Research Note

Toward Better Outcomes in Audiology Distance Education: An Educational Data Mining Approach

Bruno Elias Penteado,^a Paula Maria Pereira Paiva,^b Marina Morettin-Zupelari,^b
Seiji Isotani,^a and Deborah Viviane Ferrari^b

Purpose: This article introduces concepts and a general taxonomy used by the educational data mining (EDM) community, as well as examples of their applications, with the aims of providing audiology educators with a referential basis for developing this area.

Method: A narrative review was carried out to present an overview of EDM and its main methods. Some of these methods were exemplified with analysis of real data from an Internet-based specialization course on pediatric auditory rehabilitation.

Results: The review introduced EDM main concepts and applications and described methods from its area. Real

data examples illustrated EDM use to predict interpersonal help-seeking, model interpersonal interaction, analyze students' trajectories within a course's module, and understand how students approached group assignments. Some of the insights provided by EDM to support teaching and learning processes were also described.

Conclusions: EDM methods offer new tools to discover knowledge from digital traces (i.e., logs) and support key stakeholders (students, instructors, or course administrators) to raise awareness about course dynamics. This approach has the potential to foster a better understanding and management of educational processes in audiology distance education.

The World Health Organization (2013a) estimates the global shortage of health care workers to reach 12.9 million by the year 2035, this being one of the most critical constraints to the achievement of health and development goals. Consistently, the shortage of skilled and equitably distributed audiologists, acuter in developing countries, is a critical workforce concern (Goullos & Patuzzi, 2008; Windmill & Freeman, 2013) with serious implications for hearing health care.

Tackling this complex issue requires a multidimensional process to transform and scale up the education of health professionals. One aspect of this process is related to increasing the number of professionals and, more importantly, ensuring they have knowledge, skills, and competencies relevant to the population health needs (World Health Organization, 2013b). As essential players in this transformation, educational

institutions need to increase their capacity for training a larger number of professionals. The scope and magnitude of such challenge imply, among other things, more effective and strategic use of available resources, including the adoption of innovative approaches based on information and communication technologies for health professional education (Al-Shorbaji, Atun, Car, Majeed, & Wheeler, 2015).

Over the past decade, information and communication technologies innovations made educational interventions mediated via the Internet (e-learning) more manageable, accessible, and affordable. In many developed countries, e-learning is now part of the mainstream. For instance, in the United States, one in four students is taking, at least, one course at distance. From 2012 to 2014, overall higher education enrollment fell by 2%, while growing by 9% in courses exclusively at distance (Poulin & Straut, 2016). In recent years, e-learning has also made its way into developing countries. In Brazil, the number of higher education students enrolled in distance learning increased by more than 20% between 2015 and 2016 (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2016).

^aInstitute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, Brazil

^bSpeech Language Pathology and Audiology Department, Dental School of Bauru, University of São Paulo, Bauru, Brazil

Correspondence to Bruno Elias Penteado: brunopenteado@usp.br

Editor-in-Chief: Sumitrajit (Sumit) Dhar

Editor: Ariane Laplante-Lévesque

Received January 16, 2018

Revision received June 7, 2018

Accepted June 8, 2018

https://doi.org/10.1044/2018_AJA-IMIA3-18-0020

Publisher Note: This article is part of the Special Issue: Internet and Audiology.

Disclosure: The Samaritano Association (SA) is a philanthropic organization responsible for managing the financial resources from Proadi/SUS Grant 25000.024953/2015-89. Through this grant, SA paid for expenses such as hosting the Moodle™ platform used in this research, traveling and housing for students and instructors (when applicable), producing didactic materials, and providing stipends for instructors. The authors Paula Paiva, Marina Morettin-Zupelari, and Deborah Ferrari have declared they received such stipends. The authors have declared that no other competing interests existed at the time of publication.

In higher-income countries, audiology e-learning has been widely available at different academic levels. Nonetheless, research on outcomes of e-learning for the education of audiology professionals or students is still scarce. At the time of this writing, only two publications were found, showing distance learning efficacy for teaching counseling skills (Crandell & Weiner, 2002; English, Rojeski, & Branham, 2000).

For health professions, recent systematic reviews found that, in general, e-learning studies are very heterogeneous and their methodological quality needs improvements. Still, a body of evidence indicates e-learning allows for increased accessibility and convenience, portability, and interactivity with instructors and peers. Several findings indicated that e-learning is at least as effective as traditional learning methods when considering knowledge and skills acquisition (Al-Shorbaji et al., 2015; George et al., 2014; Sinclair, Kable, Levett-Jones, & Booth, 2016). Advantages for educational providers include cost savings, coverage, scalability of instructional materials, and enabling students to practice skills prior to experience with real patients (Al-Shorbaji et al., 2015). Moreover, e-learning enables “in-service” education, so that health professionals’ knowledge is built according to relevant population needs, thus strengthening the development of new health practices (Silva, Santos, Cortez, & Cordeiro, 2015).

Despite e-learning advantages, variation in knowledge outcomes exists, depending, among other factors, on the pedagogical approaches utilized in the courses (Al-Shorbaji et al., 2015; George et al., 2014; Sinclair et al., 2016). E-learning also carries other issues that are critical for a successful end point—for example, keeping students engaged and motivated online (e.g., Schaeffer & Konetes, 2010), detecting potential dropouts (e.g., Bayer, Bydzovská, Géryk, Obsivac, & Popelínský, 2012), reducing the sense of isolation (e.g., Al-Samarraie, Selim, Teo, & Zaqout, 2017), and identifying students who are struggling with the content and are in need of support (e.g., Staikopoulos & Conlan, 2017). A deeper understanding of the design of e-learning instruction and its efficacy and effectiveness, in order to improve practice, is essential.

Currently, the most popular learning management systems (LMSs) include at least an analytic panel (dashboard; Macfadyen & Dawson, 2010) where it is possible to visualize information such as user profile, number of accesses (e.g., in a given course, content, or activity), session duration, activities performed (e.g., content read, messages written, assignments, and tests taken), as well as academic and interaction outcomes (Romero, Ventura, & García, 2008). Although highly valuable for administrative tasks, such as monitoring students’ virtual attendance or usage of tools, these reports alone do not respond to the more complex questions that concern educational phenomena. To deal with this problem, a relatively new field has emerged, and it is known as *educational data mining* (EDM). EDM explores statistical and computational methods for data analysis and theory testing, aiming to help a better understanding and improvement of teaching and learning process.

To the best of our knowledge, there is no research exploring EDM in audiology education. Given that knowledge generated through EDM allows for the development of new and more effective teaching methods (R. Baker, Isotani, & Carvalho, 2011), this two-part article aims to provide audiology educators with a referential basis to further develop this area. Part I presents a narrative review with introductory concepts and a general taxonomy commonly used by the EDM community. Part II integrates some of these concepts in an example of audiology e-learning, thus complementing other articles that compiled different aspects of EDM (R. Baker et al., 2011; R. Baker & Siemens, 2014; Koedinger, D’Mello, McLaughlin, Pardos, & Rosé, 2015; Romero & Ventura, 2013) and addressed learning analytics in health care education (Ellaway, Pusic, Galbraith, & Cameron, 2014).

Part I—EDM: Introductory Concepts

EDM is a discipline “concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings and using those methods to better understand students and the settings which they learn in” (Educational Data Mining Society, 2018). It combines knowledge coming from different disciplines, including cognitive science, computer science, cognitive psychology, education, and statistics (Koedinger et al., 2015), to understand how people learn and to help with the creation of tools to support this learning process.

Currently, growing amounts of educational data have been generated daily, and frequently, they are available at a reasonably low cost. One of the reasons is the spread of mobile and online technologies, including hardware and software, that support several users in different educational contexts. When users interact with a digital device or use an online learning environment, this interplay can be captured (logged) and made available for subsequent analysis (R. Baker & Siemens, 2014; Koedinger et al., 2015). For example, students can use interactive learning environments or computer-supported collaborative learning tools to engage with learning materials and increase social bonds among them. Teachers can use technology to keep materials organized, carry out online tests (with automatic assessment), and apply different pedagogical approaches according to students needs. Furthermore, administrators can use several tools to gather data from their schools/universities processes. In all of these situations, EDM can be used to analyze the large volume of data generated by these interactions to detect patterns, which otherwise would be almost impossible to find (Romero & Ventura, 2013).

The availability of large amounts of data, along with other technological factors such as standardized data formats, the increase of computational power, and the development of analytics tools, have been essential for the emergence and widespread adoption of EDM. EDM uniqueness stems from particular characteristics of data coming from educational contexts, such as multilevel hierarchies (e.g., data from students within the context of a course, within the context of an institution) and nonindependence

in data (e.g., an answer to an exercise depends on a previous answer to the same exercise). Issues of time, sequence, and context also play important roles in the study of educational data (Educational Data Mining Society, 2018).

A similar research area is called *learning analytics*, which provides a complementary perspective on the analysis of educational data. Both share similarities and a few differences—EDM is more concerned with automated methods for discovery within the data, whereas learning analytics is more interested in informing humans to use the insights within educational data in an exploratory manner (R. Baker & Siemens, 2014).

According to Romero and Ventura (2013), the application of data mining techniques over LMS data can be seen as an iterative cycle of hypothesis formation, testing,

and refinement—these steps are summarized and exemplified in Figure 1. The generated knowledge can be used not only by educational designers, instructors, and administrators but also by end users—students.

Statistics are at the core of data mining algorithms. Still, when comparing data mining techniques and traditional statistics, one must consider that, in data mining tasks, the goal is to maximize the predictability of the variable of interest, producing reliable future predictions (generalizability). On the other hand, in a more traditional use of this tool, the interpretability of the resulting model matters most in order to extract useful information about how the associations occur between independent and dependent variables (Breiman, 2001; Hand, 2006).

Although outside the scope of this work, certain EDM aspects are to be considered, such as techniques for data

Figure 1. The basic educational data mining process. Adapted with permission from Romero and Ventura (2013). Copyright © 2012 John Wiley & Sons, Inc.

gathering, standardization, and storage; data ownership and control; and forms of analysis (Ellaway et al., 2014). In addition, respective consent has to be obtained, and applicable privacy regulations need to be followed.

Koedinger et al. (2015) made a review of the area of EDM, taking into account the commonly studied applications related to the psychology of learning, dividing the research into four major categories:

- Assessment, modeling of learning, and prediction of achievement: concerns the quantification of learning and performance to predict the probability of correct answers in future problems on the same topic. By using different models, it is possible, for instance, to detect when a student masters a specific skill, avoiding its overpractice, saving time, and allowing him or her to move forward in his or her studies.
- Transfer of learning and discovery of cognitive models from data: aims to create or improve cognitive models of student knowledge or skill from data, as well as to assess their empirical validity. These data-driven models attempt to structure knowledge components from a domain and can be used for the design of improved educational processes so as to yield better learning outcomes.
- Affective states, motivation, and metacognition: affect, motivation, and metacognition indirectly influence learning—EDM methods can be used to model such states. These models may then be incorporated into the LMS for automated interventions (e.g., reengaging bored students).
- Language data and collaborative learning support: analyzes how social processes can impact cognitive processes, mediated by language and collaboration data. Works in this category investigate how the application of machine learning techniques can be used to detect and induce automatic actions to support students' learning as they interact in textual or collaborative environments.

The aforementioned applications are mostly based on the collection of structured data from sources such as LMS and tutoring systems. In addition, Blikstein and Worsley (2016) present a collection of recent studies in learning analytics that aim to explore multimodal data sources (e.g., analysis of text, speech, handwriting, sketch, actions and gestures, affective states, eye gaze, and their integration) as well as data mining techniques to research more complex and open-ended learning environments, such as computer programs, robots, and engineering challenges. These data and techniques are suitable to examine learning in realistic, ecologically valid, social, mixed-media learning environments.

A Taxonomy for EDM Methods

A taxonomy for EDM methods was developed by R. Baker and Siemens (2014), combining EDM, learning

analytics, and other fields' existing approaches that are used to understand educational problems (see Figure 2). Some examples are provided for these methods to illustrate possible applications.

Part II—A Practical Application Example

In this second part of this manuscript, four methods displayed in Figure 2 will be addressed, with examples of applications on an actual online specialization course taught for hearing health care professionals in Brazil during the period of 2015–2017. These post hoc analyses aimed to further explore how teaching and learning processes took place within this course. Some of the insights provided by EDM will also be described.

“Auditory Rehabilitation in Children” Specialization Course

The Auditory Rehabilitation in Children (ARC) was a free-of-charge, e-learning course developed through a partnership among the Speech-Language Pathology and Audiology Department (University of São Paulo–Bauru Campus), the Samaritano Association, and the Technical Area of Health for People With Disabilities–Attention to Health Secretariat, Ministry of Health. It aimed at providing continuing education for speech-language pathologists and audiologists (in Brazil, professionals hold dual certification) and ear, nose, and throat physicians working in public hearing health care facilities accredited by the Brazilian Unified Health System.

The ARC course had a duration of 18 months, with a 400-hr course load (365 hr online and 35 hr face-to-face). A total of 105 participants, working in 68 hearing health care facilities, were selected in a joint process with the state health secretariats across all Brazilian regions. The participants were divided into three groups (G1, G2, and G3), each accompanied by a single tutor throughout the course. These groups were further divided into seven subgroups, of five students each, for performing collaborative tasks.

The tutors supported their respective students' activities, promoting integration as well as monitoring participation and performance as means to address potential issues that could interfere with learning. In addition, 45 instructors (lecturers, researchers, and clinical specialists) were involved in the definition of content, development of instructional material, mediation of the learning activities, and assessing students' performance (formative and summative). All communication and interactions between students and instructors were asynchronous.

ARC was learner-centered, encompassing collaborative, active, and problem-solving methods that enabled reflections about the challenges faced in learners' own clinical practice, while also supporting the construction of knowledge and skills acquisition needed to improve such practice. The courseware was hosted in the Moodle™ LMS (Moodle, n.d.). It consisted of 27 modules, launched as

Figure 2. Taxonomy of educational data mining methods.

Method	Aim	Approaches	Example
Prediction	To estimate the unknown value of a variable by means of a combination of other variables (predictors)	Classification Predicted variable can be binary or categorical.	Prediction of a student's likelihood to pass or fail, complete or drop out of a course, based on other variables (e.g. his/her interactions with instructors and peers).
		Regression Predicted variable is continuous.	Prediction of a student's score based on other variables collected in the learning environment.
		Latent knowledge estimation	Assessment of a student's proficiency on specific skills by his/her patterns of correct performance over time.
Structure discovery	To discover structures not known a priori, seeking to find out structures that emerge naturally from the data	Clustering Finding data points that naturally group together	Grouping students by their behavior profiles in the learning environment.
		Factor analysis Finding variables that naturally group together	Creating measurement scales based on items responded by the students.
		Social network analysis Developing models of the relationships and interactions between individual actors, as well as the patterns that emerge from those relationships and interactions	Understanding communication structure and flow among students and/or instructors.
		Domain structure Finding the structure of knowledge in an educational domain.	Discovering how skills being assessed in a topic could be modeled in a sequence that facilitates learning.
Relationship mining	To find significant relationships among different variables in data sets with a large number of variables.	Association Formulation of if/then rules - if variable y occurs, then variable z occurs.	Discovering conditions in the student's behavior in order to suggest better topics of study.
		Correlation Finding positive or negative linear correlations between variables	Discovering variables that are correlated when a student "games the system" (e.g. systematically guessing answers in order to obtain the correct response).
		Sequential patterns Finding temporal associations between events.	Define student's behavioral patterns that lead to a learning event (e.g. to achieve a given score, a student must watch videos 1 and 2, complete exercises 1-7 and solve his/her doubts through the standby support).
		Causal data Finding whether one event/construct was the cause of another event/construct	Understanding how different variables interact to cause poor learning on a given topic.
Distillation of data for human judgment	To visualize complex data to assist in decision making or in understanding phenomena	Data visualization (heat maps, learning curves, learnograms)	Using learning curves to represent how fast a user is learning a specific topic, avoiding over-practice and enabling the human understanding of patterns in student learning.
Discovery with models	To use an existing EDM/analytics model as a component in a new EDM/analytics	Modeling of constructs	Modeling affective states from the user's actions and use this model to predict learning behaviors in the course.

Figure 3. Suggested sequence for the course modules.

Section		Description
<i>Context</i>		Provision of context to facilitate the attainment of learning goals. Presentation of short texts and/or videos with information about the topic under study as well as short clinical cases with the description of an endpoint for reflection. Administration of surveys and diagnostic assessments, aiming to organize the learning and teaching processes.
<i>Challenge (Problemization)</i>		Scored assignment (usually group assignments). In essence, students were introduced to a problem related to the topic under study and to their work context. Students had autonomy to collect information, build hypothesis and propose possible(s) solution(s). Goals: developing skills in reasoning and self-directed learning, integrating theory and practice, making learning more meaningful, building knowledge collaboratively, developing and widening course contents, among other possibilities.
<i>Content</i>		Videos, tutorials, infographics, guidelines and other didactic and instructional materials developed or selected by the instructors. Goal: present fundamental concepts of the topic under study, from which the knowledge would be built or developed.
<i>Formative assessment</i>		Multiple choice, 'true or false' questions (automated feedback) or essay questions with feedback provided by the instructors.
<i>Frequently asked questions</i>		Compilation of frequently asked questions from the course's previous edition and their respective answers.
<i>Interaction (Asynchronous)</i>	<i>Forum</i>	Participation required. Individually scored. Instructors posted a generating question related to the subject under study. Students should reply to the generating question as well as comment on, at least, two of their colleagues' responses, using the knowledge acquired/developed on that module to support their arguments. This activity was configured so that a given student could only read a colleague's response after posting his/her own. Instructors stimulated the discussions, mediated different points of view and provided feedback to each student in a timely manner, by encouraging the deepening of correct responses and offering opportunities for revision of incorrect answers.
	<i>Standby support</i>	Non scored. Voluntary participation. Students could, at any time, ask the instructors questions related to the subject under discussion in a given module or related to the use of the tools available in the LMS. The open format also allowed students to read questions and answers exchanged among participants.
<i>References</i>		Bibliographic references, external links to video classes and websites of interest to the topic under study.
<i>Library</i>		Collaborative space, created and maintained by students, with references and contents found by them - these content were also analyzed by the teachers, giving more reliability to the shared material.
<i>Self-Assessment</i>		Self-assessment questionnaire answered by the student, composed of six items: attendance, punctuality, autonomy, engagement.
<i>Module's Assessment</i>		Assessment of student's satisfaction regarding contents and activities, the performance of tutors and instructors, visual communication in didactic materials as well as navigation and technical difficulties with the LMS.

the course progressed, covering topics from hearing loss identification to intervention. For each module, a designed instructional sequence, composed of 10 different sections, was suggested (see Figure 3). However, the navigation was open, and the student could access the different sections of

a module in any order. The nature and design of instructional materials, activities, and assignments varied according to each module's learning goals.

Ninety-three students (95.7% women, 4.3% men; 96.7% speech-language pathologists and audiologists; avg.

age = 36.3, $SD = 7.7$) successfully completed the course, and their responses and behaviors were considered in the present work. The data set is composed of the log records registered by the LMS and the results of validated self-assessment instruments, where specified.

Prediction

Learning strategies are mediators of learning outcomes, developed by individuals during their learning activities to control their own psychological processes, such as attention, acquisition, memorization, and transference (Zerbini & Abbad, 2008). They are procedures used by individuals in learning activities to ensure they meet all of the criteria set out at the beginning of the instructional practice (Warr & Allan, 1998).

From a theoretical framework (Warr & Allan, 1998), the Learning Strategies Scale (LSS) was created (Warr & Downing, 2000) and, later, adapted and validated to the Brazilian context (Zerbini & Abbad, 2008). The LSS consists of 28 statements divided into seven factors: emotion control, motivation control, interpersonal help-seeking, rehearsal and organization, elaboration, help-seeking from written material, and comprehension monitoring. For each statement, responses are given on a Likert scale, ranging from 0 (*never*) to 10 (*always*). The individual score—here, the variable to be predicted—is the average of the responses to the items related to each factor and the overall scale.

According to the process shown in Figure 1 (Step 1), the following question was generated: “Can a model be developed to predict interpersonal help-seeking behavior from students’ actions in the course?”

In the ARC context, the LSS was administered to 93 students at course completion. In this work, only the interpersonal help-seeking factor (six items, Cronbach’s $\alpha = .89$) was chosen to be predicted. Interpersonal help-seeking behavior is considered an important indicator of students’ engagement with the course, and in many cases, it is associated with students’ dropout rates—that is, the more help a student seeks, the better are the chances of him or her finishing the course successfully (Karabenick & Zusho, 2015). It is a self-regulatory strategy, which assesses one’s own learning process, being particularly important in e-learning, in which the interaction between student and instructor may be more sparse. Because self-regulatory strategies can be learned (Zerbini & Abbad, 2008), we sought to build a prediction model to support the monitoring of students’ interpersonal help-seeking behavior and to provide means so that it can be better developed throughout the course (Penteado, Isotani, Paiva, Morettin-Zupelari, & Ferrari, 2018).

In order to operationalize interpersonal help-seeking construct, three sources inside the course were chosen: the challenge with wiki-based activity, the forum, and the standby support (see Figure 3). The wiki tool allows for group members to edit and develop content collectively; for this reason, it was often used in case of group assignments. Only modules with all these three sources were analyzed ($n = 14$).

The behavioral metrics were derived and computed from the log files, originally recorded by the LMS. Table 1 shows the features considered for modeling (corresponding to Figure 1, Step 2).

Some preprocessing steps (see Figure 1, Step 3) were needed to properly balance the data, which was highly skewed (Penteado et al., 2018). Weka¹ tool was selected to run the analysis with some state-of-the-art regression algorithms of different families. The aggregated data were formatted for this tool (see Figure 1, Step 4). The ZeroR algorithm was used as the baseline, which only outputs the mean value of the sample, discarding information from all variables. This is an important step in the data analysis process, in order to evaluate the usefulness of the prediction over a simple benchmark. In our sample, the mean value for the interpersonal help-seeking factor was 8.3 (min = 4.5, max = 10, $SD = 1.2$) with a mean absolute error of 1.158, meaning most students reported frequent use of interpersonal help-seeking strategies.

The respondents’ scores were divided into two sets (67% for training the model and 33% for testing it), and the model was trained using k -fold cross-validation ($k = 10$). The cross-validation method alleviates the problem of overfitting the model to the training data, hindering generalizability to new data. The steps of selecting algorithms and performing the training test split are central to the data mining method (see Figure 1, Step 5).

The most common metrics used to analyze the performance of regression algorithms were selected: the mean absolute error, the root mean square error, and the Pearson correlation between true and predicted scores. The results (see Table 2) were computed by applying the model obtained from the training data to the unseen test data (see Figure 1, Step 6).

Linear regression was chosen to explain the results, because it is more interpretable than the other algorithms and because it also showed good results, including the best correlation. The resulting model is shown in Figure 4.

In our results, the models could predict the interpersonal help-seeking scores with a variation of, on average, 1 point in the scale and approximately 7% of improvement over baseline (see Table 2). The baseline presented low errors in the selected metrics, indicating a highly concentrated distribution of the self-reported scores. The correlation of .34 from the logistic regression model can be considered moderate in some areas, particularly in natural settings of social sciences, in which multiple variables influence the outcome.

Improvement of this model could lead to an unobtrusive monitoring of the interpersonal help-seeking strategies used by the students throughout the course, thus helping instructors to detect those who have developed this behavior, thus favoring the construction of their knowledge or those who are in need of assistance (see Figure 1, Step 7).

¹<https://www.cs.waikato.ac.nz/ml/weka/>

Table 1. Features selected for the modeling process.

Forum	Standby support	Wiki	Other
Average length of posts	Average length of posts	Average number of edits	Age
Average time spent on the tool	Average time spent on the tool	Average time spent on the tool	—
Number of posts' visualizations	Number of posts' visualizations	Number of pages' visualizations	—
Number of posts	Number of posts	Average number of characters in the edits	—
Average scores	Median of days for posting questions	Average scores	—
—	—	Median of days for editing the wiki page	—

It must be noted these are insights from an exploratory phase, and several new hypotheses emerged (see Figure 1, Step 8), for instance, on what point of the course would the application of such model be reliable? Are the skewed results obtained due to limitations of the LSS? Further analyses are needed in order to answer such questions and refine the model.

Structure Discovery

Interpersonal interaction is a central component in education, and its positive effects on learning outcomes have been argued by different learning theories, such as collaborative learning (Johnson, Johnson, & Smith, 2007). In addition, interaction fosters engagement, and this has been successfully used in models to predict attrition (Bayer et al., 2012; Macfadyen & Dawson, 2010).

In the ARC course, the Forum assignment (see Figure 3) was one of the tools for sharing of knowledge, ideas, and experiences, related to the topic under study, among the students. For this reason, it was important to identify how the interaction occurred within all groups in this activity and who were the most active students.

In this work, the analysis was carried out only for the module “Outcome measures in pediatric auditory rehabilitation.” The participants were asked to comment on their experience regarding the introduction of aided speech perception measures in their work environment (e.g., factors that helped or hindered). In addition to sharing their experience, students were also asked to comment on at least two of their colleagues' posts.

Social networks analysis (SNA) method was used. This flexible method is suitable to discover structural characteristics, relationships, and flows between different entities

(people, computers, web pages, organizations), connected by some sort of interaction between them. It models a problem by setting out entities and their relationships and uses graph theory to represent and compute its metrics. It maps the entities as nodes in a graph and the relations as edges.

In this scope, the context of forum interactions was modeled with nodes representing students, tutors, or instructors that commented on this forum and edges representing a reply given by one person to another previous post (a directed relation). The results of the modeling are depicted in Figure 5—the thickness of an edge represents the number of messages exchanged between two people. From Figure 5B, it can be seen that the edge can be directed, having a meaning of *A follows B*, but not the other way around. When directions are not present, they have an identical meaning in both ways.

Figure 5A shows all students and instructors that interacted in this forum. At the center of the network is one of the instructors responsible for the module. It was possible to distinguish the strong influence of the instructors, acting as hubs among the students, suggesting they are mediating the discussions. When instructors were removed from the graph (Figure 5B), some strong edges between certain students, who exchanged large numbers of comments among them, were noted. This information is relevant because it may indicate the engagement of such students in the course. The level of student's cohesion within a given group was also observed. A few students only replied to the generating question and did not interact directly with their peers at all—the isolated nodes in the periphery. SNA can guide instructors' intervention in order to make participation of these students more effective, bringing them closer to their colleagues' discussions, thus facilitating interpersonal interaction and learning.

Furthermore, a set of metrics can be computed, for the nodes (in- and out-degrees, centrality, betweenness, closeness), the edges (weights), and the graph (density, diameter, connected components). These metrics can, in turn, be used to create software components to respond automatically in a given situation. For instance, a software component may act as a learning assistant, providing the student with feedback in real time, thus creating opportunity for reflection and, possibly, change in his or her interpersonal interaction behavior. Alternatively, some LMSs (e.g., Moodle™) have rewarding mechanisms,

Table 2. Summary of the performances using the test set.

Algorithm	MAE	RMSE	Correlation
ZeroR	1.158	1.337	.000
Linear regression	1.088	1.354	.339
Support vector machine for regression	1.025	1.233	.244
Random forest	1.075	1.266	.233
M5 rules	1.072	1.326	.220

Note. MAE = mean absolute error; RMSE = root mean square error.

Figure 4. Resulting model of the learning regression algorithm.

$$\begin{aligned} \text{IHS} = & (0.0057 * \text{number of visualizations in standby support}) \\ & + (0.0019 * \text{average number of characters in posts of standby support}) \\ & + (-0.0067 * \text{average time, in seconds, spent on forum}) \\ & + (0.0849 * \text{the median of days of editing in Wiki, before due date}) \\ & + 7.3321. \end{aligned}$$

such as badges, that can be programmed to be awarded either automatically or by tutors once a student meets completion criteria (Deanmade, 2015). This could be used to stimulate the behavior of communicating with peers.

It must be noted that, although SNA reveals the structure of interactions, it does not detail the nature of exchanges or the impact of connectedness. The combination of SNA with other approaches (e.g., discourse analysis) can help with this matter (R. S. J. d. Baker, 2010).

Relationship Mining

As previously mentioned, in each module of the ARC course, an instructional sequence was suggested for all the students (see Figure 3). Still, navigation within a module as well as within each section, when applicable, was open. This form of navigation offered students more autonomy and responsibility to choose their own learning path, according to their needs, prior knowledge, skills, and experiences, among others. Furthermore, as per University of São Paulo regulation, a minimum of 75% of attendance (e.g., access to contents or activities available in a given module) was required.

In this example, we sought to understand more about the students' learning paths, through the analysis of most frequent sequences of web pages (e.g., activities, videos, texts, forum, quizzes) visited in a given module. In other words, we wanted to find patterns of recurring behaviors when the students navigated the LMS.

Sequential pattern analysis aims to find relationships between occurrences of sequential events, that is, to find if any specific order of the occurrences exists. For instance, a student in a session visits pages A and B and then, in a second session, visits C, D, and E. They are considered two transactions and are represented as $\langle A, B \rangle$, $\langle C, D, E \rangle$, in this order. This arrangement supports subsequences such as AC, AD, AE, BC, BD, and BE, with the interpretation “after visiting A, the user visited C” and so on. In this sort of analysis, we try to find ordered subsequences with a minimum threshold of occurrence in the data set, also called *support*—the proportion of time that a sequence occurred. Also, we can determine how much noise we can admit in considering that two subsequences are the same: $\langle A \rangle \langle B \rangle$ and $\langle A \rangle \langle C \rangle \langle B \rangle$ may be considered the same sequences with $\langle C \rangle$ as noise—so we adjust this “gap” accordingly.

For this analysis, we chose the module “Pediatric auditory diagnosis—Part 1,” consisting of 18 different pages that could be navigated freely, in any order. The generalized sequential pattern algorithm, implemented in the RapidMiner^{TM2} tool, was applied (Hofmann & Klinkenberg, 2013). In this case, only the order in which the events occurred was considered, not their actual time stamp. We considered the user as unit of analysis, pages as the items abstraction, and every action in the LMS as a separate transaction (window size = 0), considering a gap of up to 10 transactions to be computed (maxgap = 10) and filtering out those transactions that occurred too early in time (mingap = 2), avoiding duplicity, because it raised a lot of noise where the same tool had different pages. The minimum support was set to 0.7, that is, it should bring the subsequences that occurred for at least 70% of the users.

Table 3 summarizes the output sequences discovered. These results must be analyzed cautiously. They are sequenced order of events with higher support in the data set than the threshold set (0.7). Depending on the parameters chosen, a lot of redundant or obvious rules may appear. A postprocessing step is often needed to select only those associations that carry relevant knowledge.

In our example, five out of 13 sequences were obvious choices, because they represented sequential pages in the module's layout, that is, the students followed the suggested instructional sequence. One example is the sequence $\langle \text{Video Overview of Behavioral Assessments} \rangle \langle \text{Video Visual Reinforcement audiometry} \rangle$, which occurred for 98% of students at least once.

On the other hand, alternative patterns were found, representing behavior common to several students. For instance, the “Forum” was present in six out of the eight remaining sequences. For this Forum, participants were asked to critically discuss the physiological and behavioral auditory assessment pediatric protocol employed in the clinic in which they worked, including proposals for improvements, when necessary. This may suggest that this assignment was very important and central because most users navigated from and to it.

Instructors can analyze which trajectories led to more effective learning (R. Baker et al., 2011). Monitoring and interpreting sequential learners' activities have the potential

²<https://rapidminer.com>

Figure 5. (A) Communication structure between the Forum’s participants: Each node represents a student in the graph, and each exchanged message is represented by a directed link between the people involved. (B) Communication structure resulted from removing instructors from the analysis, making the three groups visible.

to improve adaptivity and personalization within educational environments, both for improving the course design or by automatically recommending actions to the students.

Distillation of Data for Human Judgment

The collaborative construction of knowledge was one of the goals of the group assignments (“Challenge” activity) within the ARC course. Traditionally, such assignments are assessed only via analysis of their final product (e.g., coherence with the proposed task, achievement of the learning goals, and accuracy and completeness). However, in order to support the course’s pedagogical model,

assessment should also take into account the work process. The understanding of how online collaboration takes place makes the planning of strategies for supporting students to achieve collaborative skills easier (MacDonald, 2003). Distillation of data for human judgment brings tools, particularly in the form of data visualization, to help stakeholders (e.g., instructors, students, administrators) to explore and understand educational data. In this example, the module “Outcome measures in pediatric auditory rehabilitation” was selected. The “Challenge” assignment, destined for the subgroup of five students, consisted of writing a critical report from a series of videos where professional actors reenacted audiologist–parent interactions for the administration of outcome measures (e.g.,

Table 3. Sequences found in the data set of the module under study, ordered by their support.

Support	Temporal sequences
0.986	⟨Video Overview of Behavioral Assessments⟩ ⟨Video Visual Reinforcement audiometry⟩
0.917	⟨Self-assessment⟩ ⟨Module assessment⟩
0.861	⟨Video Overview of Behavioral Assessments⟩ ⟨Video Immittance measurements in babies⟩*
0.778	⟨Text Main considerations for immittance measurements in babies⟩ ⟨Forum⟩*
0.736	⟨Video Overview Behavioral Assessments⟩ ⟨Forum⟩*
0.736	⟨Text Assessment of auditory responses to speech⟩ ⟨Forum⟩*
0.736	⟨Forum⟩ ⟨Challenge/Group activity⟩*
0.722	⟨Video Visual Reinforcement Audiometry⟩ ⟨Text Main considerations for immittance measurements in babies⟩*
0.722	⟨Video Immittance measurements in babies⟩ ⟨Text Main considerations for immittance measurements in babies⟩
0.722	⟨Challenge/Group activity⟩ ⟨Forum⟩*
0.722	⟨Forum⟩ ⟨Self-assessment⟩
0.708	⟨Text Main considerations for immittance measurements in babies⟩ ⟨Text Assessment of auditory responses to speech⟩
0.708	⟨Introduction⟩ ⟨Video Overview of Behavioral Assessments⟩

Note. The sequences indicated by (*) represent transitions in an order different than suggested by the course design.

Figure 6. Wiki document evolution along the weeks for some groups (left). On the right, a particular group (“3C”) strategy is shown. In general, the peaks represent copy-and-past behavior, where a student commented upon a colleague’s work; the valleys represent the synthesis of ideas, by trimming the content.

inventories to assess auditory and language development). As this was a scored group assignment, all students were supposed to contribute to the report construction. We sought to understand the strategies adopted by the subgroups in order to accomplish the construction of such report (e.g., how writing and editions were negotiated among group members, how individual contributions were handled).

The Moodle™ native wiki tool was used—this tool includes features that are designed to facilitate collaboration among group members by creating and editing their reports. The wiki records every action taken by the students in the creation of the shared document—essentially, who made edits, what was edited, and by whom and when. The number of characters in the document, throughout the module’s duration, was also monitored.

Figure 6 (left) shows examples of four groups—one can see the diverse strategies used, both in the time taken and written content. The data suggest that some groups, such as 3C, whose strategy is detailed in the right panel (see Figure 6), adopted a more collaborative approach, with engagement of all participants in making contributions, in an effort to accomplish the task together (McInnerney & Roberts, 2004). Other groups, such as 1D, seem to have chosen a cooperative approach, possibly splitting up the task and working individually and then combining the results when they got closer to the deadline. Last but not least, in some groups (e.g., 3G), contributions were made very late in the task, making the possibility of collaboration unlikely.

The examples in Figure 6 also indicate that the designing of collaborative learning activities and employment of the so-called collaborative tools are not an assurance that collaboration will indeed take place. Methods for data visualization can be used to infer how different patterns emerge from a complex setting of interaction. However, they are limited to a quantitative analysis of the document, disregarding the quality of what is written. For instance,

some changes may be aesthetic, duplication of excerpts, or comments. Other EDM techniques, such as text mining, could be used to infer more features from the text and complement the analysis (Koedinger et al., 2015). A more comprehensive analysis should use both forms, quantitative and qualitative, to triangulate the results toward some instructional goal.

Concluding Comments

This article introduced concepts and applications and described quantitative methods from the EDM area. It also provided some insights into analysis and interpretation of real data through EDM methods.

EDM involves several steps, with different professional skill sets: technological knowledge (e.g., for database structure and storage, data transformation, statistical analysis, and software design), domain knowledge (in our example, audiology), and pedagogical knowledge (e.g., for analysis modeling, theoretical contextualization, and interpretation of the outcomes). Audiology educators interested in using EDM should at least understand enough of the methods and their purposes to be able to make or support a choice that is most suitable for a given context/research question, as well as how to interpret the results.

On what relates to the tools used for EDM, most academic articles use general purpose free data mining tools (e.g., RapidMiner™, R [R Core Team, 2013],³ Weka [Hall et al., 2009], and KEEL [Alcalá-Fdez et al., 2011])⁴, equipped with components such as preprocessing tools, state-of-the-art algorithms, statistical validations, and visualizations. Other tools, such as Gephi™ (Bastian, Heymann,

³<http://www.r-project.org/>

⁴<http://www.keel.es>

& Jacomy, 2009)⁵ for graph visualization and calculation of metrics for social network analysis, Tableau⁶ (Wesley, Eldrige, & Terlecki, 2011) for building analytic reports and data visualization, and LightSIDE⁷ (Mayfield & Rosé, 2013) for text mining in English, have also been standing out for being free of charge and meeting varied purposes.

In summary, EDM offers new tools to discover knowledge from digital traces and has the potential to foster a better understanding and management of processes in audiology distance education. In addition, the most reliable evaluation of an EDM discovery is whether it yields better student learning when applied. Such an evaluation has been referred to as “closing the loop,” as it completes the cycle of system design, deployment, data analysis, and discovery leading back to design (Liu & Koedinger, 2017). First, statistical models are created from data; then, adaptive systems are built to use these statistical models; and finally, the experiments are executed, comparing the original system to the adaptive system, which is used to extend and refine the underlying models and theory and design new or improved features.

The methods approached in this article can be used or combined with others, so that software components may automatically detect certain situations in the learning environment to intervene properly, as in the case of intelligent tutoring. In the field of intelligent tutoring systems, a lot of effort has been put to automatic interventions such as personalized feedback, prevention of hint abuse, and fine-grained assessments, among others (for a comprehensive list of examples, please see Van Lehn, 2006). Still, because not all instructional situations can be properly contextualized automatically, another solution is to provide the course instructor with the knowledge discovered, thus supporting his or her decision making. Furthermore, this knowledge can be used by instructors to refine the instructional design of a course for its subsequent offering, as is the case with the ARC course.

To the best of our knowledge, there is no research exploring EDM to educational processes in audiology. The authors hope the reader will be motivated to develop this area in audiology.

Acknowledgments

The Auditory Rehabilitation in Children (ARC) course was funded by the Brazilian Ministry of Health, Support Program for Institutional Development of the National Health System (Proadi/SUS Grant 25000.024953/2015-89).

The authors thank the Samaritano Association for managing the Proadi/SUS Grant and for the administrative coordination of the ARC course.

The authors Bruno Elias Penteado and Paula Maria Paiva thank the Coordination for the Improvement of Higher Education Personnel for the provision of scholarship for their graduate studies.

Portions of this article were presented at the 3rd International Internet & Audiology Meeting, Louisville, KY, July 2017, which was funded by National Institute on Deafness and Other Communication Disorders (NIDCD) Grant 1R13DC016547 and the Oticon Foundation.

Bruno Elias Penteado also thanks the 3rd International Internet & Audiology Meeting for the provision of the travel award.

References

- Alcalá-Fdez, J., Fernandez, A., Luengo, J., Derrac, J., García, S., Sánchez, L., & Herrera, F. (2011). KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing*, 17(2–3), 255–287. Retrieved from <http://www.keel.es>
- Al-Samarraie, H., Selim, H., Teo, T., & Zaqout, F. (2017). Isolation and distinctiveness in the design of e-learning systems influence user preferences. *Interactive Learning Environments*, 25(4), 452–466. <https://doi.org/10.1080/10494820.2016.1138313>
- Al-Shorbaji, N., Atun, R., Car, J., Majeed, A., & Wheeler, E. (2015). *eLearning for undergraduate health professional education: A systematic review informing a radical transformation of health workforce development*. Retrieved from <http://whoeducationguidelines.org/sites/default/files/uploads/eLearning-healthprof-report.pdf>
- Baker, R., Isotani, S., & Carvalho, A. (2011). Educational data mining: Opportunities for Brazil. *Revista Brasileira de Informática Na Educação*, 19(02). <https://doi.org/10.5753/RBIE.2011.19.02.03>
- Baker, R., & Siemens, G. (2014). Educational data mining and learning analytics. In R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (Cambridge Handbooks in Psychology) (pp. 253–272). Cambridge, United Kingdom: Cambridge University Press. <https://doi.org/10.1017/CBO9781139519526.016>
- Baker, R. S. J. d. (2010). Data mining for education. *International Encyclopedia of Education*, 7, 112–118.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media. Retrieved from <https://gephi.org>
- Bayer, J., Bydzovská, H., Géryk, J., Obsivac, T., & Popelínský, L. (2012). Predicting drop-out from social behaviour of students. *Proceedings of the 5th International Conference on Educational Data Mining*, 103–109.
- Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*, 3(2), 220–238. <https://doi.org/10.18608/jla.2016.32.11>
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199–231. <https://doi.org/10.2307/2676681>
- Crandell, C. C., & Weiner, A. (2002). Counseling competencies in audiologists: Efficacy of a distance-learning course. *Hearing Journal*, 55(9), 42–47. <https://doi.org/10.1097/01.HJ.0000293929.72235.86>
- Deanmade, N. (2015). *Gamification with Moodle*. Birmingham, United Kingdom: Packt Publishing Ltd.
- Educational Data Mining Society. (2018). *Home page*. Retrieved from <http://educationaldatamining.org/>
- Ellaway, R. H., Pusic, M. V., Galbraith, R. M., & Cameron, T. (2014). Developing the role of big data and analytics in health professional education. *Medical Teacher*, 36(3), 216–222. <https://doi.org/10.3109/0142159X.2014.874553>

⁵<https://gephi.org/>

⁶<http://www.tableau.com/products/desktop>

⁷<http://www.cs.cmu.edu/~cprose/LightSIDE.html>

- English, K., Rojeski, T., & Branham, K. (2000). Acquiring counseling skills in mid-career: Outcomes of a distance education course for practicing audiologists. *Journal of the American Academy of Audiology*, 11(2), 84–90.
- George, P. P., Papachristou, N., Belisario, J. M., Wang, W., Wark, P. A., Cotic, Z., ... Car, J. (2014). Online eLearning for undergraduates in health professions: A systematic review of the impact on knowledge, skills, attitudes and satisfaction. *Journal of Global Health*, 4(1), 1–17. <https://doi.org/10.7189/jogh.04.010406>
- Goulios, H., & Patuzzi, R. B. (2008). Audiology education and practice from an international perspective. *International Journal of Audiology*, 47(10), 647–664. <https://doi.org/10.1080/14992020802203322>
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *SIGKDD Explorations*, 11(1), 10–18. Retrieved from <https://www.cs.waikato.ac.nz/ml/weka/>
- Hand, D. J. (2006). Classifier technology and the illusion of progress. *Statistical Science*, 21(1), 1–14. <https://doi.org/10.1214/088342306000000060>
- Hofmann, M., & Klinkenberg, R. (2013). *RapidMiner: Data mining use cases and business analytics applications*. Boca Raton, FL: CRC Press. Retrieved from <https://rapidminer.com>
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2016). *Censo da Educação Superior 2016: Notas estatísticas* (p. 17). Retrieved from http://download.inep.gov.br/educacao_superior/censo_superior/documentos/2016/notas_sobre_o_censo_da_educacao_superior_2016.pdf
- Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, 19(1), 15–29. <https://doi.org/10.1007/s10648-006-9038-8>
- Karabenick, S. A., & Zusho, A. (2015). Examining approaches to research on self-regulated learning: Conceptual and methodological considerations. *Metacognition and Learning*, 10(1), 151–163. <https://doi.org/10.1007/s11409-015-9137-3>
- Koedinger, K. R., D'Mello, S., McLaughlin, E. A., Pardos, Z. A., & Rosé, C. P. (2015). Data mining and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(4), 333–353. <https://doi.org/10.1002/wcs.1350>
- Liu, R., & Koedinger, K. R. (2017). Closing the loop: Automated data-driven cognitive model discoveries lead to improved instruction and learning gains. *Journal of Educational Data Mining*, 9(1), 25–41.
- MacDonald, J. (2003). Assessing online collaborative learning: Process and product. *Computers and Education*, 40(4), 377–391. [https://doi.org/10.1016/S0360-1315\(02\)00168-9](https://doi.org/10.1016/S0360-1315(02)00168-9)
- Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers and Education*, 54(2), 588–599. <https://doi.org/10.1016/j.compedu.2009.09.008>
- Mayfield, E., & Rosé, C. P. (2013). LightSIDE: Open source machine learning for text accessible to non-experts. In M. D. Shermis, & J. Burstein (Eds.), *Handbook of Automated Essay Grading* (pp. 124–135). New York, NY: Routledge Academic Press.
- McInerney, J. M., & Roberts, T. S. (2004). Collaborative or cooperative learning? In T. Roberts (Ed.), *Online collaborative learning: Theory and practice* (pp. 203–214). Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-59140-174-2.ch009>
- Moodle. (n.d.). Moodle (Version 3.5). Retrieved from <http://moodle.org>
- Penteado, B. E., Isotani, S., Paiva, P. M., Morettin-Zupelari, M., & Ferrari, D. V. (2018). Prediction of interpersonal help-seeking behavior from log files in an in-service education distance course. In Rosé C. Penstein et al. (Eds.), *Artificial Intelligence in Education*. AIED 2018. Lecture Notes in Computer Science, Vol. 10948. Springer, Cham.
- Poulin, R., & Straut, T. (2016). *WCET distance education enrollment report 2016*. Retrieved from <http://wcet.wiche.edu/initiatives/research/WCET-Distance-Education-Enrollment-Report-2016>
- R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org/>
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12–27. <https://doi.org/10.1002/widm.1075>
- Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers and Education*, 51(1), 368–384. <https://doi.org/10.1016/j.compedu.2007.05.016>
- Schaeffer, C. E., & Konetes, G. D. (2010). Impact of learner engagement on attrition rates and student success in online learning. *International Journal of Instructional Technology & Distance Learning*, 7(5), 3–9.
- Silva, A. d. N., Santos, A. M. G. d., Cortez, E. A., & Cordeiro, B. C. (2015). Limits and possibilities of distance education (EAD) in the permanent education in health: Integrative review. *Ciência & Saúde Coletiva*, 20(4), 1099–1107. <https://doi.org/10.1590/1413-81232015204.17832013>
- Sinclair, P. M., Kable, A., Levett-Jones, T., & Booth, D. (2016). The effectiveness of Internet-based e-learning on clinician behaviour and patient outcomes: A systematic review. *International Journal of Nursing Studies*, 57, 70–81. <https://doi.org/10.1016/j.ijnurstu.2016.01.011>
- Staikopoulos, A., & Conlan, O. (2017). Proposing an evaluation task for identifying struggling students in online courses. In *Adjunct publication of the 25th conference on user modeling, adaptation and personalization—UMAP '17* (pp. 203–208). New York, NY: ACM Press. <https://doi.org/10.1145/3099023.3099049>
- Van Lehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, 16(3), 227–265.
- Warr, P., & Allan, C. (1998). Learning strategies and occupational training. *International Review of Industrial and Organizational Psychology*, 13(2), 83–121.
- Warr, P., & Downing, J. (2000). Learning strategies, learning anxiety and knowledge acquisition. *British Journal of Psychology*, 91(Pt. 3), 311–333. <https://doi.org/10.1348/000712600161853>
- Wesley, R. M. G., Eldridge, M., & Terlecki, P. (2011). An analytic data engine for visualization in tableau. *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data* (pp. 1185–1194). New York, NY: ACM. Retrieved from <http://www.tableau.com/products/desktop>
- Windmill, I. M., & Freeman, B. A. (2013). Demand for audiology services: 30-yr projections and impact on academic programs. *Journal of the American Academy of Audiology*, 24, 407–416.
- World Health Organization. (2013a). *A universal truth: No health without a workforce third global forum on human resources for health report*. Geneva, Switzerland: Author.
- World Health Organization. (2013b). Transforming and scaling up health professionals' education and training: World Health Organization guidelines 2013. *Guidelines*, 124. <https://doi.org/10.1613/jair.301>
- Zerbini, T., & Abbad, G. (2008). Strategies for distance learning: Validation of a scale. *Psico-USF*, 13(2), 177–187. <https://doi.org/10.1590/S1413-82712008000200005>