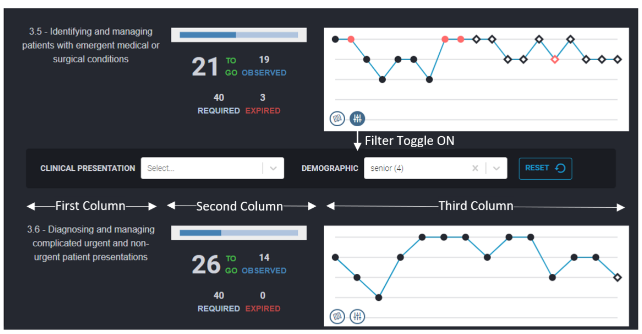

Figure 5: **EPA 3.5** in phase *Core of Discipline* with date filter set to August 1st – October 1st 2019 (diamond points) and demographic filter set to Senior patients (red points).

4.1.3 EPA Overview Table

The entire list of EPA types that residents are required to complete are broken down into four groups based on the training phase during which a resident is supposed to complete them and are numbered accordingly.

The serial ordering of the EPAs provides a natural choice for presenting them in a tabular format. We provide 4 blocks arranged vertically with the oldest starting from the top for every training phase as shown in Figure 4. For every phase, we present a completion rate which is calculated as the ratio of the sum of acquired EPA observations to the sum of required EPA observations for every EPA in that phase.

Although residents generally complete the EPAs of their current training phase before they pick up EPAs of later phases, there are exceptions. Due to various external factors such as their rotation schedules and the nature of medical cases of the patients they attend to, residents can occasionally end up completing EPAs which are not in their current training phase. This means residents can have a non-zero completion rate for training phases that they haven't yet started.

To present this information, we use a combination of icons, numbers and textual labels, as shown in Figure 4. The icons are used for quick lookup on which phases are complete and which are in progress. The labels are used to annotate the icons for the users unfamiliar with the meaning of the icons. The numbers, which represent the completion rate, are not shown for phases that are already complete (since they are not important to the users). Every row in the overview table acts as an accordion that can be opened and collapsed (Figure 1(left)), and the encoding of content in the accordion is discussed further in the following section.

4.1.4 Individual EPA Overview

When a user clicks on a row in the overview table, a list of all EPA types in that training phase are shown in a tabular layout with three columns, as shown in Figure 5.

First Column: In a particular training phase, the first column provides the identifying serial number for all the EPAs in this phase, as well as textual information describing the corresponding medical scenarios.

Second Column: In a particular training phase, the second column depicts the resident's position in each EPA of this phase. Each row represents a single medical scenario, and the second column of the row consists of two items (Figure 5). The first is a bullet chart [9] that visualizes whether a resident has met the target number required for that EPA. We visually encode the percentage of completion for a given EPA as the length of the bar inside the bullet chart. We do not use any scale on the chart, as it is there to assist users in gaining an approximate idea of the percentage just by a quick lookup. We also cap the bullet chart at 100% so if a resident over performs in an EPA it does not stand out. A common visualization strategy is to stack the excess performance using the idea of a Horizon chart [15], but we did not use this due to the unfamiliarity of such a chart among the users. However, this information can be inferred (if needed) from a section of numbers beneath the chart, as described below.

A series of four numbers bellow the bullet chart provides additional information (Figure 5) such as the number of EPAs still need to complete, the number of EPAs completed, the total number of EPAs required, and the expired EPAs (i.e., the EPAs that were not completed by reviewers within approximately one month). Since the number of EPAs that still need to be completed is an important metric for a resident, we calculate this number and present this information (with the text 'X to go') at the beginning of the other numbers in the series. In addition, we emphasize it by choosing a slightly bigger font-size than the rest.

The numbers are associated to the labels 'to go', 'completed', 'required', and 'expired', respectively. We choose the color based on what the label represents. For example, the 'expired' label is coded in red to draw attention of the user because every time an EPA expires, it is an EPA that even though a resident has completed was never reported. This is to emphasize that the residents should track their EPAs and make sure that they are completed on time. The label 'completed' is colored in blue, to match the color of the bar in the bullet chart as they both represent the same information. The labels 'required' and 'to go' are coloured in a lighter shade of blue and green respectively to bring a colour balance between the 4 numbers.

Third Column: In a particular training phase, the third column visually depicts a resident's progress for all the EPA observations in that training phase. Columns 2 and 3 together answer Q_3 (i.e., the position and performance of a resident for a specific EPA). We opt for a simplistic line chart design with minimal clutter. Every record is represented as a point with the oldest record starting on the left. The points are arranged vertically using the O-Score Entrustability scale with 5 being the highest (resident managed the situation independently) and 1 being the lowest (senior doctor had to completely take over the situation). The better a resident performs in an EPA, the higher is the vertical position of the point in the chart. We use background lines to show the 5 levels, instead of labelling the points, to reduce visual clutter as the levels are easy to understand without providing additional context.

A common design choice for representing a time series data is to

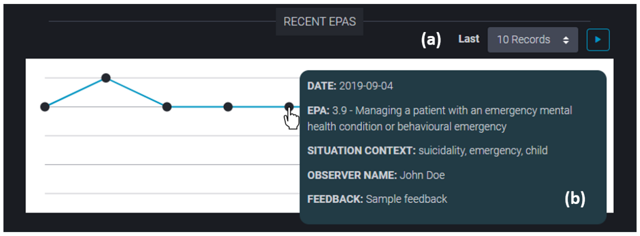

Figure 6: Recent History Chart. **a)** Filter to select the number of points to show in terms of last X months or last Y records. **b)** Popup screen displaying additional information about the point that is being hovered on by a mouse cursor (anonymized data).

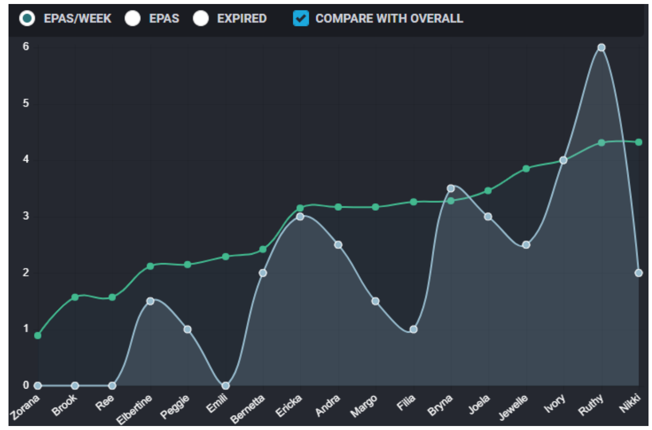

Figure 7: Review View comparing all residents with the green line representing overall EPAs/week and the blue line representing EPAs/week in the period August 15th-September 1st 2019 (anonymized data).

use a line chart with a date/time axis, where the axis is scrollable or zoomable. However, for the x-axis, instead of using an exact date/time scale, we use a set of equally spaced points in chronological order, with the oldest being on the left of the chart. The reason for not using the observation dates is that the observation dates are irregular, not important for the users, and furthermore, the observations often span a large period of time, i.e., residents often complete EPAs in chunks with long periods of time gaps in between.

Every point (observation score) can be hovered using a mouse to bring up additional information about that record in the form of a pop-up. This includes observation date, feedback, the patient demographics and medical context of the case.

Finally, to answer Q_4 (i.e., diversity of the demographics covered by the resident), we provide two small buttons at the bottom-left corner of the chart. The first one can be clicked to see all the records in a table that can be sorted and filtered by columns. The second button brings up two filter lists that can be used to visually identify a particular record based on patient demographics or medical context. For example, if a user wanted to see which of the records were for senior patients, they could select the 'Senior' option from the drop-down list and the corresponding points (observation scores) would turn light red. We used a categorical color palette for the demographics.

Because every EPA observation record has an observation date associated with it, we can filter the records to look at a resident's performance in a given time period. However, this information is only useful when looked at in the context of the other records the resident has completed, because, every EPA observation is a test demonstrating competence and this can only be acquired over time. Hence instead of showing only those scores on the line chart, we change the shape of the observation points that lie within the selected date range. Regular records are shown as circular disks, while records that fall in the selected date period are shown as hollow diamond points (Figure 5). The two visual encodings (shape and color) can be combined intuitively. For example, we can lookup all the Senior cases in EPA 3.5 that fall in a time period from August 1st to October 1st, 2019, using a combination of the date filter and the demographic filter, as shown in Figure 5.

4.1.5 Recent History Chart

This chart is meant for quick lookup of a resident's recent performance with the option to view records in the following ranges: last 10, last 25, last Month and last 3 Months. The chart does not visually distinguish the different EPA types (i.e., EPA-2.1 vs EPA-3.2), instead, provides this and other additional information in a pop-up menu that can be invoked by hovering over a point, as shown in Figure 6.

4.2 Reviewer View

While assessing a resident in a residency program, a reviewer is often interested to see an overview of the resident's performance. Therefore, we made the resident view accessible to the reviewers. We were also interested to know whether comparing residents with their peers will be valuable, and we found mixed opinions. Some reviewers mentioned that usually they do not compare a resident's performance with peers, but some others mentioned that a high-level overview of how a resident is performing compared to the other residents (i.e., Q_5) in the same training phase might give some insight about the diversity in the residents' performances.

In the comparative chart that we developed for the reviewers, we present three metrics for each resident: *Total EPAs, Total Expired EPAs,* and *EPAs/Week*. The *Total EPAs* and the *Total Expired EPAs* are the count of all EPA observations and the count of all expired EPA observations across all EPA types, respectively. The metric *EPAs/Week* is an indication of a resident's weekly EPA observation acquisition rate, and is defined as the average number of EPA observations a resident acquires in a week across all EPA types. The reason for choosing a week as a time interval is our pre-analysis of the data, where we noticed that residents often acquire EPAs in clusters with multiple EPAs in a single day followed by several days with no EPAs.

A line chart is used to represent the data and users are provided with the option to select any one of the above mentioned metrics using radio options in the panel above the graph, as in Figure 7. Users can look at the data in a particular time period, and also compare the value of a metric in a time period against the entire time a resident was active in the program. For example, in Figure 7, the green line represents the EPAs/week of all residents over their entire program so far, and the blue line represents the EPAs/week that the residents acquired in the period from August 15th to September 1st, 2019. The checkbox *Compare with Overall* can be checked on or off to toggle the visibility of the green line showing the overall trend.

4.3 Program Evaluation View

Program evaluation view consists of four linked views with two filters on the top (Figure 1(top-right)). The first filter lets users select a particular academic year. This is because the rotation schedules are planned out to be a year long activity. The second filter is for selecting the training phase. This filter can also be set to all phases to look at the data collected across all the four different training phases.

4.3.1 Monthly Distribution of EPA Count

Program evaluators were interested in examining the rate at which the EPAs are being completed over time and how the rate varies over

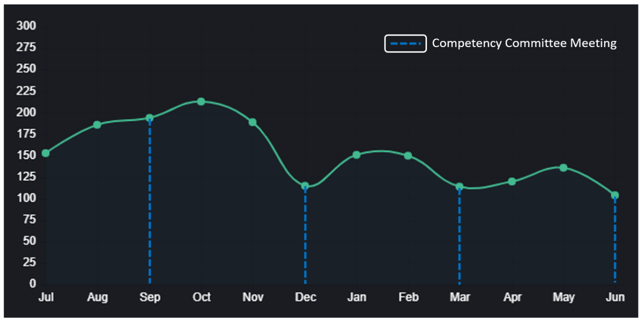

Figure 8: Monthly Distribution for a particular Academic Year with vertical lines indicating competency committee meeting dates.

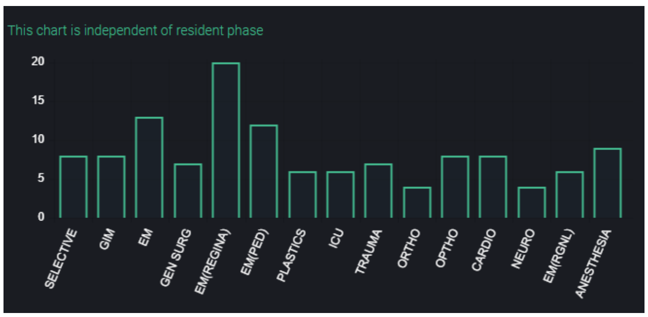

Figure 9: EPA Count per Rotation for a particular academic year.

various learning phases (i.e., Q_6). The rate of EPA completion can be dependent on several factors such as the variation in the residents' rotation schedules and training phases. We created a chart that bins all the EPAs into monthly intervals based on their *Observation date* and then display them as a trend line, as shown in Figure 8. We adopt a timeline that starts from July and ends in June to match the medicine academic year cycle. This chart works in conjunction with the academic year and training phase filters described above to answer Q_6. In addition to presenting the residents' activities in various training phases, the chart can also help understand the influence of the quarterly competence committee meetings on the total EPA observations being acquired.

4.3.2 EPA Count per Rotation

The program evaluators are also interested in knowing how many EPA observations are being acquired per rotation (i.e., Q_7). This is helpful in identifying rotations in which residents are under performing or simply not receiving adequate feedback to support their learning and progress in the residency program. In addition, evaluators need to come up with a target EPA count (i.e., the required number of EPA observations to demonstrate competency) for every rotation, based on the performance of the residents in the previous academic years.

Although every rotation is 4 weeks long, the number of EPA observations that residents acquire in every rotation can be different. For example, the average number of EPA observations completed per rotation for residents on an *Emergency Medicine* rotation can be very different from the number of EPA observations completed on a *Cardiology* rotation. To calculate this information, we first select the records that lie in a given academic year, and then group them based on their rotation tags, which were created in the pre-processing stage discussed in Section 3.4.2. Since we know the rotation schedules of all the residents in the selected academic year, we can get the number of times a particular rotation is assigned to a resident. To compute the average EPA count per rotation, we divide the total number of EPA observations acquired in a particular rotation by the number of times that rotation was assigned. Since this information is distinct for every rotation, we present it as a bar graph, as shown in Figure 9.

4.3.3 EPA Specific Rotation and Rotation Specific EPA Distribution

To help program evaluators understand the relation between individual EPA types and rotations (i.e., Q_8), we construct two charts that work in tandem (Figure 10). The first chart is an EPA specific rotation chart, where users can choose an EPA to inspect the rotations in which the selected EPA gets filled out. The second chart is a Rotation specific EPA chart, where users can choose a rotation to see all the EPA types that get filled out when residents are in that rotation. For example, in Figure 10(left), we can see that *EPA 3.6* was filled out the most in the rotation EM (Emergency Medicine) compared to other rotations, and Figure 10(right) shows that out of all the EPA types in EM, *EPA 1.2* was filled out the most.

5 Implementation Details

In this section we describe the implementation details of the data access layer and the presentation layer of our visual platform.

5.1 Back-end Development

We developed our visual platform as a single page web application that is linked to a NodeJS REST API in the backend, which in turns connects to a NoSQL MongoDB database, where the actual data is securely stored. We adopted this distributed architecture by using a REST API instead of the more traditional dynamic approach with pages being rendered on the server for two main reasons: loose coupling, and easier scalability. Since the project involved rapid prototyping with the design of the dashboard changing constantly, we wanted to keep the server code separate and untouched, as we iterate over different designs on the UI and deploy several versions in parallel. We also took into consideration the possible increase of the server load, as the project gets wider adoption within the university across multiple medicine programs (even though we only started with emergency medicine). Since the server does not hold any state at a given point of time, multiple instances of the backend server can be spawned behind a common gateway, thus drastically improving the response time. Since the system handled sensitive medical data, we restricted access to the database based on user roles (resident, commitee reviewer, and program director). However, instead of creating a new user management system, we adopted the University's central authentication service and assigned users JWT (JSON Web Tokens) on the UI when they logged in. These tokens were then used to screen all API calls and restrict unauthorized access.

5.2 Front-end Development

On the front end side, the single page application was built using a combination of *React.JS* [20] (a web library used at Instagram and Facebook for handling rapid data updates to the screen) and *D3.JS* [2] (a web visualization library). Our design was implemented with the latest web standards and has been open sourced on GitHub under an MIT license.

5.3 Responsive Design

Users access our system using a wide range of devices, i.e., smartphones, tablets and desktop computers. To ensure that the user experience was uniform across various screen sizes and orientations, we adopted a responsive design as shown in Figure 11.

The idea of using a collapsible overview table for visualizing blocks of information is very useful in devices with small screen, as the cells can be collapsed and expanded as needed. All visualizations

Figure 10: Relation between EPAs and rotations represented through two charts.

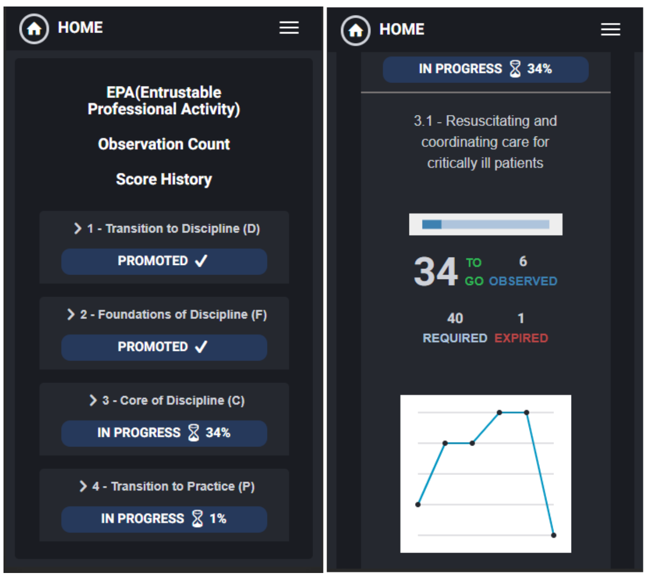

Figure 11: Screenshots of the system on a mobile screen showing responsive design.

in the system adapt based on the width of the screen, and are rendered in a scale and transform invariant SVG (Scalable Vector Graphics) format.

6 User Evaluation and Feedback

We deployed a stable version of the visualization system online for the department of emergency medicine at the university, and provided access to all three user groups: Residents (16), Reviewers (6) and Program Evaluators (2). The users used the system for two weeks, which made them familiar with the visualization platform. We then conducted an uncontrolled longitudinal case study by recording user logs for a period of 5 weeks on the website. We then analysed these logs to understand whether the users interact with the system, and if so, then how various real-life scenarios influence the system use. The reason we were interested to examine whether and how the usage pattern varies, is because it is an indicator that the users found the system useful for their real-life decision making.

We also collected feedback from different user groups through feedback sessions with seven residents after they had a considerable exposure to the system. We also showed our system in meetings with five reviewers and two program evaluators.

6.1 Evaluation Methodology

To quantify user engagement, we logged every user visit to the website. The logs included the name of the user, time of visit, the user type and the page visited on the system. We recorded the logs for a period of 5 weeks (around a month) from 18th August to 21st September, 2019. We selected the time period to intersect with the quarterly competency committee meeting date for emergency medicine (September 10th) as it would help us understand whether or how the usage changes during and around these meetings.

We only kept the logs for the intended user groups, i.e, filtered out the logs of the developers, who were involved in the maintenance of the system, and administrators who periodically uploaded resident data onto the website.

6.2 Evaluation Results

Based on the user logs, we found that the resident view was the most visited part of our system. This is because of the larger number of residents, and also that it can be accessed by multiple user groups (residents to view their own data and reviewers to view data of all residents individually). The program evaluator view and the reviewer view have only a small number of users. Furthermore, these views are not meant for day to day analysis, but rather for understanding the big picture once every quarter or a year. We thus focus our analysis on the usage of the resident view.

6.2.1 Usage Pattern Influenced by Assessment Meetings

To understand how user engagement changes over time we binned our logs into one week long intervals. We then grouped all the logs in each weekly interval based on the user type (resident or reviewer) and counted every log as one visit. The total number of weekly visits made by all users to the resident view reveals that the count peaked during Week 4 (Sept. 8 – Sept. 14), as in Figure 12(left). This could possibly be because the competence meeting happened on the September 10th. This pattern is also inline with comments made by residents during the feedback discussion such as *"I have looked at it prior to usually like my competency committee meetings. That is kinda when I mainly use it"*(P4).

The break-down of the visits based on the user types reveals that resident visits remained fairly consistent with slight increase during the meeting week, but the number of visits made by reviewers changed drastically. As shown in the stacked bar chart of

Figure 12: (left) Number of visits made by different user groups to the Resident View during the Study Period. (right) Visits made by residents to the system distributed over days.

Figure 12(left), visits made by the reviewers were very high during the meeting week and the week before it and then dropped again after the meeting period.

6.2.2 Usage Pattern Influenced by Information Updates

The residents' weekly EPA scores are imported into our platform every Tuesday. Although the residents get their EPA scores and feedback over their email earlier than Tuesday, several residents mentioned that they login every week especially on Wednesdays to check their progress, which is also an indication that they found the system useful to track their performances. We examined this by analyzing usage logs based on the days of a week and found a peak on Wednesdays, as shown in Figure 12(right). This clearly demonstrates the level of engagement some of the residents have with the visualization system, and how it has become an integral part of their learning program.

6.2.3 Residents' Feedback

Residents can be broadly broken into two groups: some who visit the website only during review meetings, and the others who use it more frequently on a weekly basis. The difference between the groups can be partially explained by how well versed they are with the different aspects of the system.

Some residents seem to use the system extensively -*"I use it just as an interface just to get a better kind of gestalt view of how I'm doing. I can track my progress and EPA numbers per week which is important to me...The other thing I like to use it for is using it to kind of plan my life out because it has all of the dates of my rotations kind of in a convenient space"*(P2). We believe that as residents gain more experience in using the different components of the system, they may transform into frequent users of the system. Some residents mentioned that they use the system for casually looking at their EPAs - *"...I would go onto it just to peek at which EPA, to look at like how they've been trending and how I've been doing with most of them. But me personally, I haven't used it a ton. But that's probably 'cause I don't know all of the aspects of it that I can utilize"*(P1).

We have also seen that the residents found their own ways of using the system. For example, they sometimes used the visualization *"to identify gaps ... Like needing more assessments for a certain EPA"*(P6), or to *"plan for how many EPAs I need to get done per shift in the coming year, with a pretty good idea of which ones I'm more lacking*"(P7), or to use *"as a documentation of the conversation happening"*(P6).

6.2.4 Usage Frequency vs. Expired EPAs

Since we observed two types of residents (frequent and non-frequent) users, we examined whether there is any difference between these two groups in any of the EPA related metrics.

We classified the residents into two above mentioned groups as follows: residents who visited the website at least once a week during the study period, and residents who did not. We then looked at different metrics of EPA acquisitions for the two groups and found that the rate of expired EPA observations for residents who visited the dashboard at least once a week was 3 times lower than the expiry rate for residents who did not login as frequently. A potential reason could be that residents who frequently looked at the dashboard, ensured that all EPA observations they completed were submitted on time.

We did not notice any interesting pattern in other metrics such as the number of EPA observations acquired and average EPA score.

6.2.5 Feedback from Reviewers and Program Evaluators

We demonstrated the system to the reviewers and program evaluators in individual meetings. The reviewers appreciated the visual platform mentioning that the *"layout is good"*, *"it's way more user-friendly"* than the existing system, and also that they are getting used to the visualization platform over time: *"every time it's getting easier so that's a good thing"*. Reviewers also liked the simplicity of the design, e.g., *"just keeping it simpler and cleaner is better in my mind"*.

The visualization platform greatly helped the competence committee meeting: *"especially this last meeting was the best one I thought that we've had overall. It was the easiest meeting to get the data and put everything together"*.

The reviewers appreciated various visualization components. For example, they mentioned that the rotation schedule view (Figure 2) was *"very helpful just to actually visualize the number of EPAs they've had per block"*. A common use for the EPA overview (Figure 5) was to *"look at all the outliers, if any of the EPA ratings that are 3 or less... then go through the actual feedback with those just to see if they're actually correlated"*. Both the demographics filter and the recent history chart (Figure 6) were helpful for the reviewers in their review. Some reviewers mentioned that they are able to perform even deeper investigations, e.g., *"to see if the numbers are low, why they're low, what rotation they were missing... if a rotation didn't go well... so-and-so was struggling... then you could see that their numbers were also low, then it would be a flag to talk to the resident to see what's going on"*.

On a question about whether they compare resident performances (Figure 7), some reviewers mentioned that they prefer *"comparing resident to a standard"* rather than peers because *"everybody's different"*. But some reviewers mentioned that *"I'll kinda compare them. Just to see if anybody's kinda really an outlier, falling behind"*. Sometimes they mentioned this to be useful for looking up *"how many EPAs they should be getting in that four-month timeline. So I compare them to their cohort"*.

The enthusiasm of the reviewers were eminent from their various expectations to further enrich our visualization system. For example, to have *"a summary since the last competency committee meetings"*, to integrate other datasets such as *"self-reflection of residents "*, and even to extend the system for the evaluation of the reviewers' feedback.

The program evaluators appreciated that the *"competence committees are making evidence-based decisions"*, however, whether or how these visualizations could bring a change in a program would require data over a much longer period of time, and also the EPAs acquired by the residents in similar programs in other universities.

7 Design Choices and Insights Gained

Professional training largely differs from a structured course in its academic activities. One of the most important insights that we gained in our study is how such differences demand a design that is different from a dashboard designed for a typical structured course.

The key design issue is that a resident is assessed based on his unique rotation schedule and need, whereas in a structured course context, the learners are assessed altogether by the same assignment or test. Hence we had to make the design choices focusing largely on individual residents rather than on their comparative performance analysis. We thus recommend designers to gain a better understanding of the training context, which would help them verify whether to adopt an existing visualization framework or to develop a new one.

Since resident view was accessed and used by both residents and educators, the usage varied widely among the users. Although we attempted to have a minimalist design, people needed some time to get used to the system, e.g., sometimes they did not notice how various filters could be used, and after pointing those out they mentioned that *"that is amazing!"* Through our iterative development, we observed that starting with a simple and intuitive design, and then adding new features incrementally, is often useful to help users to get used to the system. We came across several other important design considerations while developing the system, as follows:

- Using a combination of icons and labels to improve the readability of various icons.
- Highlighting of the important assessment criteria for both the residents and reviewers.
- Grouping visual charts in collapsible blocks with headers to create an overview that can be expanded on demand.
- Ensuring a responsive design when the visualization platform would be accessed on small devices such as smartphones.
- Rendering the simplest chart first and then facilitate filters such that advanced users can use them for deeper investigations.

8 Limitations and Future Work

A natural limitation of our work comes from our exclusive focus on the emergency medicine department. However, we believe our findings will be applicable to visualize training data for various other departments. Recently, our visual platform has on-boarded 3 new programs (Internal Medicine, Pathology, and Obstetrics/Gynecology) and several other programs within and beyond the university have also shown interest in adopting our system. One possible avenue for future research is to examine the generalizablity of our visualization approach for other professional training scenarios.

The focus group participants were all from North America, and their feedback and opinions may differ from the medical professionals from other other parts of the world. Whether the differences in the cultural aspects in how residents should be trained require different design considerations, would be an interesting direction to explore.

We only had a small number of reviewers (6) and program evaluators (2) available. It would be interesting to examine the usage pattern with a larger number of users when other medical departments start using our platform. Although the program evaluators were positive about the program evaluation view, this view is expected to be viewed infrequently, i.e., the program evaluators are expected to look at it once every quarter of a year or less. Consequently, we could not analyze the usage logs for the program evaluation view. We expect that a future longitudinal study spanning years would help us better understand the use cases for this view.

9 Conclusion

In this paper we presented a visualization platform to assist various user groups (residents, reviewers, and program evaluators) to analyze residency training data in academic medicine. We conducted an uncontrolled longitudinal study over five weeks. Our analysis of the user logs and feedback from various user groups showed that the developed visualizations were effectively used by the users in various learning and training processes. We believe that our study in this paper will inspire further visualization research in learning analytics of professional education.

References

[1] C. Boscardin, K. B. Fergus, B. Hellevig, and K. E. Hauer. Twelve tips to maximize the value of a clinical competency committee in postgraduate medical education. *Medical Teacher*, 40(11):1110–1115, 2018.

[2] M. Bostock, V. Ogievetsky, and J. Heer. D3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, Dec. 2011.

[3] Q. Chen, Y. Chen, D. Liu, C. Shi, Y. Wu, and H. Qu. Peakvizor: Visual analytics of peaks in video clickstreams from massive open online courses. *IEEE Transactions on Visualization and Computer Graphics*, 22(10):2315–2330, 2016.

[4] Q. Chen, X. Yue, X. Plantaz, Y. Chen, C. Shi, T.-C. Pong, and H. Qu. ViSeq: Visual analytics of learning sequence in massive open online courses. *IEEE Transactions on Visualization and Computer Graphics*, 2018. In press.

[5] Convergence LMS. `https://www.convergencetraining.com/default.aspx`. Accessed: Sept. 2019.

[6] Coursera. `https://www.coursera.org/`. Accessed: Sept. 2019.

[7] edX. `https://www.edx.org/`. Accessed: Sept. 2019.

[8] EHS Software. `https://www.trainanddevelop.ca/ehs-software/`. Accessed: Sept. 2019.

[9] S. Few. *Information Dashboard Design: Displaying Data for At-a-Glance Monitoring*. Analytics Press, 2013.

[10] S. Fu, Y. Wang, Y. Yang, Q. Bi, F. Guo, and H. Qu. Visforum: A visual analysis system for exploring user groups in online forums. *ACM Trans. Interact. Intell. Syst.*, 8(1):3:1–3:21, 2018.

[11] S. Fu, J. Zhao, W. Cui, and H. Qu. Visual analysis of mooc forums with iforum. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):201–210, 2017.

[12] FutureLearn. `https://www.futurelearn.com`. Accessed: Sept. 2019.

[13] W. T. Gofton, N. L. Dudek, T. J. Wood, F. Balaa, and S. J. Hamstra. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. *Acad Med*, 87(10):1401–1407, Oct 2012.

[14] K. E. Hauer, N. Iverson, A. Quach, P. Yuan, S. Kaner, and C. Boscardin. Fostering medical students' lifelong learning skills with adashboard, coaching and learning planning. *Perspectives on Medical Education*, 7(5):311–317, Oct 2018.

[15] A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, and H. Pfister. Upset: Visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.*, 20(12):1983–1992, 2014.

[16] L. Lim, S. Dawson, S. Joksimovic, and D. Gašević. Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference? In *Proc. of the 9th International Conference on Learning Analytics & Knowledge*, LAK19, pp. 250–259. ACM, 2019.

[17] T. Lockspeiser, S. Li, A. Burke, A. Rosenberg, A. Dunbar, K. Gifford, G. Gorman, J. Mahan, M. McKenna, S. Reed, A. Schwartz, I. Harris, and J. Hanson. In pursuit of meaningful use of learning goals in residency: A qualitative study of pediatric residents. *Acad Med.*, 91(6):839–46, 2016.

[18] M. Mouaici, L. Vignollet, C. Galez, and M. Etienne. Learning analytics dashboards for professional training - challenges and proposal. In *Proc. of the 13th European Conference on Technology Enhanced Learning*, Lifelong Technology-Enhanced Learning. University of Leeds, 2018.

[19] F. A. Orji, J. E. Greer, and J. Vassileva. Exploring the effectiveness of socially-oriented persuasive strategies in education. In H. Oinas-Kukkonen, K. T. Win, E. Karapanos, P. Karppinen, and E. A. Kyza, eds., *Proc. of the 14th International Conference on Persuasive Technology (Persuasive)*, vol. 11433 of *LNCS*, pp. 297–309. Springer, 2019.

[20] React. `https://reactjs.org/`. Accessed: Sept. 2019.

[21] O. Ten Cate. Entrustability of professional activities and competency-based training. *Medical Education*, 39(12):1176–1177, 2005.

[22] C. Vieira, P. Parsons, and V. Byrd. Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, 2018.

[23] C. Vieira, P. Parsons, and V. Byrd. Visual learning analytics of educational data: A systematic literature review and research agenda.

Computers & Education, 122:119–135, 2018.

[24] J. Zhao, C. Bhatt, M. Cooper, and D. A. Shamma. Flexible learning with semantic visual exploration and sequence-based recommendation of mooc videos, 2018.