

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE ESCUELA DE INGENIERÍA

PROCESS-ORIENTED METRICS THAT ASSESS THE PERFORMANCE OF STUDENTS WHO ARE LEARNING SURGICAL PROCEDURES: THE PERCUTANEOUS DILATATIONAL TRACHEOSTOMY CASE JUAN JOSÉ MARTÍNEZ

Thesis submitted to the Office of Research and Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science in Engineering

Advisor: MARCOS SEPÚLVEDA

Santiago de Chile, September 2021

© MMXXI, JUAN JOSÉ MARTÍNEZ



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE ESCUELA DE INGENIERÍA

PROCESS-ORIENTED METRICS THAT ASSESS THE PERFORMANCE OF STUDENTS WHO ARE LEARNING SURGICAL PROCEDURES: THE PERCUTANEOUS DILATATIONAL TRACHEOSTOMY CASE JUAN JOSÉ MARTÍNEZ

Members of the Committee: MARCOS SEPÚLVEDA JORGE MUÑOZ-GAMA EDUARDO KATTAN MAURICIO LÓPEZ

Thesis submitted to the Office of Research and Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science in Engineering

Santiago de Chile, September 2021

© MMXXI, JUAN JOSÉ MARTÍNEZ

To my parents and grandparents, my family and my friends that supported me all these years

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Marcos Sepúlveda, for the continuous support, encouragement and patience. Since I started the thesis, he gave me the opportunity to choose an area of research that suited my interests, giving me all the tools and knowledge to do so. He always trusted in my abilities and helped me to improve them immensely thanks to his devotion for his work and the motivation that produces in the others. Always with a sincere sense of humor and a great empathy.

I would also like to thank René de la Fuente for his great work in this research and his patience. He always had the kind attitude of looking for ways to improve the research so that I could increase my knowledge. His constant questions motivated me to look for more challenges to improve my thesis. He proved to be a great professional.

Thanks to Jorge Munoz, who helped me professionalize my research through the new tools that I learned in his course. I also want to thank the time he gave in the development of the research to improve the concepts worked in this thesis and his willingness to collaborate solving my doubts.

Thanks to Victor Galvez, who was always there when I needed help. His previous jobs served me a lot as a motivational factor and as a support guide. Victor has a great ability to support and also teach, delivering great value to the privileged who work alongside him.

Finally, I would like to thank Eduardo Kattan, who carried out the guiding research for my thesis and helped me to approve the basic concepts for it. And to Luis Leiva for his project, which was the technical support of the research.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	vii
LIST OF TABLES	viii
ABSTRACT	ix
RESUMEN	X
1. CHAPTER 1	1
1.1. Context	2
1.2. Research thesis	3
1.2.1. Research hypothesis	3
1.2.2. Research objectives	3
1.3. Background	3
1.3.1. Classic metrics in training	4
1.3.2. Feedback in medical education	4
1.3.3. Business Process Modeling	5
1.3.4. Process Mining	6
1.4. Materials and methods	7
1.4.1. Model definition	7
1.4.2. Data collection	7
1.4.3. Procedure analysis	9
1.4.4. Research's paper	9
2. CHAPTER 2	10
2.1. ABSTRACT	10
2.2. INTRODUCTION	11
2.3. MATERIAL AND METHODS	13

2.3.	1. Model Definition	13
2.3.	2. Data collection	13
2.3.	3. Procedure analysis	15
2.4.	RESULTS	18
2.4.	1. Process-oriented metrics at high level	18
2.4.	2. Process-oriented metrics by stage	18
2.4.	3. Process-oriented metrics in activities	20
2.4.	4. Progress of process-oriented metrics in activities	22
2.4.	5. Comparison between process-oriented metrics and classic metrics	23
2.5.	DISCUSSION	24
3. CHA	PTER 3	28
3.1.	Conclusions	28
3.2.	Limitations	29
3.3.	Future work	29
REFERE	NCES	31

LIST OF FIGURES

1.1	BPMN example	5
1.2	Main Process Mining techniques(Source: Process Mining Manifesto)	7
2.1	Reference model used in the study (Adapted from the original BPMN model)	14
2.2	Process-oriented metric, grouped by session	19
2.3	Process-oriented metric, aggregated by session, broken down by stage	19
2.4	Process-oriented metrics for selected stages, broken down by activities	21
2.5	Evolution of all process-oriented metrics in the activities	22
2.6	Evolution of the aggregated process-oriented metric and the classic metrics .	23

LIST OF TABLES

2.1	Process metric definition	16
2.2	Correlation results between classic metrics and process-oriented metrics	23

ABSTRACT

Assessing competency in performing a surgical procedure training is key for instructors to distinguish whether a resident is qualified to perform a surgical procedure. Currently assessment techniques provide a global result and do not provide feedback focused on the procedural stages or on the order in which activities need to be performed. For optimal feedback, it is recommended that it has to be very specific so residents can detect exactly where they are making mistakes in the most objective way possible.

Process Mining is a discipline that generates knowledge from data of the executions of processes registered in information systems, facilitating the analysis of these processes. Since surgical procedures can be understood as a progression of steps, they can be viewed also as a process. In this research, three process-oriented metrics are proposed to generate a more detailed analysis of the variability with which residents carry out their training sessions. This allows obtaining feedback from a global level of the procedure, at a level of stages and finally, at the level of activities.

The research was carried out on the data of a Percutaneous Dilatational Tracheostomy (PDT) surgery training study. Several visualizations were delivered that allow understanding the progress of each resident through their sessions from a process approach. Finally, the behavior of the process-oriented metrics was compared to the classical evaluation methods. In the future, it is expected to be able to validate these metrics for professional use as an evaluation method and as a complement of the current ones.

Keywords: Process Mining, Medical Training, Feedback, Technical Skills, Surgical Procedures.

RESUMEN

Evaluar las competencias en la realización de un entrenamiento de procedimientos quirúrgicos es clave para que los instructores puedan distinguir si un residente está calificado para realizar un procedimiento quirúrgico. Actualmente, los métodos de evaluación en simulación brindan un resultado global y no brindan un feedback centrado en las etapas del procedimiento o en el orden en que se deben realizar las actividades. Para que el feedback sea óptimo, se recomienda que este sea específico, para que así los alumnos detecten exactamente donde cometen errores de la forma más objetiva posible.

Process Mining es una disciplina que genera conocimiento a partir de datos de las ejecuciones de procesos registrados en sistemas de información, facilitando el análisis de dichos procesos. Dado que los procedimientos quirúrgicos se pueden entender como una progresión de pasos, estos pueden ser vistos como un proceso. En esta investigación se proponen tres métricas orientadas a procesos para generar un análisis más detallado de la variabilidad con que los residentes realizan sus entrenamientos. Esto permite obtener un feedback a nivel global del procedimiento, al nivel de etapas, y finalmente, al nivel de actividades.

La investigación fue realizada utilizando los datos de un entrenamiento del procedimiento quirúrgico Traqueostomía Percutánea por Dilatación (PDT), para el cual se entregaron diversas visualizaciones que permiten entender el progreso de los alumnos a través de sus sesiones en base a un enfoque de proceso. Por último, se estudió el comportamiento de las métricas orientadas a procesos en compración a a los métodos clásicos de evaluación. A futuro, se espera poder validar estas métricas para su uso profesional cómo método de evaluación y así complementar los actuales.

Palabras Claves: Minería de Procesos, Entrenamiento Médico, Feedback, Habilidades Técnicas, Procedimiento Quirúrgico.

1. INTRODUCTION

Surgical procedures training is an essential component within the scope of medical education, in which it is defined whether or not the residents reach the appropriate level to perform surgery. Within the training methods, simulation stands out as one of the most effective (Seam et al., 2019) to teach and evaluate procedural skills (Scalese et al., 2008), due to its proximity to a real patient scenario. However, today it is still an expensive method, so the quality of the residents' learning obtained in the training is proving to be absolutely crucial. Besides, the opportunity cost of having a doctor teaching surgical procedures to students instead of treating patients is very high (Aitken, 2012). Therefore, it is critical to generate maximum learning for residents from the information obtained in each training session.

In simulation, several evaluation methods are used to measure the level of competence acquired by residents during training. Commonly, two types of approaches are used in evaluation: Global Rating Scales (GRS) (Anderson et al., 2006) and Checklists (Epstein, 2007). Both have proven to be effective in establishing the level of proficiency in the execution of procedures (Morgan et al., 2001). However, they are designed exclusively for that purpose (Williams et al., 2002), without capturing information explicitly about the variability with which the resident performs the sequence of steps of each execution and the evolution of this sequence of steps through the training sessions.

This research contributes to the analysis and evaluation of the residents' progress in their simulation training processes, increasing the information obtained about the training and providing new instances of feedback with a process-oriented analysis. By understanding where residents make mistakes, we can intervene in the methodologies and training processes to improve the quality of learning, which is the main motivation of the research.

1.1. Context

The application of process modeling tools is commonly used in engineering to analyze processes and find points of improvement in them so as to deliver better outcomes. In turn, surgical procedures, given their nature, consist of a progression of steps, which allows them to be understood as a process (Neumuth, 2017). This interdisciplinary way of studying surgical procedures has raised great interest in both areas (medical and engineering), generating promising projects that mix tools from both fields and seek to make their learning processes more efficient.

This thesis is part of a long project of the HAPLAB laboratory of the Pontificia Universidad Católica de Chile. This seeks to standardize surgical procedures through the elaboration of well-defined models. Along with this, in the practical part, the residents who have participated in the project have been taught different procedures using this perspective to later be evaluated in training sessions. The goal is to obtain more objective information from the training and, in turn, more objective evaluation methods.

Within the context of surgical procedures, there are numerous articles (Nakai, 2018; Kristensen et al., 2015) that explain different techniques to perform different types of procedures, however, they are oriented toward describing the technique of the procedure rather than discussing the most effective way to teach it or provide feedback to residents. Therefore, this research comes to propose metrics to reinforce the analysis of the training of the procedure and thus, complement the traditional evaluation methods and the feedback generated in all the learning process.

To visualize the impact of this approach, this research was carried out on the data generated by the training process of residents of the Pontificia Universidad Católica de Chile in the Bronchoscopy Guided Percutaneous Dilatational Tracheostomy (BG-PDT) procedure, which was taught based on a predefined process model made by experts in this procedure (de la Fuente et al., 2021).

1.2. Research thesis

Given, that there is information on the variability with which the students are carrying out the sequence of steps of the procedure that is not being captured, the use of processoriented metrics that can measure the change of the student's learning throughout the sessions is suggested. The objective of these metrics is to improve the analysis of surgical procedure training to optimize resident learning and to make more effective the way to teach surgery.

1.2.1. Research hypothesis

Thus, the following hypothesis is proposed: Through process-oriented metrics, an analysis of the residents' progress can be obtained which indicates which sections of the procedure are more difficult to learn without deviating from the results of classic evaluation methods

1.2.2. Research objectives

The goal of this thesis can be broken down into two specific objectives:

- (i) Allow a standardized and specific analysis, from a global level of the procedure to one activity by activity, of the progress of the residents in process-oriented metrics, complementing the feedback provided by the instructors.
- (ii) Compare the behavior of process-oriented metrics against the results delivered by the classic evaluation metrics.

1.3. Background

To get a better understanding of the research, the most important concepts discussed within it are briefly detailed below.

1.3.1. Classic metrics in training

Classic metrics represent the most common evaluation methods in the surgical procedure training field. They define which resident is qualified to carry out the procedure on their own. Among these metrics, the GRS (Global rating scale) and the Checklist stand out as the most used in surgical procedures training (Morgan et al., 2001).

The GRS follows a predefined scale for the procedures which responds to certain criteria where the evaluator sets a score for each over an expected performance, thus, it is a subjective method (Anderson et al., 2006). These criteria are related to the quality, safety and comfort with which the resident executes the procedure. On the other hand, the Checklist is an objective evaluation method, in which each step of the procedure is detailed (Epstein, 2007). In training, instructors check if the resident performs these steps in every execution.

1.3.2. Feedback in medical education

Feedback within the context of medical education is defined as the delivery of specific information on the performance of a trainee compared to an expert level, given with the intent to improve the trainee's performance (van de Ridder et al., 2008). Instructors are usually tasked with delivering it together with the reports of the classic evaluation methods. Also, feedback can be delivered during the training as well as when it has finished.

Quality feedback is essential to promote learning in residents during their training. For this to be optimal, it must be delivered in a specific and descriptive way. If an instructor points to specific errors in performance, students are less likely to perceive the information as personal criticism and they will be able to integrate it more easily (Ghaderi and Farrell, 2020), understanding the exact points where the procedure is not being well performed. On the other hand, it should be delivered in a simple and not bulky way, since this could have negative effects on the residents' working memory (McGraw et al., 2018), worsening the outcomes of the learning process.

1.3.3. Business Process Modeling

A business process is the combination of a set of activities within an enterprise with a structure describing their logical order and dependency, whose objective is to produce a desired result (Aguilar-Savén, 2003). A Business Process Model (BPM) is the graphical representation of these business processes or workflows. All the steps of the process are presented in such a way that an end to end visualization of it is obtained.

A BPM delivers a standardized version of the process, of high or low abstraction, so that the actors of an organization can understand the activities and workflows that compose it in a more efficient and didactic way. In addition, this allows for greater transparency within those involved in the process, indicating the requirements of each activity and those responsible. Among the most widely used types of BPM we can find the Business Process Modeling and Notation (BPMN). This is characterized by having an intuitive notation (see an example in Figure 1.1) and, at the same time, being able to represent extremely complex processes.



Figure 1.1. Example BPMN explaining an order submission of a customer in a store. Source: https://www.flokzu.com/blog/en/bpm/what-is-bpmn/

1.3.4. Process Mining

Process Mining (van der Aalst, 2016) is an emerging discipline that allows analyzing the execution of a process based on the knowledge extracted from event logs created from the data stored in information systems. Event logs record the executions made by one or more actors of the process, which has been previously modeled with the activities and flows that compose it. Process mining provides an important bridge between data mining and business process modeling and analysis (van der Aalst et al., 2012). It allows to get fact-based insights derived from the data that is already in the organization processes. Then you can audit, analyze, and improve your processes.

For the logs of an event log to be valid (van der Aalst, 2016), they must be composed of at least three elements: an identifier (Case Id), which represents the number of the execution; the activity executed (Activity); and finally, the start an end time of the activity execution (Timestamps). In this way, the event log can be ordered sequentially. To obtain more detailed information, it is common to add other descriptive elements to the event log, such as the actor that performs the activity, devices used in the execution, among others.

There are three different ways to use an event log, these are called the main process mining techniques (van der Aalst et al., 2012). These are the following (see Figure 1.2):

- (i) **Discovery**. A discovery technique takes an event log and produces the model without using any apriori information.
- (ii) Conformance. Here, an existing process model is compared with an event log of the same process. Conformance checking can be used to check if reality, as recorded in the log, conforms to the model and vice versa.
- (iii) Enhancement. The idea is to extend or improve an existing process model using information about the actual process recorded in some event log.



Figure 1.2. The three ways that event logs can be used with their corresponding output: (a) discovery, (b) conformance checking, and (c) enhancement. Source: (van der Aalst, 2012).

1.4. Materials and methods

Understanding the background in which the research was developed, we proceed to detail the materials and methods that were used to carry it out. The research is based on the POME methodology (Munoz-Gama et al., 2021) which basically consists of three main stages: Model definition, Data collection and Procedure analysis.

1.4.1. Model definition

Before starting the training sessions, the POME methodology indicates that a model of the procedure is required, from the beginning to the end, to extract the activities and then generate an event log which refers to these activities. For this research, a BPMN model elaborated by a group of experts was used as a reference of the PDT procedure. This was extracted from an article (de la Fuente et al., 2021).

1.4.2. Data collection

After defining the model, we proceed to carry out the training sessions and collect the generated data. In this research, data from a BG-PDT training from the Faculty of Medicine of the Pontificia Universidad Católica de Chile was used. 8 senior residents from Anesthesiology, Emergency Medicine, Internal Medicine or first year Intensive Care participated. Each one carried out 7 sessions, 6 in the simulator and 1 session with cadaver (session 7), adding a total of 56 training sessions.

It is important to note that prior to the start of the training, the BPMN model of the procedure was explained to the residents so that there were no doubts regarding the activities that they should perform in the simulator and cadaver stages.

1.4.2.1. Classic metrics

For the training sessions, the executions of the residents were evaluated with 4 different metrics to define the degree of proficiency achieved in each session, so the instructors can establish whether or not the student is able to carry out the procedure. These were: OSATS, Execution time, TPL distances and the number of movements executed. For this research, OSATS and Execution time were taken as reference.

1.4.2.2. Video Tagging

To identify the activities completed by the residents in each training session, they were video recorded from the beginning of the execution of each resident until it was finished, regardless of whether they reached the last activity or not of the procedure. Then, each video recording was tagged by experts indicating what activity was carried out throughout the session using the POME software tool (Leiva et al., 2019). This is a software that helps to synchronize the time execution with execution events making tagging easier for residents.

1.4.2.3. Log generation

With all the tags of the videos inside the POME tool, it automatically generates the event log. For each activity executed by the resident, an event log was generated. These

event logs stored the Case Id, the activity name, the start and end times of each activity (Start and Complete Timestamps), the resident who performed the procedure, and finally, the respective training session.

1.4.3. Procedure analysis

This research proposes the use of process-oriented metrics that allow a top-down analysis of the collected event log. Three process-oriented metrics were defined, which measure the degree of non-compliance with the execution with the procedure model, showing the differences with the expected execution.

Based on the defined model and the data collected, a top-down analysis was developed with the proposed process-oriented metrics. It begins from a broad perspective, analyzing the process at its highest level of abstraction and then progressing in level of detail, analyzing at the level of stages, and then, at the level of activities. This analysis is accompanied by visualizations that show, from the perspective of process-oriented metrics, the progress of the residents throughout the sessions. Subsequently, the correlation between the classic metrics of the procedure and the process-oriented metrics was studied.

1.4.4. Research's paper

The results and conclusions of this research are reflected in the paper "Process-oriented metrics to provide feedback and assess the performance of students who are learning surgical procedures: The Percutaneous Dilatational Tracheostomy case" that will be presented in the next chapter. The progress of the residents according to the proposed process-oriented metrics is presented in detail, detecting the stages of greatest difficulty and the activities that could be generating bottlenecks during training. Also, the evolution of these metrics was compared with the classic ones as the training progresses. This paper was submitted to the Medical Teacher journal.

2. PAPER: PROCESS-ORIENTED METRICS TO PROVIDE FEEDBACK AND ASSESS THE PERFORMANCE OF STUDENTS WHO ARE LEARNING SURGICAL PROCEDURES: THE PERCUTANEOUS DILATATIONAL TRA-CHEOSTOMY CASE

2.1. ABSTRACT

Purpose

Assessing competency in surgical procedures is key for instructors to distinguish whether a resident is qualified to perform them in patients. Currently, assessment techniques not always focus on providing feedback about the order in which the activities need to be performed. In this research, using a Process Mining approach, process-oriented metrics are proposed to assess the training of residents in a Percutaneous Dilatational Tracheostomy (PDT) simulator, identifying the critical points in the execution of the surgical process.

Materials and methods

A reference process model of the procedure was defined, and video recordings of student training sessions in the PDT simulator were collected and tagged to generate event logs. Three process-oriented metrics were proposed to assess the performance of the residents in training.

Results

Although the students were proficient in classic metrics, they did not reach the optimum in process-oriented metrics. Only in 25% of the stages the optimum was achieved in the last session. In these stages, the four more challenging activities were also identified, which account for 32% of the process-oriented metrics errors.

Conclusions

Process-oriented metrics offer a new perspective on surgical procedures performance, providing a more granular perspective, which enables a more specific and actionable feedback for both students and instructors.

2.2. INTRODUCTION

Medical education has shown a significant increase in the use of simulation to teach and evaluate procedural skills (Scalese et al., 2008). This training method allows instructors to conduct more learner-centred training (Lammers et al., 2008) and has also been shown to be an effective method for residents to reach an adequate level of proficiency prior to patient contact (Seam et al., 2019). However, its use has been limited due to the high costs of using simulation models for specific skills (Lichtenberger et al., 2018). There are studies that have sought to reduce their costs for some procedures with new technologies (Kattan et al., 2019; Lichtenberger et al., 2018). Even so, many procedures still remain very expensive to teach and evaluate using simulation. Besides, the opportunity cost of having a doctor teaching surgical procedures to students instead of treating patients is very high (Aitken, 2012). Therefore, it is critical to generate maximum learning for residents from the information obtained in each training session.

Two complementary goals of the instructional process of surgical procedures are to asses performance and provide feedback. The purpose of assessing competency in the performance of a surgical procedure is to define whether a person is capable of performing it under certain conditions. In turn, the objective of feedback is to provide specific information comparing the student's performance and a standard (Ghaderi and Farrell, 2020).

In simulation, several evaluation methods are used to measure the level of competence acquired by residents during training. Commonly, two types of approaches are used in evaluation: Global Rating Scales (GRS) (Anderson et al., 2006) and Checklists (Epstein, 2007). GRS are subjective, but have the flexibility to be adapted to any surgical procedure. Checklists, on the other hand, seek to demonstrate whether or not each of the steps of the procedure is performed (Lammers et al., 2008). Both have proven to be effective in establishing the level of proficiency in the execution of procedures (Morgan et al., 2001). However, they are designed exclusively for that purpose (Williams et al., 2002), do not

capturing explicitly information about the variability with which the resident performs the sequence of steps of each execution and the evolution of this sequence of steps through the training sessions.

Process Mining is an emerging discipline that generates knowledge from process execution data recorded in information systems (van der Aalst, 2016), facilitating the analysis of the observed processes. This discipline has been used in several healthcare specialties (Rojas et al., 2016). Since surgical procedures can be understood as a progression of steps, they can be viewed as a process (Neumuth, 2017), so the inclusion of process mining for their analysis has emerged as an opportunity to deliver new information about learning in these procedures.

Recently, this approach has been applied (de la Fuente et al., 2020) to analyse the variability with which experts and residents perform the installation of Ultrasound-guided Internal Jugular Central Venous Catheter (UGIJCVC), identifying patterns that show the difficulty that residents have in learning this procedure. On the other hand, for the same procedure, the use of process mining techniques was proposed to identify desired and undesired process patterns, in order to complement personalized feedback to students using a process perspective (Lira et al., 2019).

In this study, we analysed a Percutaneous Dilatational Tracheostomy (PDT) training case extracted from a simulation study (Kattan et al., 2020) that, like any common surgical training, is evaluated with metrics that measure the proficiency of the skills of performing the procedure as a whole. Within these, one can find the OSATS (Objective Structured Assessment of Technical Skills, a type of GRS) performance metrics and the execution time, which evaluate the process as a whole. However, the feedback obtained from these metrics may not be sufficient to know where residents may be having problems with the order of steps, so the need arises to look for metrics that complement them and make better use of the available information.

Our hypothesis is that, by using metrics obtained using process mining, we can assess in a more detailed way the proficiency with which residents perform their training sessions and, at the same time, provide more accurate information about where they are making mistakes. Through a top-down analysis focused on stages and activities of the surgical procedure, critical stages and activities in the execution of the procedure are detected, and the variability with which residents perform them, which is not fully captured by classical evaluation metrics, becomes evident.

2.3. MATERIAL AND METHODS

In this study, the use of process-oriented metrics is proposed to analyse the performance of residents during their training in a PDT procedure simulator. The study is based on the POME methodology (Munoz-Gama et al., 2021), which is basically composed of three stages: Model Definition, Data collection and Procedure Analysis.

2.3.1. Model Definition

A reference process model was established to extract and analyse the activities of the procedure. For its definition, a generic model of the Bronchoscopy-Guided PDT (BG-PDT) procedure was used, which was developed based on the consensus among experts using the Delphi method (de la Fuente et al., 2021). This BG-PDT model was adapted to represent the procedure incorporating the limitations of the simulator. Finally, a more reduced model was obtained (Figure 2.1), which considers a total of 23 activities (21 mandatory and 2 optional), which are grouped into 4 stages: procedural preparation, tracheal puncture, tracheal dilatation and cannula placement.

2.3.2. Data collection

Videos of residents performing PDT training in a low-cost BG-PDT simulation and in a cadaveric model (Kattan et al., 2019) were collected. 8 senior (postgraduate year



Figure 2.1. Reference model used in the study. Adapted from the original BPMN model (de la Fuente et al., 2021)

3) residents from anaesthesiology, emergency medicine, internal medicine or first year intensive care from the Faculty of Medicine of the Pontificia Universidad Católica de Chile participated in the training. Before the evaluation, residents were shown relevant BG-PDT literature and a step-by-step video of the complete procedure performed in the simulation. A total of 56 sessions were performed, 7 sessions per resident, 6 in the simulation and 1 cadaver session (session 7).

The videos recorded the execution of each of the steps of the procedure performed by the residents, from the beginning of the procedure, including the preparation of instruments to the connection of the cannula. To generate the event logs, the videos were labelled using the software POME (Leiva et al., 2019). For each activity executed by the resident, a unique identifier of the execution (Case Id) is stored, in this case an identifier that combines the corresponding resident and a correlative of each training session of the procedure; the executed step (activity); and the start and end of each activity (start and complete timestamps).

2.3.3. Procedure analysis

2.3.3.1. Process-oriented metrics definition

This paper proposes the use of three process-oriented metrics that allow a top-down analysis. These metrics measure the degree of non-compliance of the execution with the procedural model, showing the differences with the expected execution. Given the sequence of activities $S = \langle a_1, a_2, ..., a_{n-1}, a_n \rangle$ described by the defined model, and a_k the activity expected at position k in the sequence, Table 2.1 presents the process-oriented metrics to be used in the study. It should be noted that it is assumed that there are no parallel paths in the execution of the procedure; the metrics could be generalized to consider the existence of parallelism.

Table 2.1. Process-oriented metrics that quantify errors made by residents during training, compared to the model defined for this procedure (PDT)

Name	Objective	Formal definition	Example
Omission	Torecordwhetheranobligatoryac-tivityisnotperformedduringthe execution.	Count how many times a mandatory activity a_k , $1 \le k \le n$ is not performed in the execution.	→ Withdraw syringe Advance j- tip guidewire Withdraw needle →
Repetition	To record whether an ac- tivity is repeated consecutively during execution.	Count how many times an activity a_k , $1 \le k \le n$ is followed by the same activity a_k in the execution	Withdraw syringe
Deviation	Torecordwhetheranactivityisper-formedinincorrect	Count how many times an activity a_k , $2 \le k \le n$ is followed by an activity a_j , $1 \le j \le k$ in the execution sequence.	Withdraw syringe
Aggregated	Torecordtheresultofallprocess-oriented an metricsfor an activity. an	The sum of all process-oriented metrics involving an activity a_k .	

It should be noted that in order to record whether an activity was involved in the deviation metric, it was considered whether that activity was the origin of the incorrect pattern, the destination of the incorrect pattern, or if it was in between them in the expected model and was performed in the respective execution.

2.3.3.2. Process mining top-down analysis

Based on the defined model (Figure 2.1) and the data collected, a top-down analysis was developed, starting from a broad perspective, analysing the process as a whole, and then moving to a more detailed level, analysing at the stage level, and then at the activity level. Subsequently, the correlation between the classic metrics of the procedure and the process-oriented metrics was studied.

The top-down methodology used begins with an analysis of the outcome of each resident's process-oriented metric aggregated into a total outcome per session. The Wilcoxon test was used to measure whether the progress between sessions was statistically significant at the group level ($p \le 0.05$). Then, the above result is broken down for each of the stages of the procedure. From the previous result, the two stages with the worst results in the proposed metrics were selected, and their activities were analysed according to the process-oriented metrics. Finally, the evolution of two particular activities during training was analysed.

2.3.3.3. Classic metrics and process-oriented metrics

The linear correlation between process-oriented metrics and classic metrics was measured to study their behaviour as the sessions progressed. To represent the performance of the residents in the classic metrics, the results of the execution time and OSATS metrics from the simulation study (Kattan et al., 2020) were used. These two classic metrics were analysed in this research since they are not restricted to any specific procedure (Niitsu et al., 2013). Execution time was measured from the moment the resident started the first activity in the model until the last activity was completed. To measure OSATS, the videos were reviewed by two blinded experts. The time metric is considered to improve as it decreases, while OSATS ranges from 1 to 25, with 25 being the maximum expected score. To obtain the correlation values, Pearson's correlation index was used, comparing each processoriented metric with each classic metric for each procedure execution separately. The result was considered significant with a p-value of less than 0.05.

2.4. RESULTS

2.4.1. Process-oriented metrics at high level

The results show a tendency for the total errors of the residents to decrease as they progress through their training sessions (Figure 2.2), with the exception of the cadaver session (session 7). In this last session, in particular, the deviations increase with respect to the sixth session with simulator. At the end of the training, the residents performed the procedure without repetitions, and omissions decreased. On the other hand, deviations did not improve significantly over the sessions. For the omission metric, there was a statistically significant decrease in sessions 5, 6 and 7 with respect to session 1 (Wilcoxon, $p \leq 0.05$). On the other hand, no evidence was found to establish a statistically significant difference in the other metrics.

2.4.2. Process-oriented metrics by stage

Continuing with the top-down analysis, Figure 2.3 shows how the results of the processoriented metrics are distributed in each of the 4 stages of the procedure.

Procedural preparation stage

For the Procedural preparation stage, a maximum of 7 total errors were observed, which occurred in the first session. Throughout the 7 sessions, there were 0 repetitions.



Process-oriented metrics results grouped by sessions

Figure 2.2. Results obtained in each process-oriented metric, grouped by training session. (*) There is a statistically significant decrease in the omission metric with respect to session 1 (Wilcoxon, $p \le 0.05$).



Figure 2.3. Results obtained in each process-oriented metric, aggregated by training session, broken down by stage.

After the third session, there were no more omissions, and after the fourth session, there were no more deviations.

Tracheal puncture stage

In the Tracheal puncture stage, the maximum number of errors occurred in the fourth session, with a total of 25 errors. Throughout the 7 sessions, there was only 1 repetition in session 3. Omissions decreased in the last three sessions, but not substantially. Deviations, on the other hand, remained more or less constant throughout the sessions. In all sessions, omissions predominated over the other metrics.

Tracheal dilatation stage

For the Tracheal dilatation stage, the maximum number of errors occurred in the first and second session, reaching a total of 30; then they decreased throughout the sessions, having a slight upturn in the session with cadaver. It is observed that the repetitions decreased as the sessions progressed. However, there was a repetition in session 6, after having observed that they had ceased to occur in session 5. Omissions did not tend to decrease. Deviations and omissions were detected in all sessions. Also, in all sessions the value of the deviation metric predominated over the others.

Cannula placement stage

For the cannula placement stage, the maximum number of errors occurred in session 1, where all the errors corresponded to omissions of procedural activities. In session 3 the omissions ceased and the deviations started to happen, which were maintained until the last session.

2.4.3. Process-oriented metrics in activities

We proceeded to analyse the activities of the two stages that did not show a clear tendency to improve with the development of the sessions (Tracheal puncture and Tracheal dilatation). Figure 2.4 shows the sum of the results obtained in the process-oriented



Figure 2.4. Sum of the results obtained in the process-oriented metrics considering all training sessions and all residents (56 executions in total) for the selected stages, broken down by activities. *The activities are abbreviated and are displayed according to their order in the model.

metrics considering all the sessions and all the residents (56 executions in total) for the selected stages, broken down by activities. This shows which activities can be the most difficult for the residents to learn.

All activities in these model stages have scores greater than 0 on all three processoriented metrics, with the exception of *Perform 1.5 cm wide horizontal incision*, which is not a mandatory activity and therefore has no omissions. Of the 13 activities with errors, 10 are dominated by deviations. In the remaining 3, omissions predominate.

Tracheal puncture

The activities with the highest number of errors in the Tracheal puncture stage are: *Palpate 2nd tracheal ring* (56 aggregated errors; 48 omissions), *Stabilize the larynx with the middle finger and thumb* (53 aggregated errors; 39 deviations) and *Puncture between 1st and 2nd tracheal ring* (50 aggregated errors; 50 deviations).

Tracheal dilatation

The activities with the highest number of aggregated errors in the Tracheal dilatation step are: *Hold trachea with the non-dominant hand* (66 aggregated errors; 49 deviations), *Advance guiding catheter* (56 aggregated errors; 51 deviations) and *Advance dilator in 45°* (56 aggregated errors; 50 omissions).

2.4.4. Progress of process-oriented metrics in activities

The evolution of two specific activities was analysed, *Advance dilator in 45° until positioning marks meet*, which belongs to a stage that did not show a relevant improvement during training (Tracheal dilatation) and *Withdraw dilator, leaving guidewire and guiding catheter*, which belongs to a stage that did improve (Cannula placement). Figure 2.5 shows that the first one does not present a clear positive evolution, since it starts with 11 aggregated errors (6 deviations; 1 omission; 4 repetitions) and, as the sessions progress, the deviations tend to be maintained. However, from session 5 onwards, there are no more omissions or repetitions. In contrast, the *Withdraw dilator, leaving guidewire and guiding catheter* activity shows that the residents did decrease their errors during training. They start with 4 aggregated errors (2 deviations; 2 omissions) in session 1 and then, from session 5 onwards, they do not present any more errors.



Figure 2.5. Evolution of all process-oriented metrics in the activities Advance dilator in 45° and Withdraw dilator, leaving guidewire and guiding catheter, by training sessions.

2.4.5. Comparison between process-oriented metrics and classic metrics

All process-oriented metrics have a negative correlation with the OSATS metric, showing that the reduction of errors in the execution of procedural activities correlates with an increase in the achievement of competence (Table 2.2). In turn, all process-oriented metrics have a positive correlation with the execution time metric. For all comparison cases, the correlation is statistically significant with a p-value less than the established p-value (p = 0.05). Figure 2.6 shows that, like the classic metrics, the process-oriented metrics showed improvement as the sessions progressed.

Table 2.2. Correlation results between classic metrics and process-oriented metrics considering all executions (56). Significance value (p-value) in parentheses.

Classic metric	Repetition	Omission	Deviation
OSATS	-0.53 (<i>p</i> < 0.01)	-0.63 (<i>p</i> < 0.01)	-0.45 (<i>p</i> < 0.01)
Execution Time	$0.57 \ (p < 0.01)$	$0.71 \; (p < 0.01)$	$0.54 \ (p < 0.01)$



Figure 2.6. Evolution of the aggregated process-oriented metric and the classic metrics, by training sessions.

2.5. DISCUSSION

This study proposes the use of process-oriented metrics that allow the execution of the PDT surgical procedure by residents in training sessions to be analysed using a process perspective. The top-down analysis identified which stages and which activities of the procedure are the most difficult for residents to learn regarding the order of the steps. Overall, process-oriented metrics show similar behaviour to classic metrics while providing a higher level of detail to understand how residents evolve during the training process.

Finding the critical points of the procedure is key to designing an effective training plan for residents. The cognitive load theory (Sweller, 1988) points out that trainees' memory is limited when receiving new information and, when overwhelmed, their learning capacity is diminished. This study, through top-down analysis, shows that it is possible to identify the stages and activities in which the residents might be overwhelmed and, consequently, at the end of the course they do not achieve a proficient result according to the processoriented metrics. At stage level, the stages Tracheal puncture and Tracheal dilatation were identified as those where most procedural errors occur. Then, at the activity level, those activities with the highest number of errors in the stages analysed in detail are: *Palpate 2nd tracheal ring* (Tracheal puncture), *Advance guiding catheter* (Tracheal dilatation), *Advance dilator in 45°* (Tracheal dilatation) and *Hold trachea with the non-dominant hand* (Tracheal dilatation). Finally, the process-oriented metrics allow the evolution of the activities through the sessions to be reflected, showing which activities do not show progress as the training progresses.

For the training process to be effective, all its parts must be well balanced cognitively (McGraw et al., 2018), so as not to overload the residents' working memory. The results show that not all activities and stages have the same cognitive load, which could be generating cognitive bottlenecks (Young et al., 2014), unbalancing the learning process. A new decomposition of activities and stages could have positive effects on student outcome (Nicholls et al., 2020). By limiting the amount of information delivered per stage of

the procedure, a more balanced model of the procedure could be conveyed (Young et al., 2014), with the cognitive load better distributed, thus improving the learning experience of the residents.

Feedback in the context of medical education is defined as the delivery of specific information on the performance of a trainee compared to an expert level, given with the intent to improve the trainee's performance (van de Ridder et al., 2008). For feedback to be optimal, it should be as specific as possible, since, when an instructor points out specific errors in performance, trainees are less likely to perceive the information as personal criticism (Ghaderi and Farrell, 2020), allowing the trainee to focus on specific points to improve in the performance of the procedure. The results of process-oriented metrics deliver standardized information that, from the highest to the lowest level, would allow residents to understand in which stages/activities of the specific procedure they are failing and how they are progressing in those stages/activities during training. In addition, the feedback provided by process-oriented metrics would be less emotionally charged and stressful than a global evaluation itself, given the impartiality with which it is presented. Minimizing the burden of affective factors that are directly related to cognitive load could optimize procedural learning (Szulewski et al., 2020). In addition to the above, processoriented feedback has been well received by residents (de la Fuente et al., 2020), who state that the standardization of the process together with concrete feedback improves their training experience.

To the best of our knowledge, while there are studies on tracheostomy training (Nakai, 2018; Kristensen et al., 2015), they are oriented toward describing the technique of the procedure rather than discussing the most effective way to teach it or provide feedback to residents. This study proposes the use of process-oriented metrics to reinforce the analysis of this procedure and thus complement traditional evaluation methods, increasing the information generated for both instructors and residents. These metrics do not contradict the classic metrics, since, at a global level, these metrics presented an evolution similar to the classic metrics used in the simulation study (Kattan et al., 2020). Throughout the training

sessions, the process-oriented metrics improved as well as the classic metrics. Even for session 7, in the switch to cadaver, both get worse. Coherently, a statistically significant correlation was observed between each of the process-oriented metrics with the classic metrics. However, process-oriented metrics are capable of identifying that there are activities and stages of the process in which there is still room for improvement, according to the process perspective, which is not detected by classic metrics. For session 6, the classic metrics are already proficient in OSATS and execution time, while among the process-oriented metrics, deviations and omissions still have non-optimal values. This makes us observe, as is pointed out in (de la Fuente et al., 2020), that the process perspective helps to find hidden information that is not reviewed by traditional evaluation methods or by the observation of the instructors themselves.

There are certain limitations to this study. In the first place, the model used does not represent the complications that a real execution could have. In addition to the above, no analysis has been carried out on the execution of experts, so the conception of proficiency could vary if, when observing the execution of an expert under real situations, it is concluded that it is good in classical metrics, but not in process-oriented metrics. On the other hand, in the definition of the model used in the study, several of the activities of the original model were left out, excluding the bronchoscopist from the procedure. This was due to the fact that this study focuses on the learning of procedural skills, while the bronchoscopist's activities are monitoring activities, therefore, considered secondary for this study.

In conclusion, the process-oriented metrics capture new information to analyse the performance of the residents in the execution of the PDT procedure, which allows us to provide them with adequate and more detailed feedback as they progress through the training sessions, both at the global level and at the stage and activity levels. These process-oriented metrics showed a statistically significant correlation with the classic metrics, so it is promising to propose them as a tool to complement the analysis of the results provided by the evaluation methods currently used. Additionally, through these process-oriented

metrics, instructors can be made aware of the weak points of procedural learning. With this information, they can evaluate whether instructional resources should be redeployed to make the teaching process more efficient.

3. THESIS CONCLUSIONS AND FUTURE WORK

In this last chapter, the main conclusions derived from the research are detailed according to the the objectives established. Finally, some possible future works are discussed to strengthen the validity of the metrics presented.

3.1. Conclusions

In this research, the use of process-oriented metrics was studied to improve and complement the analysis of surgical procedures training. A top-down analysis was generated that allowed us to know how the students were carrying out the process according to these process-oriented metrics, first at a global level of the whole procedure, then at the level of stages and, finally, at the level of activities. Finally, the behavior of these process-oriented metrics was compared with the classic metrics that are used to evaluate and give feedback to the students.

The top-down analysis identified which stages and which activities in the procedure are the most difficult for residents to learn from a process perspective. This provides clear insights to a better understanding of how the training process is being carried out and thus, find the crucial points where there could be space for improvement. Along with this, it was evidenced that these metrics allow to graphically visualize the progress of each student at a global level as well as at the level of stages and activities.

The process-oriented metrics presented showed a similar behavior to the classic metrics at the global level, obtaining a statistically significant correlation, but also they provide a greater level of detail to understand how residents evolve during the training process.

Finally, from the conclusions described together with the results presented in the paper included in Chapter 2, it is possible to affirm that the hypothesis raised "Through processoriented metrics, an analysis of the residents' progress can be obtained which indicates which sections of the procedure are most difficult to learn without deviating from the results of classical methods" can be confirmed.

3.2. Limitations

There are certain limitations to this study. In the first place, the model used does not represent the complications that a real execution could have. In addition to the above, no analysis has been carried out on the execution of experts, so the conception of proficiency could vary if, when observing the execution of an expert under real situations, it is concluded that it is good in classical metrics, but not in process-oriented metrics.

Secondly, the feedback obtained in the research has not been applied directly to the students who participated in the study, so the impact of this is not known. However, this showed great potential and it could be interesting to carry out a future research that can capture its value.

Finally, several of the activities of the original model were left out of the model used in the study, excluding the bronchoscopist from the procedure. This was due to the fact that this study focuses on the learning of procedural skills, while the bronchoscopist's activities are monitoring activities, therefore, considered secondary for this study.

3.3. Future work

The metrics presented yielded promising results in relation to the detection of errors by residents regarding the variability with which they perform the procedure in their sessions. Although, these have not gone through a validation process suitable for a professional use, so actually, they can only be used as a complementary method to the classic metrics. Therefore, it is proposed as the next step, to begin a research that allows validating them for use in training as an evaluation method and thus, adding them to those already used to provide a broader and more objective view of traditional methods.

Another future research is to carry out the same analysis described in this thesis with an event log created based on the performance of the surgical procedure by medical experts. This would help to standardize the results of the process-oriented metrics that ensure minimal abilities to perform the surgical procedure. In this way, it might be possible to detect an adequate proficiency level in a resident when he has reached a value relative to an expert.

REFERENCES

Aguilar-Savén, R. (2003). Business process modeling: Review and framework. *International Journal of Production Economics*, 90(2):129–149.

Aitken, R. (2012). Lost opportunity cost of surgical training in the australian private sector. *ANZ Journal of Surgery*, 82(May):145–150.

Anderson, D., Long, S., Thomas, G., Putnam, M., Bechtold, J., and Karam, M. (2006). Objective structured assessments of technical skills (osats) does not assess the quality of the surgical result effectively,. *Clinical Orthopedics and Related Research*, 474(4):874– 881.

de la Fuente, R., Fuentes, R., Munoz-Gama, J., Riquelme, A., Altermatt, F., Pedemonte, J., Corvetto, M., and Sepúlveda, M. (2020). Control-flow analysis of procedural skills competencies in medical training through process mining. *Postgraduate Medical Journal*, 96(1135):250–256.

de la Fuente, R., Kattan, E., and Munoz - Gama, J. (2021). Development of a comprehensive percutaneous dilatational tracheostomy process model for procedural training: A delphi-based experts consensus. *Acta Anaesthesiologica Scandinavica*, 64:244 – 256.

Epstein, R. (2007). Assessment in medical education. *The New England journal of medicine*, 356:387 – 396.

Ghaderi, I. and Farrell, T. M. (2020). Toward effective feedback: From concept to practice. *Surgery*, 167(4):685–688.

Kattan, E., de la Fuente, R., Putz, F., Vera, M., Corvetto, M., Inzunza, O., Achurra, P., Inzunza, M., Munoz-Gama, J., Sepulveda, M., Galvez, V., Pavez, N., Retamal, J., and Bravo, S. (2020). Simulation-based mastery learning of bronchoscopy-guided percutaneous dilatational tracheostomy. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 14(6):415 – 419.

Kattan, E., Vera, M., Putz, F., Corvetto, M., de la Fuente, R., and Bravo, S. (2019). Design and evaluation of a low-cost bronchoscopy-guided percutaneous dilatational tracheostomy

simulator. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 14(6):415 – 419.

Kristensen, M., Teoh, W., and Bake, P. (2015). Percutaneous emergency airway access; prevention, preparation, technique and training. *British Journal of Anaesthesiology*, 114(13):357–61.

Lammers, R., Davenport, M., Korley, F., Griswold-Theodorson, S., Fitch, M., Narang, A., Evans, L., Gross, A., Rodriguez, E., Dodge, K., Hamann, C., and Robey III, W. (2008). Teaching and assessing procedural skills using simulation: Metrics and methodology. *Academic Emergency Medicine*, 15(11):1079–1087.

Leiva, L., Munoz-Gama, J., Salas-Morales, J., Galvez, V., Lee, W., de la Fuente, R., Fuentes, R., and Sepúlveda, M. (2019). Pomelog: Generating event logs from unplugged processes. *Business Process Management Journal*, 2420(6):415 – 419.

Lichtenberger, J. P., Peter, S. T., Gada, S., Wyn, M., B, V., and Liacouras, P. (2018). Using 3d printing (additive manufacturing) to produce low-cost simulation models for medical training. *Military Medicine*, 183(1):73 – 77.

Lira, R., Salas-Morales, J., Leiva, L., Fuente, R., Fuentes, R., Delfino, A., Nazal, C., Sepúlveda, M., Arias, M., and Herskovic, V. (2019). Process-oriented feedback through process mining for surgical procedures in medical training: The ultrasound-guided central venous catheter placement case. *International Journal of Environmental Research and Public Health*, 16(1877).

McGraw, R., Chaplin, T., Rocca, N., Rang, L., Jaeger, M., Holden, M., Keri, Z., and Fichtinger, G. (2018). Cognitive load theory as a framework for simulation-based, ultrasound-guided internal jugular catheterization training: Once is not enough, but we must measure it first. *Canadian Journal of Emergency Medicine*, 21(1):141–148.

Morgan, P. J., Cleave-Hogg, D., and Guest, C. (2001). Comparison of global ratings and checklist scores from an undergraduate assessment using an anesthesia simulator. *Academic Medicine*, 76(10):1053 – 1055.

Munoz-Gama, J., Galvez, V., Fuente, R., Sepúlveda, M., and Fuentes, R. (2021). *Interactive Process Mining in Healthcare. Health Informatics*, chapter Interactive Process Mining for Medical Training. In: Fernandez-Llatas C. (eds). Springer.

Nakai, M. (2018). *Percutaneous Tracheostomy: Pearls and Pitfalls, and How to Create a "Hand-On" Training Program Course*, chapter de Farias T. (eds) Tracheostomy. Springer. Neumuth, T. (2017). Surgical process modeling. *Innovative surgical sciences*, 2(3):123–137.

Nicholls, D., Sweet, L., and Muller, A. (2020). Teaching psychomotor skills in the twentyfirst century: revisiting and reviewing instructional approaches through the lens of contemporary literature. *Medical Teacher*, 28:1056–63.

Niitsu, H., Hirabayashi, N., Yoshimitsu, M., Mimura, T., Taomoto, J., Sugiyama, Y., Murakami, S., Saeki, S., Mukaida, H., and Takiyama, W. (2013). Using the objective structured assessment of technical skills (osats) global rating scale to evaluate the skills of surgical trainees in the operating room. *Surgery today*, 43(3):271 – 275.

Rojas, E., Munoz-Gama, J., and Sepúlveda, M. (2016). Process mining in healthcare: a literature review. *Journal of Biomedical Informatics*, 21(June):224–36.

Scalese, R. J., Obeso, V. T., and Issenberg, B. (2008). Simulation technology for skills training and competency assessment in medical education. *Journal of General Internal Medicine*, 23(supl 1):46 – 49.

Seam, N., Lee, A., Vennero, M., and Emlet, L. (2019). Simulation training in the icu. *Chest*, 156(6):1223–1233.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, 12:257–85.

Szulewski, A., Howes, D., Van Merrienboer, J. J. G., and Sweller, J. (2020). From theory to practice: The application of cognitive load theory to the practice of medicine. *Academic Medicine*, Publish Ahead of Print:371–384.

van de Ridder, J. M., Stokking, K. M., McGaghie, W. C., and ten Cate, O. T. (2008). What is feedback in clinical education?. *Medical education*, 42(2):189 – 197.

van der Aalst, W. (2016). Mining P. Data science in action. Springer.

van der Aalst, W., Adriansyah, A., Alves De Medeiros, A. K., Arcieri, F., Baier, T., Blickle, T., Bose, J. C., Brand, van den, P. C. W., Brandtjen, R., Buijs, J. C. A. M., Burattin, A.,

Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., Leoni, de, M., and Wynn, M. (2012). , *Business Process Management Workshops (BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I*), chapter Process mining manifesto. In F. Daniel, K. Barkaoui, S. Dustdar (Eds.). Springer.

Williams, B. C., Litzelman, D. K., Babbott, S. F., Lubitz, R. M., and Hofer, T. P. (2002). Validation of a global measure of faculty's clinical teaching performance. *Academic Medicine*, 77(2):177 – 180.

Young, J. Q., Van Merrienboer, J., Durning, S., and Ten Cate, O. (2014). Cognitive load theory: implications for medical education. *Medical teacher*, 36(5):371–384.