Models of Cognitive Work Analysis in Neurocritical Care and Perspectives on Expertise– An Ecological Interface Design Approach

by
Ece Üreten

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	GUY DUMONT Professor Electrical and Computer Engineering University of British Columbia
Supervisor	CATHERINE M. BURNS Professor Systems Design Engineering University of Waterloo
Internal Members	KATE MERCER Adjunct Professor Systems Design Engineering University of Waterloo
	NING JIANG Associate Professor Systems Design Engineering University of Waterloo
Internal-external Member	JAMES WALLACE Associate Professor School of Public Health Sciences University of Waterloo

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored; see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This work was part of a collaboration between the Advanced Interface Design Lab (AIDL) and Toronto Western Hospital. I was supervised and guided by Professor Catherine Burns throughout the whole research stages. Our subject matter expert, Dr. Victoria McCredie shared in-depth insights into neurocritical care, facilitated the organization of observations at Toronto Western Hospital, and helped during the recruitment process. The usability study findings were evaluated with Dr. Alberto Goffi at St. Michael's Hospital.

Students in AIDL supported this project in various stages; Kathleen Schaef (master's student) conducted the interview study with me, while Cathleen Leone Grace (Co-op student) drafted initial visualizations. Rayyan Quraishi (intern) supported the implementation of the interface prototype used for the usability study. Test runs on the interface prototype were run with multiple students in AIDL.

The currently available publications were parts of this work and have been cited in various chapters. Chapter 2 includes the following paper written with our clinical collaborator Prof. Victoria McCredie, and my supervisor Prof. Catherine M. Burns.

- Üreten, E., McCredie, V., & Burns, C. (2020, September). Cognitive Work Analysis Models of Neuro-Critical Care. In Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (Vol. 9, No. 1, pp. 191-193). Sage CA: Los Angeles, CA: SAGE Publications.

Parts of chapter 5 were included in the paper below which was co-authored by Kathleen Schaef, Prof. Victoria McCredie, Prof. Catherine Burns, and myself. Kathleen Schaef published her master's thesis on this topic and thus, most parts of chapter 5 have been published in her thesis.

- Üreten, E., Schaef, K., McCredie, V., & Burns, C. (2022, September). Assessing Intracranial Pressure Visualizations Displayed on ICU Bedside Physiologic Monitors. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 66, No. 1, pp. 2143-2147). Sage CA: Los Angeles, CA: SAGE Publications.

Abstract

Introduction

Neurocritical care is a complex and data-rich environment in which timely decisions are of utmost importance. Small changes in the neurophysiologic states may imply neurological deterioration that can be fatal if not treated timely and appropriately. A multidisciplinary team of clinicians takes care of the patients with varying expertise levels. Expert physicians in this domain often try to act proactively and prevent further injury (such as secondary brain injury often visible in traumatic brain injury patients) and manage concurrently, assessing patient data represented through various sources. Novice intensive care physicians act rather reactively to specific events and have a sequential approach to management. The scattered data sources across the unit and the lack of centralized views on the patient data make it even more challenging for physicians, in general, to understand the happenings and support the patient. Collecting and identifying all data through various locations and systems can cause a high cognitive workload and is not optimally laid out through meaningful representations on different device interfaces.

Further, such interfaces (e.g., the bedside physiologic monitor) lack the incorporation of neurocritical-relevant concept visualizations that would help novices better analyze and understand the context and significant relationships among variables. This is especially important because the complexity of neurocritical care is high, immediate, and proactive treatment is desired. It usually takes much time for novices to learn the specifics of assessing neurological deterioration in a critical context and being aware of the situation to consider the trajectory and required medical support for the patient. Inappropriate actions and long decision-making time can result in unwanted health deteriorations or even fatalities.

Bedside physiologic monitors play a significant role in assessing the vital signs of patients; however, in neurocritical care, there are additional variables, such as Intracranial Pressure (ICP), that are significant biomarkers for neurological conditions and need to be assessed in context with other variables. In this work, we have focused on this specific variable as it is a key variable showing the health state of the brain and is interconnected with various other physiologic variables, which requires a detailed examination to assess neurocritical care patients' conditions. For novice neurocritical care physicians, it is a new variable and challenging context to learn how to deal with this variable as it was often displayed numerically and as a waveform on monitors. Assessing this variable in context needs to be trained to develop an understanding of (dys)functional autoregulation in the brain. Still, interfaces don't have much flexibility in representing this variable other than in a numeric or waveform pattern. Monitors thus can benefit from new representations of ICP within its context.

This work proposes an ecological interface design (EID) approach to represent ICP together with other variables on vital signs monitor to be used in neurocritical care. This work not only shows a new design for such monitor interfaces but also integrates the additional perspective on supporting novice physicians to develop more in the direction of experts to support them in the best way possible. Differences in the context of neurocritical care among physicians are explored in this work, as they are the main decision-makers in the unit. Once the understanding of differences between novices and experts can be disseminated, the steps that are required to develop expertise can be further examined and supported through various ways. Overall, the research objective of this work is to explore how expertise development can be supported through the ecological interface design approach of bedside physiologic monitors in the context of neurocritical care.

Methods and Modeling

As a starting point, observations in a neurocritical care unit are summarized and modeled through the Cognitive Work Analysis (CWA) framework using the Work Domain Analysis (WDA) and Control Task

Analysis (ConTa). This framework is selected to identify various elements of the neurocritical care context as it provides in-depth insights into complex sociotechnical areas and their constraints. Further, it lays a basis for the design of an ecological interface that enables the end-user to become an adaptive problem solver in any situation. The CWA models thus provide the basis for identifying key relationships and aspects that would be important to display on such interfaces.

In the following step, interviews were conducted with critical care experts and novices to further dive into their perspectives on expertise development, challenges, and potential ways of addressing these and supporting novices. These insights were relevant for adding and discussing aspects represented on the CWA models and exploring potential expertise-relevant measures that could be used for assessing the development of expertise through interface design.

Identifying both CWA and interview findings, static visualizations were developed and discussed with both novice and expert critical care physicians. Their feedback was incorporated into the new designs of the visualizations and integrated into a prototype (i.e., the ecological interface). This interface was compared to a standard interface used in the neurocritical care unit within a usability study. Two neurocritical care scenarios were presented to the participants in the control and experimental groups while being asked to think out loud about assessing and treating such patient cases based on the interfaces and data shown.

Results

The CWA models showed different dimensions of the work domain and how the physical variables are interconnected to provide and support patient recovery. The decision ladders mapped on a common neurocritical care scenario showed differences in expert and novice physicians' approaches to assessing, evaluating, and acting on certain signs related to the health states of a patient. While novices cognitively go through each step and might miss out on case-relevant information or may come up with misinterpretations even, expert physicians take shortcuts as they have built certain mental models already based on their vast knowledge and experience with similar cases.

Through the interviews and discussions with physicians, we further noticed that novice critical care physicians usually have a numerical and threshold approach. At the same time, experts expect them to develop an approach that incorporates understanding the significance of the waveform, trend, and individualize patient care depending on various patient characteristics. To further develop their mental model, they must step back from the "one size fits all" mindset and explore the optimal thresholds for each patient individually. To progress, they further need to develop autonomy and communication skills with various stakeholders, deepen their knowledge and familiarity with tools, perform adequately, and reflect on their actions. The interviews also revealed various aspects of expertise often tracked by the expert physicians training the novices. It was noted that experts often have a great amount of experience, knowledge, skills, training, work where there is no evidence, have excellent performance and specific personal traits.

The inputs from both the CWA models and the discussions with the physicians provided initial design ideas that could be further developed to support aspects of the novices' expertise development process in the interface design context. Concepts that are relevant for neurocritical care monitoring and associated challenges have thus been discussed. Both expert and novice critical care physicians provided feedback on the initial visualizations, which helped to iterate the visualizations further and were then incorporated into a prototype interface (i.e., the ecological interface). A standard interface was also developed as a prototype to compare how participants in both groups could be compared.

The intention was to show one potential ecological interface for neurocritical care that would specifically support the novice physicians to develop more characteristics of experts identified in the proposed studies. Novices using the ecological interface (experimental group) showed fewer errors when presented with two different neurocritical care scenarios, compared to the control group using the standard interface. At the same time, the experimental group's reasoning was more focused and included more investigation of the cause for ICP elevations and patterns compared to the group using a standard interface. Reflections on the ecological interface further showed that participants started thinking critically about their strategies and helped them state the reasons and limitations of their actions demonstrated during the scenario presentations. The confidence in handing over the patient to the next team has been rated on a numeric scale showing that the group using the standard interface rated slightly higher on average than the experimental group. Also, the average subjective performance ratings were lower for the ecological interface than the standard interface. It is noteworthy though, that the difference between both groups on confidence and performance ratings wasn't too far away from another. However, the usability of the ecological interface was perceived as useful, easy to use, and captured key neurocritical care concepts relevant to further develop expertise in the field. The ecological interface was perceived as an improvement to current interfaces. Recommendations on future interface design are shared and can help future research to further improve interface design.

Conclusion

This research work consisted of multiple phases uniting ethnographic observations, discussions with clinicians (especially novice and expert physicians in neurocritical care), the development stage of neurocritical care-relevant concepts that are lacking in current bedside physiologic monitor interface designs, an iterative design process of potential visualizations to be displayed on such interfaces, and the usability evaluation of the visualizations in context (with novice neurocritical care physicians).

Although the findings in the usability evaluation haven't shown exact confirmation of all expertise development hypotheses we were expecting (e.g., the experimental group should have more confidence), we could still see relevant tendencies of novices using the ecological interface showing better understanding and reasoning patterns of the data shown to them in context. Different designs of visualizations with different variables can be further evaluated in the future, incorporating more and various perspectives of interface users in neurocritical care. Different scenarios can be tested while also other expertise-relevant measures can be considered. More recommendations for future research are outlined in the next chapters.

Contributions

This work contributes to the understanding of ongoing challenges and the importance of data representation on daily-used interfaces and requires in-depth dissemination to support novices in developing expertise in neurocritical care. While this research represented CWA models together with the exploration of expertise in the context of neurocritical care, there is not only a scientific contribution shown in this dissertation but also how technology and specific interfaces can be improved to support different end-users on varying expertise levels. The overall intention is to contribute to a better healthcare system in which physicians can provide best practices and dedication toward patients' health states and share their expertise with newer generations. In the long run, this also has an educational and economic impact, improving hospital resource allocations.

Acknowledgements

I would like to express my deepest appreciation and gratitude to my supervisor, Professor Catherine Burns, who supported and guided me throughout my academic path; You are one of the most important role models for me in life, and research is only one facet of that.

The collaboration within neurocritical care has widened my horizon in many ways, and I am very thankful for learning from Dr. Victoria McCredie and discussing all these very distinct insights with Dr. Alberto Goffi, as well as the unique experts and trainees who have shared their time for this research project.

I appreciate all the thoughtful comments and feedback my committee has provided: Thank you Professor Ning Jiang, Professor Jim Wallace, and Professor Kate Mercer for supporting the development of a more rigorous research project.

The deep discussions I had with various students in AIDL made this research extremely exciting and fun; To Kathleen Schaef: working together on this project made it so much more fun!
To Rayyan Quraishi, Cathleen Grace Leone, Kathryn MacQuarrie, Kirsten Leigh Christensen Anderson, Jennifer Graham, Lina Zaghdoudi, and Sana Allana: you all had unique contributions throughout this research, thank you for all your support and the wonderful insights you brought into this project.
Thanks to all AIDL students, especially Murat Dikmen and Yeti Li; the discussions with both of you have always raised my curiosity in research.

The unconditional love and support I received from my family and friends cannot fully be expressed in this writing; thank you for being there for and with me.

Last but not least, I would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and its CREATE program for providing financial support to run this research and share it with the wider audience at conferences.

Dedication

To my family.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABG	Arterial Blood Gas
ABP	Arterial Blood Pressure
AH	Abstraction Hierarchy
BP	Blood Pressure
CBF	Cerebral Blood Flow
ConTa	Control Task Analysis
CPP	Cerebral Perfusion Pressure
CSF	Cerebral Spinal Fluid
CT	Computer Tomography
CWA	Cognitive Work Analysis
DL	Decision Ladder
DURESS	Dual Reservoir System Simulation
ECLS	Extracorporeal Life Support
ECMO	Extracorporeal Membrane Oxygenation
ECoG	Electrocorticography
EEG	Electroencephalogram
EID	Ecological Interface Design
EMR	Electronic Medical Records
EVD	Extra Ventricular Drain
GCS	Glasgow Coma Scale
HOB	Head of Bed
HR	Heart Rate
ICP	Intracranial Pressure
ICU	Intensive Care Unit
INR	Interventional Radiology
MAP	Mean Arterial Pressure
MMM	Multimodal Monitoring
NIRS	Near-Infrared Spectroscopy
PRx	Pressure Reactivity Index
PtiO2	Tissue Oxygenation Pressure
RBB	Rule-Based Behaviour
SA	Situation Awareness
SAH	Subarachnoid Hemorrhage
SIADH9	Syndrome of Inappropriate Secretion of Antidiuretic Hormone
SRK	Skills-Rules-Knowledge
SSSI	Single-Sensor-Single-Indicator

StrA	Strategies Analysis
TBI	Traumatic Brain Injury
TCD	Transcranial Doppler

Chapter 1 Introduction to this work

Intensive Care Unit (ICU) use in Canada is rising quickly and growing more than acute care hospitalizations (CIHI, 2016). In 2013-2014 alone, more than 230,800 adult patients were admitted to the ICU in Canada. During the pandemic times, this has further increased. Neurocritical care is a specialized ICU and a new area not yet available worldwide. Common patient cases that are handled in neurocritical care are related to diseases or injuries of the nervous system, such as traumatic brain injury (TBI), subarachnoid hemorrhage (SAH), intracerebral hemorrhage (ICH), seizures, or ischemic stroke. It requires dedicated expertise in neurology and intensive medicine. However, physicians with various backgrounds (e.g., neurology, internal medicine, anesthesiology, emergency medicine, neurosurgery) can be trained e.g., through an ICU fellowship program where trainees are exposed to different areas such as neurocritical care. This fellowship program usually takes up to two years in Canada and may vary from country to country. It is a program including very specialized training that trains the fellows in elementary critical care but also support them in building more expertise in this unique area. Worldwide, there are around 2000 neurocritical care experts (Hemphill, 2022), which is too little for treating this area's great volume of patients. Expertise thus needs to be accessible and wider spread to facilitate appropriate and timely care of such patients.

The neurocritical care environment is a complex, dynamic, and safety-critical field. The complexity of this setting relates not only to aspects of the organization or availability of resources but also to how technology comes into play and provides the best support to facilitate needed care for patients. Time is a key factor in this setting, in which patient care must be provided without mistakes. Neurocritical care units often use neurophysiologic measurements that are unique to this unit to assess patients. These measurements need to be learned and applied appropriately to mitigate any risks and errors.

Unfortunately, many human errors happen in ICUs; one study shows 178 activities per patient on a single day in which approximately 1.7 errors per patient were observed (Donchin et al., 1995). There are various error types, but most serious errors are related to medication and can be life-threatening (Rothschild et al., 2005). Thus, the understanding and transfer of information are key; however, the great amount of data and how it is located and accessible to clinicians in critical care causes information overload, poorer patient outcomes, and might involve erroneous actions (Görges et al., 2012 with reference to Donchin and Seagull, 2002). Clinicians have to collect such data from multiple devices, documents, and locations in the unit (De Georgia et al., 2015), bring those together, and make quick decisions. There is a need for more centralized monitors integrating all information streams to understand the full context of the patients' states (Görges et al., 2012). In neurocritical care, clinicians often say, "time is brain" which emphasizes further the attention towards quick actions to prevent further deterioration as little changes in the brain can have tremendous and long-term outcomes.

Recognizing trends in physiologic data is essential (McCredie, 2019). Usually, displays in the ICU show numeric values of the variables measured, using the single-sensor single-indicator approach (Andrade et al., 2020). A representation of the waveform and some calculated values are additionally shown for a few seconds. Some non-continuous measurements are often shown as numerical values only (Andrade et al., 2020). Some studies have shown improvement in clinicians' performance (e.g., detection, response time, accuracy, or treatment efficiency) when new displays were designed and tested (Andrade et al, 2020). However, it is still a challenge for critical care novices to recognize certain patterns in data sets (Fackler et al., 2009) which requires greater attention.

Although many insights through studies have been gained in the ICU environment, this dissertation focuses on neurocritical care specifically and aims to identify neurocritical care challenges, the complexity of data representation on bedside physiologic monitor interfaces that help users make decisions, understand expertise in this field and how it contributes to the development of data representation on interfaces.

The Cognitive Work Analysis (CWA) framework is used to model findings but also helps develop an Ecological Interface Design (EID) to support novice critical care physicians in developing expertise when using different visualizations incorporating key neurocritical care concepts.

In this work, the key areas of contributions and investigations thus are:

- the dynamic and complex sociotechnical area of neurocritical care as the domain of exploration, which is understudied and only has a limited number of experts in the field worldwide. Work domain models were developed and an ethnographic approach has been taken to investigate this work domain and related challenges.
- Expertise development in the context of neurocritical care novices (trainee critical care physicians) and experts (staff critical care physicians).
- The application of CWA to demonstrate novice-expert differences, challenges faced in neurocritical care and understanding the work domain, as well as the development of an ecological interface to support novices in neurocritical care to advance in the direction of an expert.

With these insights, we contribute to scientific research in the area of CWA and EID linked to expertise development (e.g., outlining the novice-expert differences on the decision ladder which help to inform the design of an ecological interface), technology and design by proposing new ideas for a bedside physiologic monitor interface through a systematic procedure, education by providing new ways of thinking about data in a meaningful context that may support individuals learning about key concepts in a clinical context (and in the future potentially in groups or for teaching and research purposes), the economy in the long run potentially for hospital ICUs where better or more efficient resource allocation may be possible, and society who can benefit of better patient care and outcomes, including quicker, efficient, and more individualized treatment opportunities.

In the context of this work, the terms “novice” and “expert” are explored and further defined in the next chapters. However, initial assumptions were made for recruiting novice and expert physicians: novice or trainee physicians were mostly regarded as trainees in the 2-year critical care fellowship program or similar program (also called fellows or trainees), while critical care “staff” physicians who were regarded as experts were chosen based on their years of experience (minimum four years working in the ICU and as an independent physician), with experience in guiding trainees in their teams and conducting research or having a research lab.

Another aspect to mention is the different use of “neurocritical care” and “critical care” in this work; Although neurocritical care is the area of interest, critical care physicians spend a certain amount of time learning and working in the neurocritical care units. Thus, the experts recruited at multiple stages of this dissertation may be working in critical care (as shown in their demographics) but also provide extensive knowledge and experience in neurocritical care and are thus equally relevant to the study process. Trainees in Canada are often enrolled in a 2-year fellowship program in which they rotate through neurocritical care.

Although it might not be appropriate to say that such trainees are novices in medicine, we refer to them as novices in neurocritical care.

1.1 Research questions and hypotheses

This research work involves multiple research questions that lead to the next ones. It is possible to classify 5 overarching research questions involving further sub-questions.

First, we want to examine the work domain of neurocritical care and understand upcoming challenges.

1) What are common challenges faced in neurocritical care?
    - 1.1) How can these challenges be modeled through CWA?
    - 1.2) What challenges do trainee (novice) and staff (expert) critical care physicians perceive?
    - 1.3) How do experts provide support to trainees to overcome such challenges?

The second part of this section will tackle initial impressions on expertise development in the field from the researcher's perspective:

- 1.4) How do physicians develop expertise in neurocritical care?
- 1.5) How are novices tackling tasks differently compared to expert physicians?

By getting more insights into these questions, we hypothesize that there are various challenges daily (e.g., organizational, communicational, etc.) that clinicians have to overcome. Models of CWA (i.e., the Work Domain Analysis and the Control Task Analysis) can help capture, organize, and display the abstractions of this complex area and show differences in how novices and experts deal with tasks. One major challenge is the appropriate use of technology or dealing with a great amount of data to develop timely and flawless patient decisions. Trainees have to deal with a great amount of new input and are challenged by the workload, while staff physicians have more experience and are able to address the issues more concisely. As trainees learn from staff physicians and are supported through various ways of training, observing, and feedback, they approach tasks differently than staff physicians and develop expertise by learning through different sources and feedback that are not limited to the staff physicians.

Although these give a more general overview of the theme, we have to dive deeper to further investigate what it means to be an expert in critical care.

2) How can expertise be described in (neuro)critical care?
    - 2.1) What are the characteristics of experts from the perspectives of novices and experts in the field?
    - 2.2) What differences do they perceive between novice and critical care experts?
    - 2.3) What are the progress indicators for trainees in this field?
    - 2.4) How are mental models conveyed and developed?

For the second part of the questions, we hypothesize that various definitions of expertise exist and may be described differently by novices and experts. As experts may reflect on their progression path and also consider how novices nowadays learn multiple expert-relevant skills, they might be able to detail the big

picture on expertise, while novices may pick up aspects that they are being measured on in protocols and feedback sessions only. Further, we expect differences between novices and experts to be similar to common dictionary definitions but will add more distinct differences beyond years of experience or case exposure. There may be a component related to the equipment handling skills and gaining deeper perspectives into the understanding and relationship of data in multimodal monitoring.
We further hypothesize that progress indicators do not relate only to years of experience but to ways of forming cognitive structures and strategies for dealing with case variations. Lastly, there may be different ways of how mental models are developed and shared with the trainees. However, we hypothesize that experts develop distinct frameworks and try to share them through thinking out loud during ward rounds or other training sessions. At the end of this section, we want to identify ways of supporting expertise development and measures of expertise to be used in the next studies.

Once we know what expertise elements exist especially in critical care, we will merge our findings from the models and expertise development sections and consider ways of integrating them into visualization design on an interface. For that, we will explore how the interface design plays a key role in physicians' way of thinking.

3) How can expertise development be supported through interface design in neurocritical care?
    3.1) How is data on monitors used to develop a trajectory for patient care?
    3.2) Which neurocritical care-relevant variables are thus important to consider?
    3.3) Which neurocritical care-related concepts play a role in developing expertise and why are they difficult?
    3.4) How do the insights from CWA additionally shape the design ideas for the interface?

In this section, we hypothesize that certain key variables, e.g., Intracranial Pressure (ICP), Blood Pressure (BP) and more, are interconnected and considered for coming up with patient trajectory. Insights from experts in the field will help us consider more than trajectory thinking as a key concept; there will be more concepts that relate to neurology-specific indications which the trainees are trained on. These concepts are expected to be challenging to learn due to their complexity in understanding the patient's individual characteristics and data points in context. Finally, we hypothesize that the previously shown CWA models will underline and expand the overall perceptions of the key concepts. Thus, we will have a set of ideas for initial visualization designs for our so-called ecological interface.

To gain perceptions on the proposed visualizations from novice and expert physicians, we hone on the following research questions:

4) How are novices and experts perceiving the static visualizations with regards to usability aspects and how do experts think about the visual representations of the neurocritical care concepts on the interface?
    4.1) From the perspective of novices, to which extent are the visualizations easy to understand and useful in context? What can be improved or shown differently?
    4.2) How are experts thinking about the relation of the concepts to the presented static visualizations? What major concerns do they perceive that need improvement? What can stay in its shown representation?
    4.3) Overall, how can we improve the design of the visualizations?

We expect that novices might not know exactly what they need. At the same time, experts have a better insight into what trainees require to develop more advanced cognitive patterns in analyzing patient cases and thus make better treatment decisions. Therefore, we hypothesize that experts will share more feedback on the way of improving the concepts represented through visualizations. For example, they may find certain visualizations too complex or easy for novices. Novices may provide many ideas for improving the usability aspects such as e.g., customizability of the variables presented on the interface, and give us feedback on which parts need more explanation to make it easy to understand and useful in context. With all these findings, we expect to iterate further on the visualization design and incorporate as many relevant aspects as possible for the final versions of the interface.

The last set of research questions tackle the impact of the visualizations used in a clinical context on expertise development of novices.

5) To what extent does the ecological interface support expertise development in the short-term compared to a standard (currently used) interface?
   5.1) What are the expected strategies used for the presented clinical cases in neurocritical care?
   5.2) Does the group using the ecological interface (experimental group) make fewer errors, as well as less harmful (or severe) errors, when compared to the group using the standard interface (control group)?
   5.3) To which extent does the experimental group provide a more expert-like reasoning and Situation Awareness (SA, level 3), compared to the control group?
   5.4) Comparing both groups again, to which extent are the experimental groups' reflections on strategies outlined in their reasoning, the overall study, and interface design, richer than those of the control group?
   5.5) To which extent are the confidence and performance perceptions of the experimental group higher than the control group (if at all)?
   5.6) To which extent are usability aspects on the ecological interface perceived higher than the standard interface?

In the final part of the research questions, we hypothesize that there are patterns or frameworks that experts use when they deal with neurocritical care cases. Further, the ecological interface will show short-term effects that may be considered expertise development steps. We expect that novices using the ecological interface make fewer errors and fewer harmful errors being indicators for improving expertise. Further, we consider that the reasoning pattern will be closer to those of experts and that the understanding (level 2 SA) and trajectory (level 3 SA) will be better in the experimental group using the ecological interface. Another expertise indicator is the ability to reflect and reconsider other strategies that could have been taken during the scenarios presented. Thus, we expect that we will see more in-depth reflections of the experimental group and better feedback on the interface design's representation of neurocritical care-relevant concepts. Confidence and performance being further measures of expertise may show tendencies of the experimental group feeling more confident in their actions or reasoning, and more secure about their level of performance. Overall, we hypothesize that the usability aspects are better overall on the ecological interface, even if this is a new way of looking and thinking about data, it will have more opportunities to understand and dive deeply into the patient context.

1.2 Overview of the dissertation

Figure 1 below shows the structure of this research work and its overall outline.

This part of the proposed work introduces the general context, research interests, and overall outline. Chapter 2 gives insights into the modeling framework called Cognitive Work Analysis (CWA) and its application in the neurocritical care context. These models will deepen the understanding of the work domain and explore differences among novices and experts within this field of interest. To get a clearer picture of expertise and its development, chapter 3 will show an interview study conducted with critical care physicians to add their perspectives on expertise and its development within their work domain. The different considerations through the interviews on expertise and the CWA models will be merged to develop initial ideas of the interface design used in neurocritical care (chapter 4). Eventually, feedback from critical care trainees and staff physicians will be explored to iterate the visualization design or completely remove specific visualizations (chapter 5).

The final design implementation of the visualizations on the prototype will be shown and evaluated with critical care trainees in a usability study (chapter 6). A general outlook and contributions will be provided at the end of this dissertation (chapter 7).

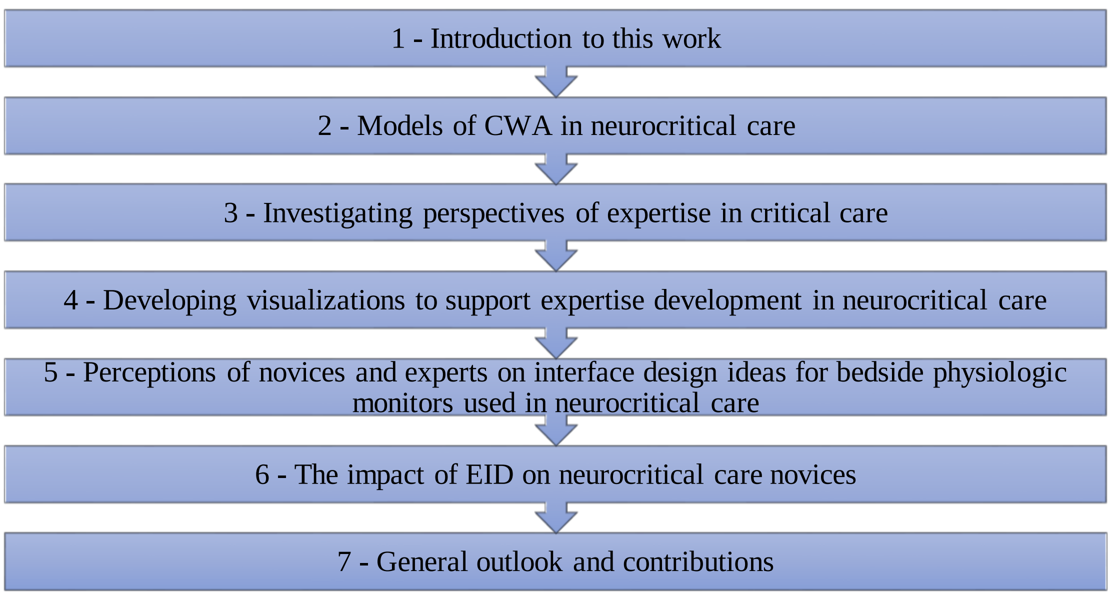


Figure 1: Dissertation outline

Chapter 2 Models of CWA in neurocritical care

This chapter provides an overview of the neurocritical care environment and the importance of expertise in this context.

2.1 Introduction

To get a richer picture of the complex socio-technical area, a learning experience at a mixed critical care unit (general and neurocritical care) was conducted, which allowed for a full immersion into the environment, observing and discussing the challenges of the dynamic space. The gathered findings are modeled through the CWA framework, incorporating the first two steps being the WDA, ConTa in this chapter, and the third step the Strategies Analysis (StrA) in chapter 6.
The CWA framework has been chosen to analyze the various relationships of the complex work domain itself. It lays a basis for the ecological interface design (EID) approach later described in this dissertation. It further provides a way to capture and organize complex information, link these meaningfully, and visualize essential tasks and strategies taken by decision-makers of a complex system.

Guided by the focus of the previously shared research questions, this chapter tries to show ways of answering the following questions. The overarching question is:

- What are common challenges faced in neurocritical care?

Challenges can be perceived in various ways on different aspects; however, we first want to identify examples of general challenges faced in neurocritical care and consider ways of modeling them with the CWA approach:

    - How can these challenges be modeled through CWA?

To also integrate the perspectives of clinicians, specifically those of physicians working in neurocritical care, we try to get insights from them directly:

    - What challenges do trainee (novice) and staff (expert) critical care physicians perceive?
    - How do experts provide support to trainees to overcome such challenges?

To explore the component of expertise in neurocritical care, the questions focus on:

    - How do physicians develop expertise in neurocritical care?
    - How are novices tackling tasks differently compared to expert physicians?

Discussing such aspects with affected stakeholders and observing common processes provides an excellent depth into the research area. However, literature searches can help gain perspective on gaps other researchers have investigated before.

2.2 Background

As a preparation for the planned observations and the procedure outlined in this chapter, background on the importance of neurocritical care and the CWA framework will be provided.

2.2.1 Neurocritical care

There are different types of intensive care units (ICUs) worldwide; some examples of the types are described as the medical ICU, surgical ICU, pediatric ICU, neurological ICU, or trauma ICU. Patients with specific needs for critical situations are admitted to these units and cared for with various technologies and methods for an in-depth inspection and treatment. Patients with neurological conditions such as traumatic brain injury (TBI), intracranial hemorrhage, seizure, stroke, or spinal cord injury are cared for in neurocritical care. Neurological deterioration is life-threatening and can cause long-term disability as there is still the need to find “effective treatment options” (Simpkins et al., 2020).

As stated by Shutter and Molyneaux (2018), there are many neurology patients for whom standard critical care management does not suffice or apply, and there is a need to investigate patients’ states through more advanced technology that is evolving as well as therapeutical options. This multidisciplinary department includes neurointensivists who work with neurosurgeons, neuroradiologists, emergency medicine, and other medical/surgical subspecialties (Mayer et al., 2006)**.** While aiming to support and improve acute neurologically deteriorated patients, a big focus lies in reducing (secondary) neurological injury, examining the overall conditions in context, and helping the patient to recover (Mayer et al., 2006). Lazaridis et al. (2019) describe secondary injury being “characterized by a cascade of biochemical, cellular, and molecular events triggered by the primary insult, and involved in inter-connected pathways of deterioration” (with cross-reference to Kochanek et al., 2000).

Neurocritical care units are growing and are considered a relatively new subspecialty in medicine (Rincon et al., 2007); Some examples can be found in North America, Europe, Asia, the Middle East, Latin America, and Oceania, most of them being academic institutions in urban areas (Suarez et al., 2020) and around 2000 experts worldwide (Hemphill, 2022). This specialized unit requires dedicated and specific expertise to manage intracranial pressure, hemodynamics, and neuromonitoring that support measuring various brain variables and cerebral flow (Rincon et al., 2007). According to Hemphill (2022), there are around “15 million cases of stroke, 27 million cases of traumatic brain injury, and 1.1 million cases of status epilepticus annually worldwide” (Hemphill, 2022) for which there needs to be ways of growing and improving the expertise in the field. Time plays a significant role (Bhardwaj et al., 2004) in this setting, as small measurement changes may indicate deterioration requiring immediate medical support. Getting a better grasp of neuromonitoring and ways to measure various aspects of the patient’s body can impact treating proactively rather than reactively. Neuromonitoring is advancing but still isn’t fully used by all neurocritical care physicians and trainees. However, multimodal neurocritical care monitoring facilitates multiple aspects, such as individualized care, early detection of neurological deterioration, and response to therapy (Yang, 2020). Considering multiple (neuro)physiologic variables in the assessment also supports getting a better sense of trajectory, proposing therapeutic interventions, and considering short- and long-term risks. All these indications and equipment need to be learned during the training in the neurocritical care specialty. The increasing trends of Artificial Intelligence used for tackling Big Data challenges are becoming more significant also in the complex area of neurocritical care. Although the information is being represented on neuromonitoring, there is a need for “making full use of all the information,” and data analytics and AI can pose support in exporting expertise and applying it to the many patients in need (Hemphill, 2022).

2.2.2 The Cognitive Work Analysis (CWA) framework

With the development of new technologies and their capabilities, we face new challenges that need to be examined carefully; especially those used in complex work domains require a different design approach and consider the interaction between humans and machines in context. It is important to consider how operators of such systems perceive and understand what is represented to them to handle possible situations in various, normal and abnormal, cases. In complex systems, risks of unexpected happenings or fault situations are high, and operators need to be ready to deal with critical and abnormal situations immediately. Measured data from the devices is represented through interfaces that have slowly advanced in the past decades; One initial way to directly show elemental data from sensors is called the single-sensor-single-indicator (SSSI) approach. However, operators need to understand the data in context and deal with higher-order state information (especially in fault situations) not explicitly represented on such displays. A lack of showing relationships between different variables while dealing with SSSI displays poses a higher burden on operators. Taking a step towards understanding human capabilities and incorporating the affordances of a work domain has thus been considered. (Vicente and Rasmussen, 1990)

The more 'integrated' display design approach pulls elemental data together and represents it more directly, potentially reducing the cognitive workload and errors while improving performance (Christoffersen, 1998).

Some of these considerations go back to literature proposed by Rasmussen (1985) in which an ecosystem can be represented and structured in higher and lower order elements through a so-called Abstraction Hierarchy (AH). This approach tries to show a more goal-oriented system that identifies constraints and the degrees of freedom available for the operator while enabling flexible adaptation. The AH is part of a more holistic approach embedded in the Cognitive Work Analysis framework used to organize and analyze complex sociotechnical systems.

Different perceptions exist on CWA being a conceptual or a theoretical framework (e.g., Fidel and Pejtersen, 2004).

Ravitch and Riggan's (2017) definition on conceptual frameworks is described as "An argument about why the topic one wishes to study matters, and why the means proposed to study it are appropriate and rigorous" (p.5), as cited in Crawford (2020). Crawford (2020) also mentions that it "informs and describes the development of research questions, design selection, data collection, data analysis and presentation of findings" (pp. 36-27).

However, a theoretical framework is defined by Ravitch and Riggan (2017) as "In the case of theoretical frameworks, the "parts" referred to in this definition are theories, and the thing that is being supported is the relationships embedded in the conceptual framework. More specifically, we argue that the parts are formal theories; [sic] those that emerge from and have been explored using empirical work" (pp.11-12), as cited in Crawford (2020).

In this work, we approach CWA from a theoretical framework perspective in which we try to elaborate on the research gaps and propose ways of further exploration by using theories to inform the understanding and analysis of neurocritical care and the research questions.

The CWA framework can be divided into five significant examination steps: Work Domain Analysis (WDA), Control Task Analysis (ConTa), Strategies Analysis (StrA), Social and Organization Cooperation Analysis and Worker Competencies Analysis as mentioned by Vicente (1999).

In this research, we mainly target a deeper understanding of the neurocritical care environment and will model our findings with the WDA to better comprehend the system boundaries. Using ConTa, we will examine the core tasks and decision flow. The strategies key decision-makers take in the critical care context will be displayed through the StrA. Further, the models created with e.g., the WDA and ConTa, can help to identify information support at certain stages; this can be explored further when using the models when designing an ecological interface (EID) that allows the user to become an adaptive problem solver, also in unanticipated situations.

Work Domain Analysis (WDA)

The Work Domain Analysis portrays the activities within a setting where required or expected goals are to be fulfilled. To capture all possibilities in such a setting, Rasmussen et al. (1990, p.41) describe that the "[…] work domain must identify the entire network of means-ends relations relevant for the system considered, […], the requisite variety which is necessary to cope with all the requirements and situations the system might face". As the constraints of a system vary, there are different levels of abstraction on the hierarchy. This is often referred to as the means-ends hierarchy or AH.

According to Hajdukiewicz and Vicente (2004), a "Work domain analysis represents information that is only implicitly captured, if at all, in a task analysis. And by identifying information requirements that are event-and time-independent, work domain analysis provides a robust basis for supporting worker adaptation to novelty and change. […] work domain analysis can play an important and unique role in shaping how well actors achieve their goals and select their actions".

The AH consists of the *Functional Purpose* (what was the work domain designed to do), *Abstract Function* **(**what are the underlying laws or principles), *Generalized Function* (what are the processes that are involved), *Physical Function* (what equipment is involved and what is its capability), *Physical Form* (what is the physical appearance and location of that equipment). (Burns & Hajdukiewicz, 2004)

The WDA has been applied in various healthcare contexts. For example, a recent study by Austin et al. (2022) examined emergency department care and identified various complex and interconnected aspects impacting time and safety prioritization. Also, system constraints such as availability of resources (e.g., computer) have an effect on the healthcare processes and impact patient safety.

Another recent study used the WDA approach to portray the complexity and constraints of the ICU environment with regard to patient mobility, in which nurses' decisions are supported through different sources. The levels of abstraction show the interconnectedness of patient and hospital unit information for which the design of decision-support tools could be beneficial. (Krupp et al., 2022)

A further field explored in healthcare is the radiotherapy technology. The WDA by Wu et al. (2012) identified the complex and safety-critical area of radiotherapy treatment. This research group explored new radiotherapy controls and an interface that would incorporate the system constraints and relationships to support radiotherapy delivery.

Control Task Analysis (ConTa)

While the WDA enables the operator to adapt to new situations and overall changes, the Control Task Analysis cannot do this. For analysing the control tasks, "[…] a class of events must be explicitly identified (or implicitly assumed) before a task analysis can even be started" (Hajdukiewicz and Vicente, 2004). The

ConTa represents the tasks or cognitive steps involved in reaching a specific goal in a process of control (Vicente, 2003). A way to visualize these steps can be done through the so-called Decision Ladder in which data-processing activities are portrayed in rectangular shape upon the model, and states of knowledge resulting from data processing as circular shape.

Strategies Analysis (StrA)

The StrA considers various ways of performing a specific task and is thus related to the ConTa. Depending on the situation, the strategies vary. Further, the strategy chosen by the decision-maker, user or operator, depends on factors such as workload, familiarity or experience etc. and may be switched throughout a task. A common way of showing this is by representing an information flow map of the selected area of interest. Compared to the WDA and ConTa, there are fewer studies conducted on StrA.

The ConTa and StrA have been applied in a study where healthcare teams need to communicate and make decisions in a surgical setting (Ashoori et al., 2011). In Cesarean sections, various teams need to collaborate while working towards their individual goals of supporting the patient. This study investigated how workflow as well as cognitive tasks are represented and shared between the team.

Another study used CWA analyses to identify "the goals, work practices, role assignments and tasks of trauma resuscitation teams." Using various data collection methods, the research group looked at team errors that happened and the overall collaborative practices. (Sarcevic et al., 2010)

In the area of cardiac nursing, a study was conducted by Burns et al. (2008) in which the WDA, ConTa and StrA were used. The Conta incorporated a closer look at the "modes of operation, the sequencing of steps, and shortcuts between tasks" that would be an addition to the WDA. With regards to the StrA, the researchers found out that there are different questioning strategies such as open-ended, standardized, topographical, hypothesis and test, or ruling out possibilities type of questions that are asked to the patient.

Few studies exist that used the StrA in the healthcare domain. One example can be shared by Shanteau and Sengstacke (2009) who have outlined a hypothesis-and-test strategy in the context of medication decision when observing nurses, physicians, and pharmacists in the medical-surgical and intensive care unit.

The CWA framework can help with various ways in domains of application. Specific examples of the CWA application in nuclear power, aviation, healthcare, transportation, automotive and more are demonstrated in Stanton et al.'s (2017) book.
In this chapter, the WDA and ConTa will be shown, whereas the StrA will be depicted in Chapter 6.

2.3 Methods

Multiple research methods have been used in this and later parts.

2.3.1 Combining research methods

We have used qualitative research approaches in this chapter and other chapters of the dissertation. According to Creswell (2014, pp. 246-247), qualitative research is defined as "[…] a means for exploring and understanding the meaning individuals or groups ascribe to a social or human problem. The process of research involves emerging questions and procedures; collecting data in the participants' setting; analyzing the data inductively, building from particulars to general themes; and making interpretations of the meaning of the data. The final written report has a flexible writing structure".

To understand the neurocritical care work domain, we have used an ethnographic approach which "[…] is a qualitative strategy in which the researcher studies an intact cultural group in a natural setting over a prolonged period of time by collecting primarily observational and interview data." (Creswell, 2014, p.242) The advantage of observing professionals in their natural work environment is that it is possible to get internal insights into the setting. However, the collected insights through the observations are limited to the specific times of entering and observing such a work environment.

In this research work (Üreten et al., 2020), the (neuro)critical care unit at Toronto Western Hospital, Ontario, has been examined within an approximately 100 hours timeframe. Various teams have been shadowed including physicians (staff, fellow, resident), nurses, and respiratory therapists.
As an observer, it is possible to see the wide extent of the work domain and the various facets from an outside and potentially more objective perspective. However, it is beneficial to also understand the depth of such observations by discussing with the professionals during the observations or interviewing them later to perceive the nuances and subjective opinions and experiences they have. During the observations, discussions took place among ward rounds or within short or longer breaks with various clinicians to better understand procedures, treatment plans, and how they learn new skills.

The two phases of CWA, i.e., WDA and ConTa, have been used to structure observational findings on the AH and DL that incorporated elements acquired through discussions with the clinicians, too. The discussions with physicians however often included topics related to how challenging the development of expertise is and thus created an important topic to explore further.
As discussions during observations were limited in time, we wanted to dive into the nuances of perceived challenges and the relation to expertise development further. Interviews were chosen as they are a "systematic form of asking people for research purposes – either in an open form with an interview schedule or in a standardized form similar to a questionnaire" (Flick, 2011, p. 249) and help to deepen the insights observed.

The interviews took place after the observational phase with physicians only, to specifically identify their perspectives on neurocritical care challenges and ways of mitigation as they are the main decision-makers in the unit. Interviews with other clinicians have been conducted too, however, are not represented in this dissertation as the focus was to first identify the main decision-makers' perspectives and can be explored further in future research. As there are different roles of physicians, some in the training phase while some being trainers, we wanted to gain different standpoints on challenges faced in neurocritical care to potentially identify ways of supporting them through interface design.

There are different types and styles of interviews (e.g., fully structured, semi-structured or unstructured). We have selected a semi-structured interview style, where "the interviewer has an interview guide that serves as a checklist of topics to be covered and a default wording and order for the questions, but the wording and order are often substantially modified based on the flow of the interview, and additional unplanned questions are asked to follow up on what the interviewee says." (Robson, 2011, p. 280)
This style of interviews gives freedom to the wording of questions, time dedicated to the questions and the order of questions asked in general, while having main topics prepared upfront. Interviews can also be combined with other methods, as done in this dissertation.
Similar approaches have been used in studies such as by Wu et al. (2012) where ethnographic observations, domain expert feedback have been combined for coming up with the WDA in radiotherapy treatment. Another example can be found in the research conducted by McNamara et al. (2015).

Two interviewers (the author of this dissertation and the master student Kathleen Schaef) were present throughout the full study duration to ask questions and simultaneously take notes on the responses. The interviews overall took around 60 minutes but had different phases for the topics covered. One part of the first 30-minute phase included questions on challenges faced in critical care training and how these are mitigated. In this chapter, only those aspects will be shared. In the next chapters, more insights and questions of these interviews will be further discussed.

We have looked for trainee physicians (novices in neurocritical care) as well as staff physicians (experts in neurocritical care) in Canada and the U.S.A. When searching for trainees, we looked for critical care physicians who are enrolled in the 2-year fellowship program in Canada. Staff physicians have been included as experts if they had at least 4 years of experience in the ICU being in independent practice, had trainees working in their unit whom they guided, had a research profile (having their own research group or conducting research in critical care) and having experience in neurocritical care.
The occupation of participants was an inclusion criterion as also literature states that professions are linked to expertise, such as mentioned by Cioffi (2012).
It shows the type of background and training a participant has completed. Educational backgrounds vary worldwide and so do the skills each individual has learned in their educational systems. Some healthcare providers in rural areas have training constraints and might lack access to training possibilities that would enhance their skills and abilities (Brems et al., 2006). This could mean that specializations or developing expertise in some areas can vary. Participants may also have gone through multiple educations or trainings (even from different domains) and thus questions around their educational background can help to get a better picture of variations in their base and sub-specialty training of the novices and experts. ICU sub-specialty trainings are often internal medicine, general surgery, emergency medicine, anesthesiology, respirology and rarely cardiology, nephrology, neurosurgery, hematology.
Gender may play a role in expertise of different domains. A study on emergency medicine showed that female and male residents received different kinds of feedback (Menchetti et al., 2022). Bakken et al. (2003) investigated physicians' perceptions of their performance and found out that female physicians indicated lower abilities to perform and apply knowledge or skills when linking to clinical research than male physicians rated themselves. Although gender might not be the primary indicator in this study to determine

expertise in neurocritical care, we might encounter some relations to gender that can be investigated in future studies.

Morse (2000) states that the sample size depends on the scope of the study, the nature of the topic, the quality of data, the study design and the use of shadowed data. Different opinions on determining the exact sample size for qualitative research exist. Findings by Vasileiou et al. (2018) show that provided justifications for sample size in qualitative health research are limited and also depend on the journal of publication. According to their findings, justification was mainly based on saturation and pragmatism. They recommend considering data adequacy and saturation parameters when coming up with sample size assessments and maintaining quality, trustworthiness, validity and generalizability. Their findings reveal that mean numbers of interviews were between 18 to 44, across 3 journals. Mostly used criteria to justify sample size were saturation and pragmatic considerations (time constraints, access to specific study populations).

In qualitative research, there are many discussions on how to judge quality of qualitative research and different approaches exist to take a step towards describing rigour. This is often related to trustworthiness and Hadi and Closs (2016) summarize strategies of ensuring trustworthiness in qualitative research. Examples Hadi and Closs provide are: triangulation, self-description/reflexivity, member checking, prolonged engagement, audit trail, peer debriefing, and thick description. (Hadi and Closs, 2016)

While triangulation incorporates aspects of conformability and credibility, self-description/reflexivity aim to decrease researcher bias. Member checking tries to support dependability and credibility, while prolonged engagement is a method to spend more time and build rapport and trust with participants. In audit trails, it is important to provide as much detail as possible to show truthfulness of findings. Peer debriefing is another way of increasing trustworthiness and credibility. Thick descriptions also support transferability and credibility by detailing out study information and procedures.

Similar rigour criteria are discussed and can be found in the list showing internal validity, external validity, reliability and objectivity with various operational techniques summarized by Tuckett (2005).

In this research, many details are shared about the study conduction, recruitment of participants, processes, and analyses to ensure rigour and "standards for reporting qualitative research" as in O'Brien et al. (2014).

The ability to connect with the researchers virtually was another factor to consider as the interviews took place during the pandemic and restrictions on in-person meetings existed. The interviews were conducted via the virtual platform Microsoft Teams. They were recorded, transcribed, summarized, and shared with participants to clarify unclear transcriptions or statements and to change or add anything to their statements.

Participants were recruited with the support of Dr. McCredie, who forwarded an email for recruitment to program directors of the University Health Network and ICU directors within Canada and the U.S.A.

All participants were offered remuneration or alternative formats, such as gift cards.

The interview study received ethics clearance from the University of Waterloo Research Ethics Board (#42892), and the materials used in this study can be found in the Appendix.

2.3.2 Data analysis

To structure the observational findings, the CWA framework has been used. The insights gained on the neurological, monitoring aspects as well as challenges have been drafted upon the AH along the five levels of abstraction. Further, the DL has been taken to reconstruct an example of a common neurocritical care case that physicians with different levels of expertise tackle daily.

The various themes related to expertise development have been identified during the observations which will be listed and explained in the next section.
To understand the insights gained through the interviews, an inductive thematic analysis approach was used (i.e., to develop a theory instead of testing a theory) in order to identify the themes or data patterns that emerged when participants shared their answers. The data was coded which is “[…] the process of organizing the material into chunks or segments of text and assigning a word or phrase to the segment in order to develop a general sense of it.” (Creswell, 2014, p. 241)
The codes were derived from the data (i.e., through an inductive approach) which offers an exploratory mindset. This was followed until saturation was reached: “Saturation is when, in qualitative data collection, the researcher stops collecting data because fresh data no longer sparks new insights or reveals new properties.” (Creswell, 2014, p. 248)
The themes were coded by the author of this dissertation and Kathleen Schaef separately and later discussed the themes they identified for the codes. By checking the perceptions of the codes, interpretations could be discussed, and the likelihood of subjectivity and bias could be reduced. Further, the clinical collaborator Dr. McCredie has been involved in the (final) discussions to ensure scientific rigour and provided guidance overall, as well as in situations where certain topics needed to be discussed due to different perspectives.

The six-phase approach of a thematic analysis is represented by Braun & Clarke (2012).

The first phase is about familiarizing oneself with the collected data. In this study, interviews were recorded and transcribed. This way, it was possible to revisit all conducted interviews by relistening, highlighting important aspects, and adding some comments to the text that help analyze the interview. It also provides the opportunity to think about the meaning critically. Thus, reading the data multiple times is beneficial.

In the second phase, initial codes are built and show “the content of the data” (Braun & Clarke, 2012). Codes can be either close to what has been described, i.e., descriptive or they can be interpretative. It is common that codes may change over time when going through the data; it is important that the codes include the diversity as well as the patterns coming up in the data.

The third phase relates to establishing themes or common patterns across the codes. This step can also incorporate generating subthemes where certain clusters are visible among codes but overall unify multiple subthemes into one major theme. Further, themes can be regarded as having a relationship with one another and provide direction for answering the research questions. There is also the possibility that some codes don’t match well with any theme for which a miscellaneous theme can be created.

Within the fourth phase, potential themes are being reviewed. This implies that the quality of the themes is being checked with regard to the fit with the associated data. The quality can be based on the themes usefulness in relation to the research question. The review process also tackles the understanding of boundaries of themes to think about what is incorporated and what is not. The data may be diverse, so the question arises if the themes are coherent and set up in a meaningful way. In this step, it is possible to get rid of themes, merge or rephrase them in the context.

In the fifth phase, themes are defined and named. These shall especially show the connection to the research question; naming the theme appropriately means that it should be "informative, concise and catchy". It is also possible to have quotes as the title as they refer to the way of how the participant has phrased and thought about it.

The sixth phase relates to the production of the report. In any thesis or report, the analysis is made to make a point and provide arguments for answering the research question or telling the story. Among the themes, there should be a logical and meaningful way of sharing this with the readers.

2.4 Results

The results portray insights of the observations that are modeled through CWA, while observed and discussed aspects of expertise in neurocritical care will be listed below. General discussions on challenges perceived by critical care trainees and staff physicians as well as the way staff physicians support the trainees will be depicted through the thematic analysis approach as mentioned by Clarke and Braun (2012).

2.4.1 The complex socio-technical environment of neurocritical care – a learning experience

Observations were made at the ICU of Toronto Western Hospital which included a neurocritical care unit, over an approximately 100-hour period. Different ICU teams were shadowed including 2 respiratory therapists, 5 nurses, 4 residents, 4 fellows and 3 senior intensivists. The weekly routines were observed including ward rounds, different procedures and treatments within the ICU, training sessions, communication and organizational coordination, and the use of equipment in the ICU. The overall intention of the observations was to gain deep insights into the complex and dynamic neurocritical care unit and get a better understanding for challenges faced. Another interest in the observations laid on the way of how clinicians deal with the vast amounts of data and how they use it to come up with decisions.

The input from the observations were firstly modeled with the WDA.
The Abstraction Hierarchy (AH) is based on Rasmussen's (1985) directions on defining the work domain as well as its constraints. As mentioned by Burns & Hajdukiewicz (2004), the AH consists of the following 5 levels of abstraction:

- *Functional Purpose* (looking at what was the work domain designed to do),
- *Abstract Function* **(**what are the underlying laws or principles),
- *Generalized Function* (what are the processes that are involved),
- *Physical Function* (what equipment is involved and what is its capability),
- *Physical Form* (what is the physical appearance and location of that equipment).

The chosen work domain that we try to represent here is the complex and dynamic neurocritical care environment which has previously been published in Üreten et al. (2020) showing the first two points below. In this chapter, we add a third dimension to the AH:

1) the nervous system and cognition, to show general physiologic relationships of the main neurological functions and link to cognitive processes (Üreten et al. (2020))
2) neurocritical care monitoring and treatment, to lay out the existing tools used for treatment options (Üreten et al. (2020))
3) neurocritical care challenges, displayed on various categories such as the organization and environment, technologies, patient care and educational opportunities for clinicians.

The full representation of the AH can be seen in Figure 2. This approach has been chosen as the complex sociotechnical environment of critical care requires a mapping and understanding of the biological systems of the patient (i.e., neurological considerations) as well as the technical system (i.e., monitoring, etc.) that the patient is surrounded with. The third dimension added observed challenges looking at patients, technology but also further aspects such as the environment, organizational and educational parts. All these elements have been intertwined. Level by level, each element is connected with others through means-end links. Similar approaches or examples are represented in other research studies such as Miller and Sanderson (2000).

Each column will be explained in the following.

1) The nervous system and cognition (also represented in Üreten et al. (2020)):

Regarding a top-down approach of the **nervous system and cognition**, the *purpose* can be described as the coordination of body movement, interpretation of sensory input, effects responses, mental activity, monitoring the internal environment and maintaining homeostasis. The regulation of the organ system, the control of thoughts, movements, emotions, desires, oxygen, as well as temperature, heart rate and release of hormones display the *balances level*.

Processes in the nervous system and cognition are the information transfer between the brain and body, the use of energy, respiration, healing processes, the involuntary control of muscles, reflexes, the contraction of skeletal muscles, as well as the control of smooth muscle and glandular tissue in the digestive system.

The *physical function* level can be divided into three parts: the body, systems and organs. The body refers to the patient, the systems describe the central system and the peripheral nervous system. The organs considered here are the brain, spinal cord, heart and lungs.

On the lowest level, the *physical form* can be represented as several entities; the patient type (as in terms of e.g. adult or pediatric), the patient conditions (obese, elderly, frail etc.), the connection of central nervous system to (sensory) organs, muscles, blood vessels and glands, the disease (in terms of type, state), medication (type, dose, form of administration, side effects), damage (type, amount, location), monitoring sensors and equipment (location, capability), as well as drains (location), ventilator (settings and type of ventilation) and their side effects.

2) Neurocritical care monitoring and treatment (also represented in Üreten et al., 2020):

Looking at **neurocritical care monitoring and treatment**, the top-down approach can be described as the following. The *purpose* is to improve the patient condition, facilitate long-term stability and enable independent living (quality of life).

What needs to be *balanced* are patient goals (such as quality of life, advance directive) and physiologic states (comfort, healing).

Different *processes* of influence can be described as medications, physical interventions (ventilation, draining) and the prevention, detection and management of secondary brain injury.

The level representing the *physical function* can be summarized as ventilators, drains, sensors, monitors, medication, and medical staff. The *physical form* level can be described and summarized as the same way as mentioned in the nervous system and cognition.

3) Neurocritical care challenges

The previously described **challenges** observed within critical care have been categorized into the levels of abstraction. Enabling the best care for patients is always the *purpose* of critical care units.

This can be *balanced* by different factors such as looking after the patient, providing a good organization of processes and environment, enabling up-to-date training and education of staff as well as technologies that support the staff within their processes of care.

For each of these categories, several *processes* are involved such as processes of medical care and communication among different stakeholders, quality improvement, sharing knowledge and maintaining or acquiring needed systems or devices.
Each of these aspects come along with their capabilities and *physical functions* which are for example treatment plans for patients, providing a safe, clean, silent, and comfortable environment, enabling training opportunities and conferences for knowledge exchange and technical platforms that support clinicians by visualizing and measuring patient health information.
The *physical forms* level thus encapsulates all details such as capacities or locations of the above-mentioned functions. Examples are alarms that contribute to noise and (dis-)comfort or visualization tools of patient data such as interfaces or monitors.

With a WDA and its representation on the AH we were able to set our first system boundaries which show the scope of the environment, e.g., we include the patient and equipment. Our constraints would be the actions and relationships between the variables represented on the AH, i.e., we have constraints by equipment, kinds of treatment, ethics, health-directives, staff or even rooms.

Although concerns with multiple aspects such as technology, learning and others have been observed and voiced by clinicians, there is great potential for not only providing clinicians with advanced technology but to support them developing expertise in specific areas such as monitoring or analyzing complex relationships. What yet has not been described are representations of decisions that are met by clinicians when they go through certain procedures of e.g., patient treatment or patient admittance. In the next section, the intention is to provide an example that was constructed through observations to address and show how clinicians act within the constraints.

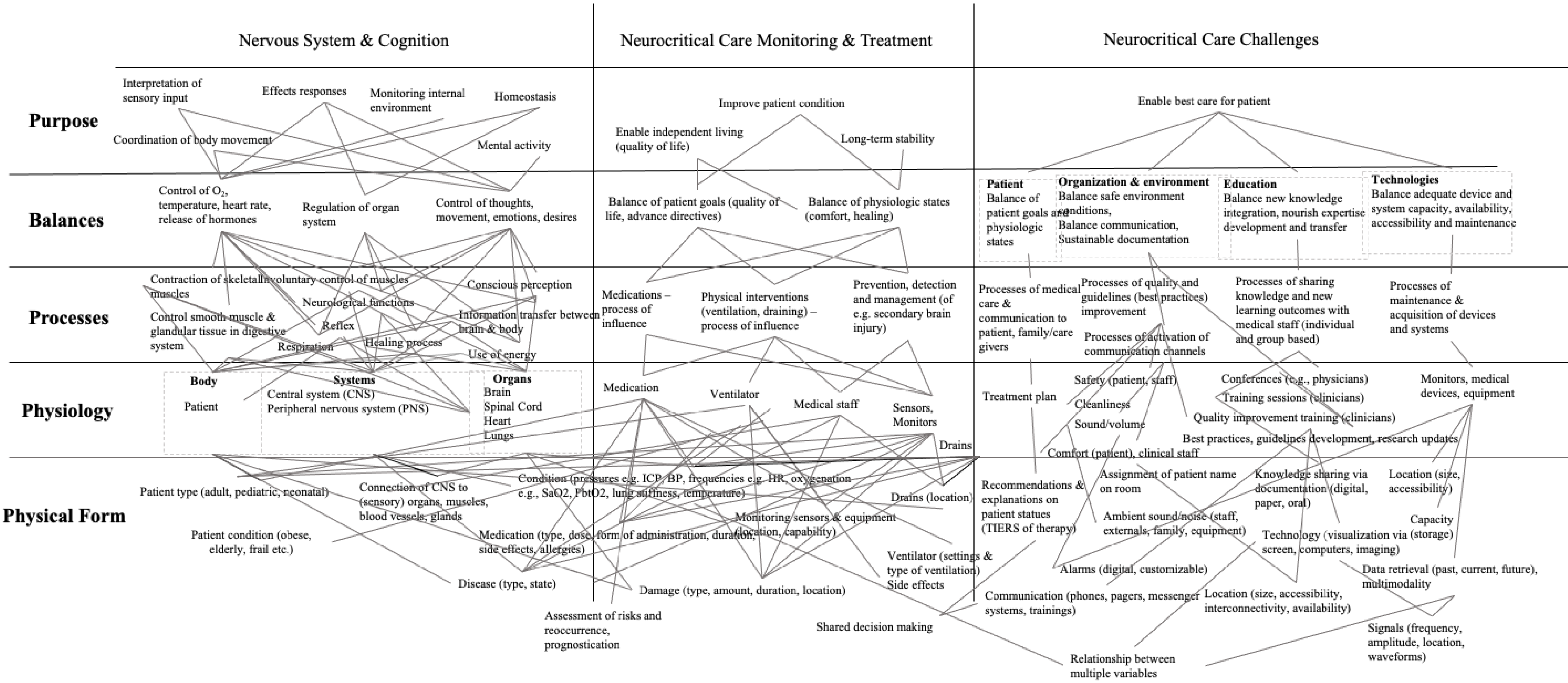


Figure 2: Abstraction Hierarchy of Neurocritical Care

2.4.2 Observations on expertise development in neurocritical care

In the context of expertise development, the following ways to develop expertise could be identified through the observational phase at the Toronto Western Hospital. Inputs from clinicians are also shown as findings below and thus sum up to a broader diversity of ways to develop expertise.

Medical study program: Medical school in Canada covers various courses (lectures, tutorials, labs, practical exercises) which support the acquisition of theories and knowledge with little exposure to practical implementation and real cases at the beginning. Students learn from medical books, online sources and observing professors and teaching assistants who are in most cases physicians themselves.

Certifications: To show that clinicians have achieved certain proficiency and acquired all necessary knowledge for a specific field, they can take part in trainings which are a type of formal measuring of learning outcomes. Within these trainings, participants learn via observations, theories, common practices, and mistakes. Trainings vary in duration and extent and thus can be performed multiple times and in various directions.

Ward rounds: Trainees observe seniors in action treating and assessing patients. The daily plan is made for the patients by the trainees and discussed with the seniors during morning rounds. All required procedures need to be performed after the ward round until the night shift takes over. Here, trainees are guided by seniors on mental strategies (e.g., how to gather data, bring them in a context and interpret them) and practical implementation (e.g., how to use sonography). Routines are built by following the discussion format: repeating patient background and reason for admission, reading data from multiple systems, listen to other clinicians' updates on data and observations, understand the current situation and develop a projection into the future. The treatment plan is made after the senior speaks to the patient and interprets bedside monitor information and other assessments.

Individual efforts for expertise development: Learning from books and journals about neurocritical care are examples of how clinicians or trainees acquire knowledge. Additional efforts such as by asking questions can be directed to seniors or fellows if procedures are unclear or trainees need support. The trainees explain their concern and approach the senior. The senior does not only provide the right answer but also fosters reasoning and adjusts the trainee's plan to current learning needs.

Trainee rounds: Seniors explain specific cases that are relevant for the development of expertise. They ask trainees questions about the treatment plan, medication and critical situations that might happen at the unit and thus foster to reflect and think independently. Seniors give feedback on their mental strategies or argumentation as well as correct them if necessary. They also provide them with similar and non-routine cases. The senior actually tries to understand during open questions and discussions how trainees logically order their answers if they present alternative solutions (e.g., treatments) and if the concepts are modeled correctly in their mind.

Guidance and instruction when introduced to a new case: Seniors and fellows explain the treatment procedure, e.g., intubation, they verbalize each step and why they do it while demonstrating the procedure practically on a mannequin (e.g., intubating the mannequin) and then ask the trainee to repeat the procedure on the mannequin on their own. The seniors observed the trainees' behavior and interfered in situations when an irreversible mistake could happen.

Orientation on Standard Operating Procedures (SOPs), checklists and advisor manuals: Common practices can vary from hospital to hospital. Trainees learn and apply the specialization- and hospital-specifics to the cases by reading the guidelines and listening to their mentors. Checklists for dealing with critical cases support the trainees in learning which data sources they need to check and conduct the assessment as well as treatment for a patient. Advisor manuals are used by seniors to guide the trainees throughout their training and give an overview of cases and expectations on the trainees' skills. These manuals also help the seniors to assess the trainees' skills and exposure to certain cases.

Absence of guidance and instruction: Trainees need to take responsibility for the patients during the night shift and seek help from on-call seniors if they need assistance in urgent cases. All important and routine procedures need to be finished during the morning shift when the trainees are guided by the seniors actively. Also, this is part of letting the trainee take over more responsibility gradually.
Further, trainees are supposed to prepare the list of all patient information (so-called "clinical messages") on a template which includes all case relevant information (e.g., patient profile, treatment plans etc.) from different systems. The summary of all information supports the learning process in gathering data, categorizing them and show most relevant aspects in a context. The trainees bring these printouts to the morning conference with seniors and practice soft skills.

Repeated practice: Some procedures are repeated until they are automatized, e.g., the intubation is performed and practiced many times. The variety as well as increasing complexity in patient cases are parts of the repeated practice to recognize and know how to deal with uncommon or varying (e.g., anatomical) cases. Through repetitions, routines manifest themselves.

Trial and error: For learning practical skills trainees undergo trial and error procedures first on one mannequin, then they repeat the procedure on multiple mannequins with varying conditions. The next step is to apply the procedure on a real patient in control of a senior.

Recall: Seniors test trainees throughout the day by asking theoretical and fundamental medical questions. They then direct to connect the theories to a current case and try to get the trainee engage in recalling important data points and bring them in a meaningful context. Thus, trainees do not only recall information but also process them and develop understanding of a context.

Deliberate practice: The expectations on the skills and knowledge increase throughout the residency and fellowship. This comes along with stepwise improvement by taking over more responsibility, dealing with more cases as well as more complex situations, and thinking as a team. These steps are usually monitored and guided by seniors.

Treating the patient: The trainees actually prepare for example by intubating the patient, and then provide treatment, e.g., in form of medication. This is done by the presence of a senior. The senior watches the practical work of the trainee and reacts if a mistake is done by the trainee or the situation gets critical. The seniors interfere by either stopping the procedure or in most of the cases take over and correct as well as finishing the treatment. Different strategies exist because some seniors react prior to the critical point, and others interfere at the critical point. Also, there seem to be differences in the way they transfer their knowledge because while some seniors take over the procedure, they communicate clearly, what the mistake was, what could happen to the patient and how the procedure is done correctly, whereas others silently accomplish the procedure in a correct way.

Mentorship: Trainees have a mentor (senior or senior fellow) throughout their training. Mentors support the learning process specifically at the beginning of training to set learning goals in accordance with their interest and capacities. They provide constructive feedback and assess whether the trainee has reached the set goals. Further, they explain how they can improve, what their weaknesses are and how to address them, where their strengths are and how to foster self-acquisition of expertise. Also, positive feedback reinforces the learning process.

Supervision: For critical treatments such as in time-critical events, the trainee is constantly guided and led by a senior. Also, the trainee is observed and controlled more intensively, more so during the residency. In less critical procedures where consequences of a wrong decision or wrong treatment are less impacting, trainees are generally given more responsibility to perform the tasks they suggest.

Communication among teams: Teams consist of different clinical personnel. Responsibilities are divided among them, i.e., nurses take care of daily patient care, while e.g., trainees learn to gather information from different systems (e.g., bedside monitors). Updates are given throughout the day (also verbally) and seniors share tips during informal talks when the team moves to the next patient bed. As all roles are fixed, trainees usually learn from staff physicians but not as much about thinking from other clinicians' perspectives. Although they learn to incorporate the suggestions of other clinicians, they might not completely be able to come up with suggestions themselves.

Conferences: Morning conferences prior ward rounds deal with urgent cases that need to be discussed (e.g., attach drainage). Other conferences take place at noon where external staff give an update on the unit's performance (e.g., antibiotics). This gives the clinicians a way to critically think about the next steps of the unit and also portrays a directive on what they should try to target among the next weeks. Further, trainees can voluntarily give a short presentation about a topic they are interested in sharing (e.g., medication on ICP patients). If anything is incomplete, seniors give verbal feedback directly in the plenum and add missing information.

Feedback: Verbal, written, direct, indirect feedback as well as from various people (clinicians, patient relatives, patient) and systems (monitors, diagnostic tools) are given to the trainees. Reflecting about target goals and achieved goals are part of the provided verbal feedback from seniors. Verbal and written feedback is given at the end of training within the unit. An example for direct feedback is when seniors tell whether the performed actions were correct or incorrect. A way of indirectly giving feedback would be by e.g., correcting the trainee in performing their task. Feedback from systems can be shown in forms of incorrect actions e.g., ordering the wrong medication for a patient. In general, experts try to support the trainees to decrease the amount and severity of mistakes.

Exemplary behavior: Seniors are not only mentors but also portray role models for professional behavior and expertise. During morning ward rounds, seniors demonstrate to trainees how they interact and assess the patients' status by communicating very clearly and sympathetically with patients and their families. They often use analogies from daily life to explain to the patient and their relatives what their medical issues are. Trainees learn how to communicate with the patient and their relatives because of observing seniors' behaviors.

2.4.3 The Control Task Analysis of experts and novices in neurocritical care

A better understanding of the complex neurocritical care environment can be gained through modeling certain steps that are needed to perform an action. The ConTa shows the tasks that are involved in reaching a specific goal in a process of control (Vicente, 2003). A model of knowledge for visualizing these steps is the so-called Decision Ladder (DL).

An example is provided in the following to show the decision steps and the human information processing for a neurocritical care patient, for which different processing routes exist; the DL being part of CWA shows a clear difference between novices and experts (Burns, 2019). When expertise is further developed, information is processed quicker and thus leads to the execution faster (Burns, 2019).

Although common examples of the ConTa are shown when a process is started, shutdown or in normal operations, this example displays a typical scenario (high intracranial pressure case) and how physicians process the information provided to them.

Our model of the DL is mapped on two levels of the main decision-makers in neurocritical care, namely the *expert* (staff) and *novice* (trainee) critical care physician. These two levels are associated to their expertise in which we compare a so-called "novice" critical care physician, who would be a trainee enrolled in a fellowship program, and an expert critical care physician, whom we consider as a staff or attending physician with completion of fellowship as well as multiple years of independent practice within critical care. The DL will show states of information processing starting from an activation, up to an evaluation, and down to an execution point. All specific steps will be described in the following.

2.4.3.1 ConTa of a high ICP patient – Experts

Starting from the "Activation" box on the left bottom, the approach would be to follow the arrows up until "Evaluate" and then go down to the "Execute" box (see Figure 3). Some shortcuts can be described by the dotted arrows and are often taken by expert physicians due to great knowledge and experience of practice on similar cases. Cognitively higher demand would be visible on the upper part of the decision ladder which involves the thinking process of evaluation. Experts use shortcuts and are quicker in processing the information to execute a task. Thus, the upper levels (around "Evaluation") are often not even entered anymore due to their stronger mental models. However, in situations of uncertainty, they would run through the evaluation steps again.

In this example, the decision ladder shows that experts would be notified through an alarm that is either given by any technical system in the ICU or by team members such as the nurse. They would observe pupil dilation, ICP numeric values, its trend as well as waveform pattern, and rule out certain conditions. An example for such a condition is given here as herniation syndrome. The set of all observations can consist of multiple aspects they put their attention on, such as ICP numeric values, trends and waveform patterns. Noticing abnormalities in any of these indicators would also raise the consideration of checking the attached equipment on the patient. A common indicator for changes in ICP is equipment failure, so the physician would check whether the Extraventricular Drainage (EVD) is flushing or not and confirm this with looking at the waveform changes. The EVD requires distal flushing in this example. The physician would have multiple factors in mind that let the situation seem ambiguous; the increase for ICP might be caused through the position of the head of the bed (HOB), high-temperature levels, abnormal blood gas values, equipment malfunction or other reasons. The evaluation process is a cognitively high demanding state. For this step, the physician would evaluate the standards of care, HOB, temperature, arterial blood gas (ABG), ICP, neuroexamination, Computer Tomography (CT) scans and literature findings that show similar cases as the one presented. The ultimate goal would be to stabilize the patient which requires a decrease of ICP. Interpreting the situation again, all parameters and report findings will be taken into consideration. To best support the patient and improve the condition, it is of utmost importance to set individualized thresholds for each specific patient. This might be to get the ICP value down below 20 mmHg. The acting physician would carry out the standards of care, i.e., all TIERS that are recommended for treating intracranial hypertension. An expert physician would additionally recall non-evidence-based trials from the literature if the patient doesn’t fully react to the TIERS as intended to. Again, all gathered patient information will be reassessed and other divisions in the hospital (e.g., neurosurgeons, radiologists) will be alerted to support with optimal management options if needed to stabilize the patient again.

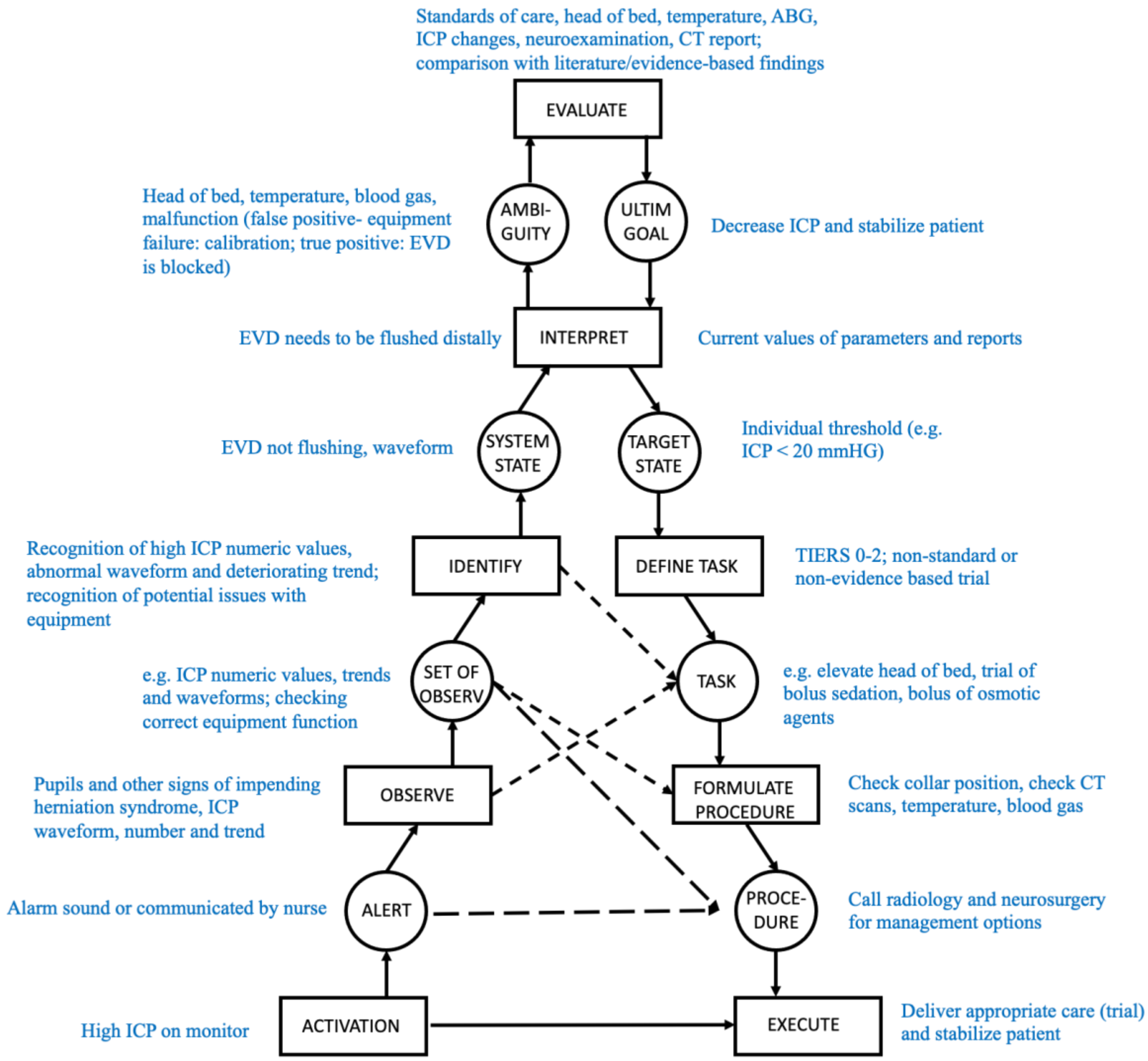


Figure 3: Decision ladder of expert neurocritical care physicians

2.4.3.2 ConTa of a high ICP patient – Trainees

Trainee physicians take the role to learn and accomplish the preliminary work for preparing the patient for the staff physician's examination but at the same time, learn to handle cases and increase autonomy. There are many parts that are challenging for trainee physicians that will show some differences in the below Figure 4 as compared to the DL of an expert physician.

In this case, it is often difficult for trainees to understand the waveforms and numeric values in a context. In most situations, they would only focus on the numeric values and not start trending as well as checking the actual waveform patterns. Further, flushing the EVD and calibrating the equipment are not straight forward tasks for them. They might not think about all limitations or potential malfunctions of the equipment attached to the patient. Trainees are made familiar with standard procedures of care and often do not keep abreast of the literature that might support the evaluation process. Depending on the trainee's level of experience, they might be able to deal with different levels of TIERS. It would be expected that advanced trainees are able to perform all TIERS, whereas trainees at the beginning of their program might be able to work with the first levels of care only. Another strong focus of trainees in general is that they stick to the standard ICP ranges that might be depicted in the literature or guidelines and lack an understanding of individualizing their approach to each patient. A crucial point for trainees to learn is to recognize when to ask for support by their team members or staff from other divisions.

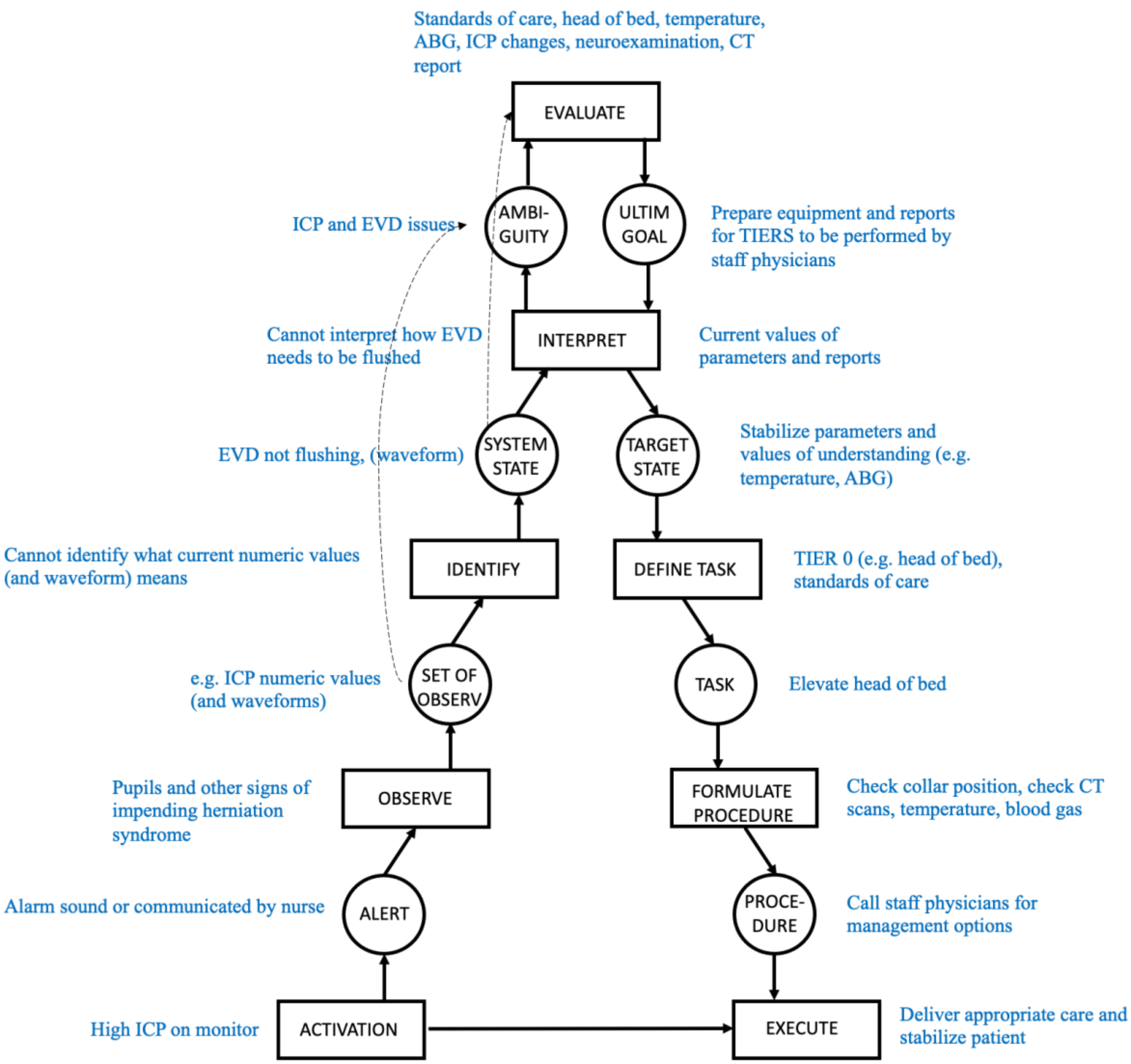


Figure 4: Decision ladder of novice neurocritical care physicians (trainees)

2.4.3.3 Shortcuts

The decision ladders also show that novice and expert physicians have different strategies in how they deal with critical cases. Experts use many shortcuts as they have seen similar situations more often than trainees and acquired ways to quickly address critical steps. Experts for instance, often anticipate the risks and know what the task would be in case they receive an alert for the patient. Or at a later stage, after gathering all observations, they might quickly set a target state or directly formulate a procedure to prevent deterioration.

Trainee physicians however might take shortcuts in situations when they are certain of their capability and know what they are equipped with in urgent situations. In some situations, they might be able to step up the ladder quicker, e.g., from the set of observations to considering ambiguity, or the system state may lead them to evaluate specific things directly. Trainees would usually try to go to the upper parts of the decision ladder and try to assess the patient with all they have learned theoretically and practically up to that point. Mental models of trainees have not yet developed to the extent of that of an expert but can be improved through multiple ways as mentioned above in literature and findings of the observation. Some of these relate to the sequence of actions, the efficient use of appropriate sources or tools as well, the interpretation of data in a context, the evaluation of various possible outcomes, the application of expected TIERs and the time when to contact other clinicians.

Although some elementary directions and differences between novices and experts can already be identified through the DL, it does not portray the specific reasons why they perceive tasks and information differently. The case represented is one typical scenario in neurocritical care, but various scenarios could be taken into consideration and compared for more similarities or differences in the information processing stages of novices and experts to have a more generalizable standpoint. At this point however, a more in-depth examination can help to identify the nuances and explore the differences further from perspectives of the physicians themselves to develop design ideas to support the development of expertise more in this context.

2.4.4 Interviews with critical care physicians

Identifying challenges through observations is one way of diving into the complex and dynamic ICU, however, trainees and staff physicians were also asked to share their thoughts on the struggles of trainees and aspects that are hard to convey and teach.

In total, 36 participants (18 trainees, 18 staff physicians) were recruited for interviews.
Demographic information was collected and is represented in Table 1 and Table 2. The average age of staff physicians is 49 years with 5 female and 13 male participants. The average year of independent practice is 16, while 5 participants are from the USA and 13 from Canada. The years of experience also varied between 4.5 years to 32 years. Participants mentioned the following patients cared for in their ICUs: medical, surgical, trauma, neuro, oncological, emergency, and postoperative cardiac surgery. All staff physicians work in university hospitals, while one is active additionally in a community hospital and another one in a tertiary care center. All staff physicians confirmed that they have trainees working in their ICUs. The abbreviation of “S” in the first (left) column stands for Staff and will be found in the text too.

Table 1: Demographic information of the staff critical care physicians

S#	Age	Gender	Years' Experience	Specialty	Subspecialty	Country
S1	39	M	4.5	Internal Med	Respirology Critical Care	USA
S2	58	M	26	Neurology* Vascular Neurology*	Neurocritical Care*	USA
S3	48	M	17	Internal Med	Critical Care Neurocritical Care	CA
S4	54	M	23	Internal Med	Critical Care	CA
S5	55	M	18	Anesthesia*	Critical Care Neurocritical Care	CA
S6	51	M	20	Anesthesia	Critical Care	CA
S7	44	F	15	Internal Med	Critical Care	CA
S8	45	M	8	Internal Med* Emergency*	Neuro-critical Care	CA
S9	39	M	8	Internal Med	Critical Care	CA
S10	46	F	15	Neurology* Vascular Neurology*	Neurocritical Care*	USA
S11	60	M	32	Anesthesia	Neurocritical Care	CA
S12	63	F	28	Neurology*	Critical Care*	USA
S13	50	M	19	Internal Med	Critical Care	CA
S14	56	M	23	Internal Med	Critical Care Neurocritical Care Epidemiology	CA
S15	64	M	30	Pediatric Med	Critical Care	CA
S16	36	F	5.5	Neurology	Critical Care Neurocritical Care	CA
S17	41	F	9.5	Internal Med	Critical Care	CA
S18	40	M	7	Neurology* Neurophysiology*	Critical Care* Neurocritical Care*	USA

* indicates certification received in another country.

The other targeted population for the interviews were critical care trainees or so-called fellows who are considered novices in (neuro-)critical care.

The trainee critical care physicians (abbreviated with a "T") were similarly asked about their age, gender, year of ICU fellowship, specialization, and subspecialty training. Their average age is 34 years excluding

T1 who did not share her age. In total, 11 females and 7 males participated, all being trainees in Canada at the time of the study conduction. The common duration for the critical care fellowship program is 2 years in Canada although there are possibilities to extend the training.

Table 2: Demographic information of the trainee critical care physicians

T#	Age	Gender	Year of training (in CC)	Specialty	Subspeciality (pending)	Country
T1	ND	F	1	Internal Med Cardiology	Critical Care	CA
T2	35	F	1	Emergency Med*	Critical Care	CA
T3	31	M	1	Emergency Med*	Critical Care	CA
T4	34	F	1	Internal Med	Critical Care	CA
T5	34	F	2	Internal Med* Pulmonology*	Critical Care	CA
T6	47	M	2	Internal Med*	Critical Care	CA
T7	39	F	1	Cardiology* Immunology*	Critical Care	CA
T8	36	F	1	General Surgery*	Critical Care	CA
T9	36	F	1	Anesthesia*	Critical Care	CA
T10	34	M	2	Internal Med*	Critical Care	CA
T11	30	M	1	Internal Med	Critical Care	CA
T12	35	F	4	Anesthesia*	Critical Care	CA
T13	31	F	2	Anesthesia	Critical Care	CA
T14	29	F	1	Internal Med	Critical Care	CA
T15	31	F	1	Emergency Med*	Critical Care	CA
T16	30	M	1	Emergency Med*	Critical Care	CA
T17	30	M	1	Critical Care*	Critical Care	CA
T18	34	F	2	Anesthesia* Critical Care*	Critical Care	CA

* Indicates certification received in another country.

2.4.4.1 Perceptions on challenges of trainees and staff

Trainees were asked to share what they perceive as struggles during their critical care fellowship training, while staff physicians were asked to comment what they find hard to teach trainees.

In this part, six main themes came up on struggles related to: training, workload, knowledge, skills, guidance, and motivation.

Training - theme:
Trainees shared that the specialty training has an impact on perceiving certain things more challenging, for example, trainees have different levels of expertise and comfort such as trauma or advanced cardiovascular ICU and thus are more advanced with airway management when compared to others (T1). The different specialty programs have various focusses and train the novices in different aspects of knowledge and practice in the ICU (T2).
In addition to that, it was mentioned that fellows are often interested in the ECMO certification but when they face the practical part, they mostly don't finish for which the reasons are unknown (T12).
Generally, another trainee mentioned that it is the knowledge and application of knowledge into clinical practice that is often a struggle for them (T13).

Workload - theme:
The physical as well as the mental side of workload has been identified as another challenge for trainees. Trainees said they deal with tough call schedules, are often tired and there are time limitations to additional readings that attending physicians recommend and teach (T2).
The time constraints to teach and consolidate learnings outside of work (T11), reacting in stressful situations (T5) and coping with a lot of uncertainty in a very high-pressure situation (T4) were discussed too.
Adjustment to the intensity of the ICU (especially with the pandemic) as well as dealing with burnout is a big challenge (T10).

From the perspective of staff physicians, one mentioned that trainees are often more distracted (S2).
But it is also hard to teach and learn how to cope with mental health and develop strategies, dealing with emotional trauma and wellness ("healthy living outside of the hospital to ensure longer sustainability inside the hospital to be successful") (S10). "Doctors do a very poor job of training and educating trainees on those aspects". (S10)
Another aspect mentioned is the challenge of making people (i.e., trainees) slow down and make sure they have a real love and precision of their words, because the precision of their words reflects their precision of thinking (S7). And the joy of that; If people aren't inspired by that, it's very hard to teach that (S7), "[…] once people have that, we can have lots of conversations. Until you've experienced the joy, then you are really committed to it." (S7)

Similarly, (S9) added: "[…] The desire not to relax especially when having seen the case before. It is the passion even when you're tired, it's not the knowledge, it's really the attitude that is the most difficult thing, and empathy with the family". (S9)

Knowledge – theme:
The subtheme content was highlighted by trainees and staff physicians. The vast amount of content knowledge (T4) and literature (S1), as well as the theoretical competence (T6) were mentioned. Further, (T8) added that trainees often don't know the main concept and lack experience, "they require some examples and maybe a summary of what to do as the first thing".

Another subtheme identified under knowledge is critical thinking.
Some assume that the theory is most challenging but usually that's not the case (T17). Procedures can also be learned; there are manual and technical skills to be learned at the beginning, but later, "the more trainees know the more they doubt". (T17)

Staff physicians however mentioned various aspects about critical thinking:
Putting all pieces of information together and understanding what that means is one aspect trainees struggle with (S4). Another staff physician (S9) also noticed that trainees often focus on the big picture and miss out the details that are significant.
Although the struggles depend on the background, it is mostly challenging for trainees to recognize how quickly things can change (might be an important warning sign), e.g., sometimes a little change in language can be a warning sign as well, or subtle things in the exam/movement (S13).

This is also related to the use of technology: there is a knowledge barrier that people (i.e., trainees) need to overcome before they touch the computer; “Having all information at the bedside available is only impactful for a few people who have gone through the one-on-one week workshop; others complain that they don't have any view form into the data. Many fear that it's way too complicated and would take too much time to gain anything from the data”. (S11)
MMM, “we do it all differently (timescale, relationships between variables, prioritization); there is a 2-week rotation in which they deep dive into each modality and how you interpret everything from a global standpoint.” (S18)
A further comment on the technology was about troubleshooting the equipment and checking if it's working properly (S13). The morphology of the waveform or absolute numbers need to be checked before extrapolating. (S13)
“The waveform analysis is hard, but it has to be considered with the bigger picture and there is very limited evidence that focusing on waveforms actually in a meaningful way changes outcomes for patients” (S12). Trainees often deal with knowledge gaps "in what treatments in reacting to changes on neuromonitoring signals are actually helpful for interpreting a monitor if there is uncertainty about what actions should follow a change on that monitor". […] It's hard to teach about any monitoring technique and I will say that beyond neuromonitoring, even just any monitoring technique, it's hard to teach the importance of interpreting a monitor if there's uncertainty about what actions should follow a change on that monitor and this is true of many monitoring techniques that we apply in critical care units or to detect acute physiologic derangement, it doesn't mean I don't buy into some of this monitoring being important, but it's hard to teach that it's important to get can't actually convincingly argue that a treatment that I will provide in response to this change, will meaningfully impact on my patients. That would come, I guess related to that the fact that there is not consensus about what neuromonitoring approaches are mandatory or necessary to improve patient outcomes because we don't have such evidence." (S12)

Skills - theme:
Multiple subthemes were identified. Both trainees and staff physicians related to procedures that seem challenging. Procedural things in challenging circumstances with life-or-death consequences (T4) are challenging: “procedures scare us the most, especially without supervision” (T7), in general the practical competence (T6) and being hands-on (T13) have been stated multiple times.

Staff physicians outlined that they see struggles of trainees on doing things consistently or continuity (S1). There are varying learners and varying career paths to know how much (patho-)physiologic and nuanced technical knowledge to teach, it's hard to know. Procedural things especially if someone is lacking visuospatial perception, technical hand skills etc. are further aspects that are hard to teach (S16).

Diagnosing is another challenge for trainees (T18). “As you realize anyone can do that by practice and that it’s just a skill, you begin to put more attention in the knowledge like clinical decisions and quality of care you are giving and what is the best practice, because everyone can do a procedure if they are trained, but not everyone has the ability to take the best decision at the best time. […] You learn and teach someone, and it gets better with time.” (T7)

Autonomy and confidence came up as further challenges: “You need to be confident to try things on your own” (T11) as well as making a major decision (T15). Some come from a micromanagement approach; Boosting confidence and morale to develop skills is challenging (T10). This aspect especially has been described in various and not always continuously increasing pattern:
One “[…] builds confidence and then it goes down, that's something trainees struggle with. Females struggle more with building confidence, building their voice. It is important to become a leader and what style of leadership you have in critical care. […] Sometimes you might have the expertise in your brain but you're not able to convey it. This is something people struggle with as they learn, a self-awareness of your level of expertise and conveying it to your team, females and minorities have more difficult with that.” (T14)
The language, culture, and system (e.g., every hospital has their own computer system) also play a role in adapting (T5). “Some trainees stay silent in stressful situations until they get confident and can deal with many things, another type of person is they ask many questions because they don't want to make it wrong. (T5)

One staff physician related to the technology-side of this aspect: there is a confidence barrier that people need to overcome before they touch the computer (S11).

Communication can be described as another subtheme under skills. (S16) highlighted the struggles trainees have with communication, especially with patient families.

Trainees however noticed there is a relation to the culture as well; “if you know the culture, it is much better with communication, and the connection to the attending physician develops over time.” (T5)
There is a personal component too, “introverts might also struggle and learn when to call for help. Sometimes you run into more trouble if you are more of an extrovert or loud especially when there is little room for errors.” (T14)

Guidance – theme:
Under this theme, trainees mentioned that having someone who guides you along the way and provides feedback by showing whether you are going in the right pace and right direction is important. Exams are there for testing knowledge and skills, but this is not the only part. There is also a part about going outside of just clinical skills and knowledge as a physician. Fellows are often not directly assigned to a specific staff. (T9)
Another trainee (T16) said that communication with the staff is important but also challenging; He said that he would prefer the trainees telling the staff directly in case they need assistance to be sure if the procedure is done in a safe way. Thus, he highlighted that showing the procedure first and explaining the steps is the

initial approach (to less experienced trainees) while he would then observe how the residents for example proceed with the previously shown actions and would then provide feedback. (T16)

2.4.4.2 The ways of how staff try to support with challenges faced

Trainees were additionally asked to share how they are being supported by staff physicians when they face challenges. Three main themes were identified: teaching, feedback, training.

Teaching - theme:
Often, staff physicians support trainees with teaching sessions and presentations (T2) for theoretical and practical competencies. Discussions about the patients (T2), creating a good environment, trying to be more empathetic and careful with the feelings of everyone has increased in the last years, "the learning process is more satisfying and less scary." (T7)

Feedback - theme:
Two types of feedback were mentioned by trainees: "[…] they would comment positively to enhance their confidence and might send a comment to the program director what should be improved. The staff is very careful" (T5). There is formal feedback provided at the end of the rotation, too. After each procedure there is a checklist report that one can send to the supervisor (T14).
Oral feedback is also provided: "The staff is very approachable and chatting with them about certain topics and challenges faced" is common (T9). Others mention, it is based on your own comfort you know when to talk to the staff (T11).
"At the end of the week, the staff often provides short feedback (what you are good at and things to improve). There is an academic advisor, they receive all feedback from the staff and provides it to you. Informal feedback can be like a conversation because you have interest in their experience. They share experiences and become friends and learn about each other." (T14)
These ways of receiving feedback also become mentorship in many cases "you develop mentorship, relationships with most of them. They mentor you on how you've grown and what you should work on. (T14)
Feedback can also be implied in different ways: "it depends, some staff let you make the decision, others direct you, some make the decision for you." (T15)

Training - theme:
Leadership has been outlined as another way of supporting trainees in developing their skills. "There is a curriculum for developing your leadership skills and personality (it can also get very emotional)." (T14)

2.5 Discussion

Various ways of observing the work environment and identifying key challenges were used in this chapter. The combination of these methods helps decrease bias and get richer insights. Some aspects overlap, and others are unique to the method applied.

2.5.1 Reflections on the study

The observations were limited in time, and there are constraints to learning from one specific unit. This could be further extended to multiple neurocritical care units to compare for similarities or differences. Further, a neurocritical care unit in Canada may have a different work culture or even access to technologies than other countries, impacting the impressions outlined in the models. The CWA models may vary from different perspectives and observations drafted by authors coming up with the AH and ConTa. Trainees included in the interviews often had international experiences and may have outlined a greater variety of insights on expert support or the type of challenges (e.g., there may be cultural challenges that trainees have to overcome). There may be further limitations related to the observations on expertise development from an external perspective; many of the expertise development ways were experienced through attending e.g., training sessions or shadowing clinicians, however, some of the listed elements came up through conversations as well. The provided list is thus not limited to those aspects only, and there may be other ways of further developing expertise. This also may vary based on the location of the unit or hospital, accessibility, economic constraints, and type of clinician.

As part of the discussions or interviews, there was only very limited time available to dive into the depth of the questions as this was part of a longer interview study with various topics explained later in this dissertation. The topic on challenges can certainly be extended in future investigations, as multiple aspects of challenges can be further examined.

2.6 Conclusion

This chapter outlined possible approaches to find answers to the research question: What are common challenges faced in neurocritical care? How can challenges be modeled through CWA? What challenges do novices and experts in critical care perceive? How do experts support trainees to overcome challenges? How do physicians develop expertise in neurocritical care and how are novices tackling tasks differently compared to experts?

Although there might be various ways to identify key challenges faced from different perspectives, a more observatory approach has been selected to look at the general work happening in neurocritical care objectively. Discussions with clinicians have then been picked up to gather specific challenge-related aspects from key stakeholders themselves. The inputs provided to answer the research questions are multi-faceted and can be pulled apart in various directions. However, this research project intends to deepen perspectives on perceived challenges, expertise development, and how these two aspects can be combined to improve neurocritical care technology (specifically, the bedside physiologic interface that plays a key role in expertise development).

2.6.1 Findings of WDA, observations, discussions and ConTa

The key findings from the observational phase were modeled with the **WDA** first, showing different layers of abstraction when looking at the more physiological side of neurology, the technology side, and the challenges related to the patient, organization and environment, education or training, and technology. All these observed areas are interconnected but show the ultimate goal of improving the patient's condition and providing the best possible care. The technology aspect (monitoring and treatment) outlines the various balances and processes enabled through nowadays available technological systems. The challenges however relate to various aspects but show that training or education plays a key part next to technology used in this dynamic area. Many overlapping aspects are identified and show that training in neurocritical care is heavily related to the understanding and use of equipment, too.

The **observations** further show that expertise is developed through various ways in neurocritical care. Examples are not only limited to having more of a trial-and-error approach or learning through repeated practice, but also when cases are discussed by going through key data points during ward rounds, and when there is absence of guidance and instruction. The latter especially entails that trainees have to learn by themselves how to make sense of the data represented to them through monitoring systems.

The **discussion** with clinicians on specific challenges they are often facing additionally showed that trainees struggle with building their knowledge, skillset, coming from different specialty trainings, dealing with the great amount of workload, and have a need for guidance.
Staff physicians' support overcome these challenges by providing more teaching sessions, giving feedback and training trainees on leadership aspects.

The **ConTa** especially showed differences among novices' and experts' process going through a common neurocritical care case when dealing with high Intracranial Pressure (ICP). One of the highlighted aspects are that experts take more shortcuts and evaluate the information faster than novices. Novices often don't recognize all relevant information displayed to them which makes the evaluation and treatment process not always ideal (often incomplete due to lacking information, or slower response to key physiologic issues).

All in all, there are various challenges that can be tackled, and further research is needed to dive into the many facets of challenges seen in neurocritical care. Many of the challenges mentioned relate to expertise and can potentially be overcome and trained when gaining more expertise in the dynamic neurocritical care area. Staff physicians try to support the trainees in multiple ways to prepare them for a more independent and advanced role in critical care.

To get further insights and details about expertise development in critical care directly from novices and experts in the field, a more in-depth investigation on expertise is portrayed in the following chapter.

Chapter 3 Investigating perspectives of expertise in critical care

3.1 Introduction

Expertise is a multifaceted topic. There is a strong need for developing expertise, especially in complex sociotechnical domains in which dealing with critical tasks is elementary.
Although some general aspects of expertise have been pointed out in the previous chapter, this chapter introduces the topic of expertise by providing insights into existing literature about expertise and investigating the perspectives of critical care physicians in the healthcare domain.

The research questions foreseen for this chapter are as follows:

- How can expertise be described in (neuro)critical care?

Exploring more details will show us how critical care physicians think about the characteristics of experts they know or have seen:

- What are the characteristics of experts from the perspectives of novices and experts in the field?

Further, they may be able to share what differentiates them from novices:

- What differences do they perceive between novices and experts in critical care?

The expectation on this matter is also to identify certain skills or indicators when expertise is being developed:

- What are the progress indicators for trainees in this field?
- How are mental models conveyed and developed?

First, the literature will show directions and general explanations on the tackled questions and will be shared before data collection.

3.2 Background

This section shows literature about expertise in general, expertise development, and how this can be measured.

3.2.1 Expertise and expertise development

In Ericsson's (2018) first chapter, a definition is given of the word "*Expert*" taken from the Webster's New World Dictionary (1968, p. 168) as "one who is very skillful and well-informed in some special field" and by referencing Wikipedia as "someone widely recognized as a reliable source of knowledge, technique, or skill whose judgment is accorded authority and status by the public or his or her peers. Experts have prolonged or intense experience through practice and education in a particular field" (Wikipedia, as cited in Ericsson, 2018) Characteristics, knowledge, and skills are different among experts and novices. These are referred to as "*Expertise*". (Ericsson, 2018)

The skills, related to expertise, can be presented by the 5-stage model of mental activities in directed skill acquisition by Dreyfus et al. (1980). They argue that a performer depends more on *concrete experience* at later stages than abstract principles required at early stages.

Stage 1: Novice. Generally, the start of the instruction process is a *decomposition of the task environment into context-free features*. *Recognition* does not require any experience. *Monitoring*, *self-observation* or *instructional feedback* help to improve the novice and thus gets in accordance with rules.

Stage 2: Competence. After a reasonable amount of experience, competence will develop. This requires dealing with real situations involving *repeating patterns*. Within this stage, an instructor can provide guidelines that include these patterns or so-called aspects and thus expects ways of *corrections* of specific conditions by applying the guidelines.

Stage 3: Proficiency. The performer at this stage has acquired more *practice* and has *experienced* multiple situations. Thus, the performer already has *memorized principles* that help to respond appropriately.

Stage 4: Expertise. Reaching this stage, the performer has already experienced various situations and learned *analytical* ways and would thus develop *immediate reactions* that are *intuitive* and appropriate as the next step. This is the *highest level of mental capacity* a performer has reached.

Stage 5: Mastery. In this stage, the performer can absorb even more moments at which point his/her performance can *exceed* common high levels. The expert who no longer requires principles can direct his/her performance towards certain perspectives and actions.

Among all five stages of skill levels, mental functions change or develop among *recollection, recognition*, *decisions,* and *awareness*. Training shall help the performer to reach the next skill level stage.

In addition to this existing model of mental activities and their relation to skills, different levels of expertise were identified by Ericsson et al. (2007). His studies showed that the *amount* and *quality of practice* display factors for showing levels of expertise. One should rather look for a coach who helps learning to coach oneself through *deliberate practice* which he defines as "considerable, specific, and sustained efforts to do something you can't do well- or even at all". Three tests are required to pass to call one an expert. These are composed of the ability to *exceed the peer's performance*, the *delivery of concrete results,* and the *ability to replicate or reproduce* and *measure expertise* in a laboratory. (Ericsson et al., 2007)

When developing expertise, one *thinks less* and acts more out of *intuition* which relates to *deliberate thinking*. Through *practice*, *reflection* and *analysis*, it is possible to improve abilities to take decisions. Expertise can thus be improved over time with *consistency* and carefully *controlled efforts*. On the other hand, experts also lose the ability to deeply analyze a situation and evaluate the appropriate responses since they react automatically to many situations and trust their intuition. (Ericsson et al., 2007)
Moreover, identifying a good *expert coach* helps to accelerate the learning process, and it is fruitful for the performer to receive *constructive feedback*. In the context of expertise, "*Expert performance*" relates to "superior reproducible performances on representative tasks". (Ericsson 2018)

Although expertise development is present in all domains, this dissertation focuses on expertise development in healthcare. Norman et al. (2018) also state that *deliberate practice* and *reasoning* in clinical diagnostic reasoning as well as surgery is a big part of medical expertise development. They point out two types of reasoning through gathered literature within clinical diagnostic reasoning. Type 1 is related to *cognitive shortcuts* or *heuristics*, being *fast*, *unconscious*, and *automatic* which are acquired through previously dealt cases. Type 2 relates to *methodical*, *systematic* and *load on working memory* and thus is *slow*, *analytical* and *effortful*, rather relating to disease processes and illness scripts (i.e., cognitive scenarios that entail signs and symptoms associated with diseases). For medical students for instance, it is important for them to practice *explaining* signs of a disease, symptoms, and their roots. Norman et al. (2018) also argue that first, *analytical and experiential knowledge* contribute to clinical competence (technical psychomotor skills till reasoning and decision-making), second, *individual attributes and aptitudes* influence expertise and third, there is a complex relationship between *expertise* and *experience*. Findings of interviews with exceptional physicians showed what they had in common were continued *deliberate learning from practice*, *skills to organize and classify patient cases*, *reflection onto performed activities* and *continuity in trying to improve performance*. Another important factor of the reasoning process is to *generate hypotheses*. Differences between students and experts were identified within hypotheses accuracy levels: experts would come up with more accurate hypotheses. (Norman et al., 2018)

Within the medical domain, expertise has been investigated at different levels. A research study by Patel et al. (2000) explored that memory is not the greatest factor that affects performance but rather *recognizing patterns*, *organizing knowledge,* and *connecting recent information to their prior learnings*. Over time and with advancing experience, performing tasks become automatic and experts become more adaptive to deal with various situations. On the other hand, to reach excellence, it is required to "counteract tendencies toward automaticity" (Ericsson, 2011, p. 117).

Differences among experts and novices can also be seen in their mental models.
"Mental models are naturally evolving models. That is, through interaction with a target system, people formulate mental models of that system" (Norman, 1983). However, mental models are often incomplete, unstable, and do not have firm boundaries. (Norman, 1983)

Patel et al. (2000) explain that experts would have more connections between the pieces of information they learned over time and thus can rule out unnecessary data. In contrast, a novice would not have great knowledge or skills and an intermediate would be in between: having knowledge but not as interconnected yet, thus he or she would engage in unnecessary "elaborations in explaining patient problems". Developing a better understanding and evaluation of tasks, issues and appropriate proceedings, an expert has acquired the required expertise to assign difficulty levels and time to manage these cases strategically. Communication, coordination of tasks, and resource allocation are other important factors required within complex, dynamic fields to solve a problem efficiently. As proposed by the authors, external support could help students and novices to enhance their learning phase, especially providing a structure.

3.2.2 Measuring expertise

Staff physicians often have the role of an academic advisor and thus train novices (i.e., trainees) daily. Manuals for staff physicians lay a basis to guide novices and capture relevant milestones. According to some hospital manuals (confidential) trainees are expected to make a certain *number* of *clinical assessments* including *related issues*. They further need to evaluate whether results (e.g., scans) are *interpreted* and *selected correctly* and whether *plans to monitor* patients' reactions to treatments are chosen appropriately.

Moreover, advisors have to evaluate whether trainees have *consulted* other clinical personnel timely and engaged in *team decision-makings*. *Prioritizing* tasks, using *appropriate medical language*, considering *feasibility* and *availability of resources* are further milestones that need to be acquired over time. Demonstrating *different skills* (e.g., visuospatial) and developing *Situation Awareness* are expected to be developed in future steps. (University of Toronto, 2019)

Lee et al. (2011) conducted a study where they assessed how participants ranked items. They measured the *accuracy* of the responses and checked how these correlated with their *self-reports*. A self-report of the level of expertise prior to task ordering was presented to the participants as well as a self-confidence assessment after the task completion.

A publication by Wickens et al. (2019) focused on pilots and the way they develop expertise. Different skills such as technical and non-technical show differences among novices and experts. Proficiency can be judged through *decision outcomes*, *objective quality of performance*, and *instructor ratings*. Further, experts show *consistency* in producing the best outcomes. *Situation awareness, automaticity, knowledge* and *expertise* in general flight performance are pointed out as other measures of expertise. Outcome quality relates to *experience, certification,* natural cognitive or *psychomotor abilities,* and *cognitive strategies*.

Endsley (2018) explains that expertise also has an impact on SA. She defines SA as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” (with reference to Endsley, 1988, p.97) and is divided into 3 levels:

Level 1 SA – Perception: identifying relevant information from the environment. This step might be challenging especially for novices to see which information is important.
Level 2 SA – Comprehension: This level incorporates an understanding of provided information.
Level 3 SA – Projection: Predicting future events helps to make decisions at the right time. This step marks the highest level of understanding of the event.

Endsley highlights that SA varies among novices and experts. Figure 5 shows that SA is more demanding, often *incomplete* and includes many mistakes in novices. Novices’ *attention* and *working memory* are described as *limited*. Experts however have a *better understanding, perform with less effort* and *more completely,* are able to *project in the future,* and *show fast SA*. They have created *mental models* within their domain, achieved *automaticity* of processes, learned skills such as *scanning patterns* or *communication*, and have a *schema* of prototypical situations.

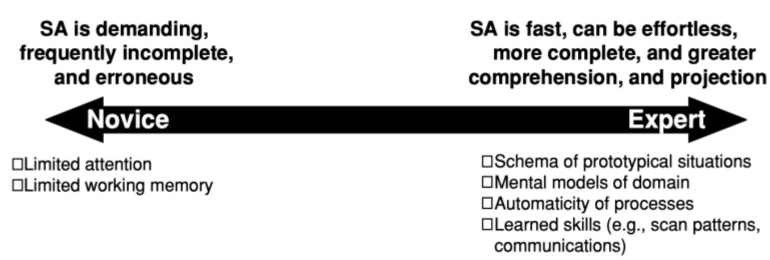


Figure 5: Factors affecting SA in novices and experts in a domain (Endsley, 2018)

According to Ericsson (2015), it is a challenge to measure objectively *reproducible superior performance* which is an indicator of expert knowledge but can be measured in everyday situations such as patient treatment outcomes. The gradual development starts with improving initial skills and requires timely, accurate, and informative *feedback* which can be either received *from others* or *self-generated* as well as learning opportunities that allow *repetition*. *Mental representations* can be used to learn superior performance which should be monitored and facilitated. (Ericsson, 2015)

The difference between mental models and SA is outlined by Andrews et al. (2023) with reference to Endsley (2016) and the National Academies of Sciences Engineering and Medicine (2022): "The key difference is that mental models are something that exists in long-term memory, regardless of present context, whereas situation awareness arises within the context of a specific situation and results from the *application* of mental models to stimuli and working memory. Mental models aid situation awareness by providing expectations about the system, which guides attention, provides 'default' information, and allows comprehension and prediction to be done 'without straining working memory's capabilities'".

The approach outlined in the methods section shows one possible way of investigating expertise in critical care directly with physicians working regularly in such units.

3.3 Methods

To get an in-depth understanding of how physicians in critical care think about expertise as well as expertise development, semi-structured interviews have been selected as a suitable, qualitative approach to get more detailed insights. The results were thematically analyzed through an inductive approach as described in Chapter 2.

This section covers another part of the interview that took 30 minutes out of a full session of 60 minutes per participant (critical care physicians). Again, the same qualitative approach as depicted in Chapter 2 has been undergone, i.e., a thematic analysis as in Braun & Clarke (2012).

The study received ethics clearance from the University of Waterloo Research Ethics Board (#42892), and the materials used in this study can be found in the Appendix.

3.3.1 Participant recruitment and data collection

The participant recruitment process has been described in the previous chapter. The interview protocol included questions around expertise posed to both staff and trainee physicians in critical care:

- Characteristics of experts in the field
- Differences between experts and novices in the field
- Progress indicators that show expertise development
- The development of mental models as an indicator and characteristic of expertise

3.3.2 Data analysis

Data was collected firstly through the recordings and transcriptions. After participants' final (voluntary written) feedback or comments on the summaries, the data was transferred to an Excel sheet for thematic coding (inductively) as described in the previous chapter.

3.4 Results

This section includes the same participants as in Chapter 2. Demographic information can thus be viewed in Table 1 and Table 2. Abbreviations for the participants are used as follows:
While “T” stands for the trainee associated with a numeric value, the “S” stands for staff physicians.

3.4.1 Characteristics of experts

One of the first questions participants were asked tackled the perception of expert characteristics in the field of critical care. The main themes coming up in the interview with the staff and trainee physicians evolved around: knowledge, training, experience, skills, communication, performance, and personal traits. Specific comments and descriptions on these themes are described below.

Knowledge – theme:
When staff referred to knowledge, they often mentioned that knowing about literature and research areas are not the only elements of expertise but also reflecting and critically thinking about the limitations of what one does not know: “I think that one characteristic of a true expert is that they know what they don’t know, and they recognize that knowledge is about knowledge discovery and is not static” (S2). Also “[…] I ask […] for a second opinion, [..] I’ve got this difficult case, would you mind fighting a second opinion or can I advance an idea [...] for validation” (S3) highlights the awareness about validating decisions with multiple physicians who are experienced.

Similarly, trainees agreed that knowing the literature (T1, T14, T16), basics of physiology, new evidence and academic work (T3), having content expertise such as diagnosis, management, new evidence and guidelines (T4) or so called technical, practical and theoretical knowledge (T7) are key indicators for expertise. “[…] They need the newer trends and new developments that are happening in the field. They would also get involved in education of the trainees or someone involved in leading …] or pioneering something” (T9). Similar aspects have been mentioned by other trainees too, relating to research (T17, T13, T11), however “[…] the number of publications doesn’t necessarily correlate with clinical experience. […], but how do they convey the message” (T13).

Training – theme:
Some staff referred to the importance of completing trainings or certification programs, “[…] the College of Physicians and Surgeons would say that anyone is expert once you achieve that level of certification” (S3) but would not be the single aspect “[…] Having a label as a cardiologist for critical care, that’s not necessarily enough to be an expert” (S4).

Trainees also mentioned specialized and structured trainings such as the ECLS training (T6), certifications (T12) or completion of trainings in general (T14, T18) as further characteristics.

When trainees were probed to reflect about what would make them an expert or the actions required to take to become an expert, one trainee added (T4): “Cognitive expertise is more nebulous, it involves continuous education, studying, reading the evidence, and staying on top of that. You start to assimilate information and come up with a treatment or diagnosis or management plan.”

Experience – theme:
Many staff mentioned that the years of practice, the type and volume of patients is one of the factors contributing to the expertise of an ICU physician (S3, S4, S6, S8, S9, S12, S13, S14, S15, S16, S17); "[…] it's a mix of knowledge and practice in volume for that perspective you need to have done many years of practice, seen a lot of patients and just reading is not enough for that perspective" (S9). Some added that the complexity (S8, S14) of cases is an adding factor.

When asking staff physicians when they would know they are an expert in the field, one staff physician mentioned being 3-4 years into practice: "Well, it's a good question. Do I even know that yet? I would say probably three or four years into practice" (S1), whereas another staff physician said 5 years after the completion of the training (S12).

Trainees mentioned the same aspects for experience (T3, T5, T6, T7, T9, T10, T12, T14, T15, T17, T18), "[…] I would say these two things: how many years of experience with knowledge and how well one researched (T13).

Skills – theme:
Having an analytical approach is another key aspect of expertise; thinking through decisions and problems: "[…] but also in a very analytical [way] I would say, also people who are patient; it requires a lot of time to think through problems, deal with families and deal with different levels of learners" (S1), being able to quickly recall and having a sense for knowing who is sick and what might happen next (S4), monitoring patients accordingly and diagnosing them accurately (S16) are further key elements. Thereby, having a "unique approach to a problem" (S18) and evaluating the information in context and portraying skills of "lateral thinking, that shows the understanding of nuances of different disciplines and not being too focused or having objectivity in the trajectory and decision making" (S14).
Pattern recognition has been mentioned by various staff physicians (e.g., S3, S7, S12): "OK, there's the knowledge area, but there's also this kind of tacit, pattern recognition that goes on, you have seen something before. You know how the trajectory is going to proceed" (S5). Acting "in a timely manner, managing multiple dimensions of the patient, the unit organization, as well as having an understanding of the dynamics of ethics and being able to operate within the dimensions of social, cultural and ethical challenges" (S7) has rounded up the skills section for staff physicians.

Trainees added certain procedures as relevant skills such as dealing with the transcranial Doppler (T2), knowing how to prioritize: "[…] But I'm learning how to prioritize things, learn how to effectively allocate resources" (T6), manage time, being able what to skip (T8), having procedural expertise relating to the cognitive and technical aspects as well as being fast with various invasive procedures: "[…] it's not only cognitive, but it's also technical. And so, someone who has a very fast style with various invasive procedures that we do. So, I would say, cognitive and procedural expertise is kind of it." (T4), and having the skill to teach (T11, T9), and receive feedback on teaching (T12). "Teasing out similar presentations and dealing with atypical situations" (T10) relates to the pattern recognition skill mentioned by staff physicians as well.

When probing trainees about when they would know that they have become an expert or what actions they need to take to become an expert, they added a few more points:

The “reflection on new cases one hasn’t seen before” (T3), as well as “thinking without trying to ask for help” (T2) are further skills. “Knowing when and how to prioritize and manage” (T8), managing multiple patients at the same time (T15), dealing with diverse patients (T13), facing difficult patients over and over again (T12) and “having the ability to treat a lot of high acuity situations” (1) are other factors that strengthen these skills. Some speak of a mentality: “[…] you develop a different mentality when you deal with a different subset of patients and you need to spend some time with these particular patients in order to develop this mentality that is required for somebody to be or perceived, or to perceive his or herself as an expert” (T6). It requires certain time and dedication that would yield to acquiring the ability to take the “decisions for the best care of the patient in that moment” (T7). This mentality also develops when “working in different countries, hospitals, academic centers or different departments which will make you mature faster” (T3). Taking part in rotations e.g. nurses would add a different perspective to learning (T3).

Communication – theme:

“Communication is hugely important” (S4)- it has been depicted as part of a skill that relates to the ability how to interact with other stakeholders such as the team, different learners, or the family (S1, S7, S10, S16). Some also added that being recognized in a field is also a factor of being an expert - “being consulted for questions or a second opinion for validation” (S3, S5) as well as networking for sharing knowledge (S6) or presenting at conferences (S11).

“[…] to communicate with both the family, the patients and team members, I think that's something which can only come with like with experience and with time, it's very difficult to learn it without actually doing it.” (T9)

Trainees also mentioned that a network is important in a specialized area (T7), as well as the ability to convey the message (T13) to various stakeholders (T16).

Performance and personal traits – theme:

Having endurance and not overreacting has been described as a characteristic of an expert, as well as having heuristics (S14): “If you're very focused, too overly focused in one area and can’t think laterally, you don't do well in critical care. We use some heuristics. […] but you have to be very objective, I think is another component. Endurance is another component, I think, because patients are followed repeatedly over time, and you have to understand the trajectories and make decisions. I think those that are not good at it, they tend to react too much to individual data points. Or overreact to aberrancies in data points like physiologic mostly.”

Being patient, thoughtful, and contemplative (S1) as well as an understanding of families (S7) have been added further. Furthermore, being a leader in the field is another aspect of expertise (S9). Also “developing approaches that are effective in managing patients” (S12) provided another dimension that could be related to performance.

Trainees commented that working every day and developing a culture (T5) as well as assessing the treatment outcome (T12) can define expertise. Further aspects related to personal traits; “[…] being humble, being approachable, being open to being wrong. And learn from everybody in the team” (T3). Further characteristics described evolved around being comfortable with what you are doing (T1), confident with own decisions (T1), being trustworthy (T14) as well as being a role model (T17).

When trainees were probed to reflect on what would make them an expert or the actions required to become an expert, they added that "talking to experts in the field and taking insights from them", as well as other clinicians (e.g. nurses) or the leadership and management in the hospital can be a great learning point: "[…] if you want to be an expert and the leader in the field, you talk to leadership and management in the hospital. You know more of the dynamics and the politics around your surface, and that's really important for a leader because it can facilitate or it can be an unopposed obstacle for the patient because there's a significant amount of what we do is not pure medical, so you need to know what's going around with that" (T3). Also, there is the legal component to being ready to "back up the decisions you made" (T14)

Some trainees also mentioned that expertise develops when "specific characteristics cannot be fulfilled elsewhere" (T7) and reflected by saying "I don't know if I'll ever get there" (T9) or "I don't know where the cut off is, I don't know what makes you an expert" (T12).

3.4.2 Differences between novices and experts

To get a clearer picture, both trainees and staff physicians have been asked to share their insights on the perceived differences between novices and experts in critical care.

The themes were similar to the previous section when they described the characteristics of experts but highlighted aspects that would clearly differentiate the expert from the non-expert. An additional theme that the staff physicians highlighted was the aspect of "working where there is no evidence".

Knowledge – theme:

Staff physicians commented that experts have more knowledge (S3, S7, S6), decipher and critically appraise medical literature: "being able to decipher and critically appraise the medical literature on the topic, you know, being able to understand: what are the limitations? When can it be applied? Or more importantly when it cannot be applied or kind of like working on the fringes of the data and really being able to find information outside of established clinical trial data that you can then still use and apply in your clinical practice" (S1). Being able to better come up with trajectory (S5) has been mentioned as another differentiating factor.

Further, (S7) mentioned that "being aware and function in all domains at once in parallel and do all at the same time" is a major difference to trainees who do that sequentially and not concurrently.

Another difference besides concurrent management is the detail to understand and treat the patient:

"Experts have more nuanced knowledge and better ability to integrate, synthesize knowledge with patients of multisystem issues that have conflicting needs. Trainees instead often don't have the knowledge or intuition to adjust the 20 normal steps to fit with the specifics of the patient, but an expert has this" (S16). This also relates to neurocritical care where multimodal monitoring, in-depth knowledge and nuanced knowledge are incorporated into the experts' ability (S16).

Trainees mentioned that "experts know the exact steps and can tell the steps" (T8) and that experts have a vast knowledge base when compared to novices who don't (T5). This knowledge base is related to keeping up with what's going on ", referring to research and evidence existing in the field (T9). Further, (T2) described that "experts know the best care for patients; they know what to do also with the current evidence as they are on top of new publications and also have basic knowledge of certain tools like imaging or ultrasound".

Training – theme:
Trainees commented mainly on the aspect of training, saying there may be experts in certain aspects of critical care (T1, T13), "you can't be an expert in everything" (T13). Another difference between novices and experts was described as "if they are learning or already are an expert in a certain thing" (T12). The years of training (T15) have been pointed out as well.
There are also ways of how novices try to learn more: peer-assisted learning or coaching as well as helping each other as trainees is important (T3).

Experience – theme:
One of the aspects that was mostly mentioned differentiating experts from novices was experience. Experts would many years of experience (T5, T6, T9, T15, T17, T18), while trainees lack experience (T3).
Experience is not only related to time though, "the time and place of practice influence the knowledge and skills" (T7), and applying the best evidence: "what you're doing is the current best evidence for this patient" (T14), and expertise comes with hard work and time, the volume and variety of patients portray another dimension (T14). "People who have the worst calls, actually develop expertise quicker because they are put under stress, and they see more" (T14).

Staff physicians also talked about the number of cases that play a role (S17, S3), as well as practice and experience (S6, S7, S10), and cohesive teamwork (S10). "Experts in training accumulate exposure to the specialty and fill the gaps" (S5).

Trainees often focus on instantaneous values whereas experts care more about trends and waveforms (S3): "I think that the challenge for the trainees is that they just don't have that experience. However, what they do have is often a much better knowledge, a recall of the guidelines, the data, and their recent evidence. And so, where that becomes a real challenge with just going back to the visualization part is that when you're looking at instantaneous data without that trend, the purely dated guideline driven like the training environment is like the value is, [for example] the value is 18, it's not high. Therefore, I don't have to do anything right. It's like a more binary approach whereas the more experienced minimize well what's the trend? I don't like where the trend is going. Let's start either preparing or let's start doing, you know kind of a soft intervention or let's do something ...And without that, without better visualization, you know it's hard to both explain and train or educate the trainees to say, like what I'm thinking about is this. This is why I'm doing it versus just devolving into that binary data-driven approach. I don't make it sound like, sorry I'm using more data or maybe an evidentiary approach like that, you know the more senior staff aren't data-driven."
Being able to individualize patient care through knowledge, experience, and expertise (S13) has further been mentioned.

Skills – theme:
Many trainees commented on certain skills that would differentiate the expert from the novice in critical care. One component of that refers to teaching: experts are often teachers and can explain the related topics to the trainees (T2, T5, T14). They also identify mistakes (T5): "[…] most of the physicians or the attending physician that I met know how to teach and know what our mistake is because they see that many years; if we say this is what we thought that [is what] we are collecting. But usually, we forget to see something, so they have that there."

Often experts are mentors to trainees (T9, T14). They communicate with the team, prioritize, and convey the main concepts to fellows (T8, T16).
Trainees, however, learn general critical care procedures (e.g., managing the ventilator) and seek advice from supervisors (experts) (T11). “More advanced trainees don't require advice from supervisors anymore” (T11). Also in uncertainty, experts are the contact person:
"I will turn to an expert opinion if there's a major conflict of what I should do and what the patients or their family want to do" (T16).
Procedures and managing patients with certain medical conditions that warrant needing to be in the ICU are further skills that experts portray (T1). They have a better level of understanding: “[…] The next level lapses like someone understanding what this patient's trajectory or prognosis is likely to be based on their medical conditions and their predisposing factors. And then the next level up might be someone who is able to think organizationally like you know, where is this patient best managed? Or should we be having conversations about aggressive resuscitation versus, you know, palliation like making those higher-level decisions that are not necessarily just treating the blood pressure?” (T4).
Experts have the ability to perform certain procedures: e.g., the echocardiogram and then translate the findings into the care of the patients: “and then it is about how one will modify the clinical practice based on what one finds” (T12).

The trainee needs to also manage different degrees of shock: “[…] they might be confused, and we know we have 100 things to do, we might be shocked at that time too” (T8), call other experts or surgeons to help with source control, or consult with others that provide a certain skill set (T1) they don’t yet have.
Further, trainees make unilateral decisions without talking to others, while experts show individualized patient care (T3).
Someone who is not an expert thus struggles with the algorithm in their mind and can't think outside the box as effectively as experts. (T10)

The staff physicians’ perspectives correlate with the trainee’s perceptions. Also, staff physicians mentioned that critical thinking on procedures and gaining perspective (S2) is a characteristic of an expert showing recognition of deterioration and knowing appropriate actions (S4), thus showing autonomy (S2): ‘I'm overall responsible for the decisions regarding care, but I will say that it is my usual attempt to allow significant autonomy to fellows and residents, students, much less so with regard to the decision making”.

Pattern recognition (S4, S5, S7, S12, S17), filtering information, and being able to prioritize (S13), as well as the ability to collect objective measurements (S15) are skills that an expert has compared to non-experts.

Further, their (experts’) actions are strategic and tell trainees what and what not to learn. They have better clinical judgement and acumen, as well as the ability for problem-solving and troubleshooting (e.g., in procedures). Experts like a lot of information, whereas trainees don’t like it and might be overwhelmed. “Residents are equipped with what they need to survive the night-call” (S16).

Experts overall function in a team and a high-stake environment with cultural, human, ethical, and social aspects of care (S7).

Performance – theme:
"[…] one needs time to develop a mentality (a métier if you will) to be able to act effectively in a specific area" (T6). Effectiveness with management has been portrayed as another component differentiating the expert from the novice; the experts show more effective management (T10, S12).

Personal traits – theme:
Experts are perceived as showing confidence (T3, T5, T12, T14), one trainee also mentioned they have more doubts and are reflective about their actions (T10). An expert is not always a natural consequence of daily practice; it is "to actually accept that you can always learn more and be in peace with that" (T17).
While trainees develop expertise, they become more comfortable (T11) with their decisions and look forward to becoming better (T16) but might also be excited (T8).
"Trainees also have fears and insecurity, they don't like to be vulnerable, we convey the message that we are in control but that's not always something we can achieve. The tendency to be in control with the insecurity of being a trainee and the pressure that you are a fellow and a senior trainee and you are strong, smart, updated, you read a lot of the work, you have seen other patients. That mixture can push you in an area that might not be ideal. This is part of the learning experience. It's something that has to be in the conscious mind, so we train and take a long time on this. Most of it is psychological safety and how we play around to reach that stability and being in a wise position that makes you an expert" (T3).

"Expertise is multifactorial and multifaceted" (S13), it is described as "functional mindfulness" (S7) and showing traits such as being calm, working well under stress, and be objective (S14). Comfort has also been mentioned by staff physicians as an indicator for expertise (S12, S17), in addition to reputation (S12).

Work where there is no evidence – theme:
"Experts try to advance science through research, while non-experts haven't reached that level of advancing science" (T10).
This has similarly been mentioned in multiple ways by the staff physicians: "work where there's no evidence, apply imperfect research results, being able to understand limitations, when it can or cannot be applied, working on the fringes of, and being able to find information outside of established clinical trial data one can use and apply" (S1).

Trainees, however, recall guidelines (S3, S8), data and recent evidence (S3) but they also struggle to deal with situations that are not presented as guidelines (S8). In contrast, experts show out-of-the-box thinking while trainees need to be willing to take the knowledge from the textbooks and learn to individualize it (S13): "I think individual components can be learned in small areas of time and you can be an expert in small aspects. But the time is what allows you to be able to take your knowledge of that area. You're an expert in and actually look at it broadly and apply it to the whole patient, not just apply it to that specific problem, but begin to apply it outside and think outside the box. And that's where you get into. If you've been doing this for years, you begin to recognize that not every patient reads the same book, and they don't always behave in the same way. So, you have to be willing to take your knowledge and apply it."
Experts further educate on the newest developments, speak and write about it (S11), they deepen "to answer questions that are out there and make those research questions a reality" (S18).

3.4.3 Progress indicators: trainees and staff

Both trainees and staff physicians have commented on indicators that show progress when it comes to expertise development. The main themes identified are knowledge in the trained area, developing mental models, performance, showing autonomy, reflecting on actions, and communicating with various stakeholders.

Knowledge in the trained area – theme:

Trainees and staff physicians commented on different sub-themes. Trainees mentioned that training through exams is an objective way to prove progress (T2, T18).

Also, the transfer of knowledge has been pointed out as a sub-theme; “knowing what the next step is, i.e., transferring knowledge from the textbook to clinical practice” (T8) and “being comfortable applying the knowledge correctly” (T11) have been identified as further progress indicator.

Staff physicians similarly commented on this sub-theme: “if I taught one thing, I check whether they remember the lesson and apply it to the next patient (not on the same day, maybe in a couple of weeks). I try to understand if they're receptive, check their engagement and ability to read on their own” (S1). Improving knowledge and pointing it out as well as discussing this if trainees make a mistake (S6), acquiring knowledge around management of certain patient populations (S12, S16, S17) over time to evaluate the quality of their plans (S17) are further comments. (S15) mentioned that constant monitoring at the bedside and asking them questions, reinforcing, teaching every day and keeping a repetitive manner as well as including iterations, he would test trainees’ knowledge (standards for management which they have to achieve and will be examined).

Staff physicians also mentioned that they let the trainees suggest their approach first; so, they would see over time how they can move independently and think in a sophisticated way (S2) or assess how they present the case and observe them (S9).

Trainees have pointed out that they would know they have progressed when they know what others might not know, or when they “ask for my opinion” (T12); “when I was diagnosing things that sometimes even the cardiologists were missing, that's when I realize maybe I've progressed” (T12), or “there is always a person that has something that can help you, I also become that person to someone else" (T7).

Mental models - theme:

Both groups commented on mental models, whereas staff physicians identified granular components of a framework and curiosity and psychic energy or joy, as sub-themes.

Few comments resonated with a framework approach with the trainees: “the way of thinking and analysis of the problem, the capacity to understand all aspects of the patient (T13), identifying priorities, and when you don't think about it (i.e., acting intuitively), when it comes naturally (T14).

The staff physicians mentioned various aspects about a framework: the “consideration of multiple aspects for differential diagnosis” (S2), the “explanation of thought process, checking if they (trainees) apply the same principles to a different, more challenging or atypical scenario (formative and submitted feedback” (S3), whether they develop pattern recognition, and comparing how they made decisions at the beginning and now (S9). Also, “if it's a procedure about 5 times, the focus is if they can break it down to the concrete many steps of doing a procedure and then tell, when something happens, what went wrong" (S11). A think and talk aloud are ways of assessing this (S11).

Similarly, (S12) mentioned that trainees should take "all of the information that is available to synthesize it quickly, correctly and appropriately to identify the key issues particularly the ones that are life threatening problems that need to be addressed. Prioritizing or developing a hierarchy for the problems that need to be assessed, and not to get lost in all the information and granular details, also don't dismiss important information. Coming up with a safe management plan that can be implemented considering the urgency of the situation in a timely fashion. The correct identification of problems, presenting greatest risks to the patient, focus on the key issues" are further aspects observed for expertise development.

Some staff physicians also mentioned they recognize progress "If they start to think along the same lines as I do, then they're getting it. But that's not always a good indicator." "I'm looking for certain key conclusions." Looking for algorithmic thinking/decision trees, probes them "why they're not leading to the same logical endpoint that I am." (S5)
Whether they can Independently mimic what not I do but is coherent that includes the known unknowns. Being practically self-aware here is my model but one piece doesn't make sense, now we can have a conversation why it doesn't make sense. If you don't even know what doesn't make sense, that is a problem." (S7)
Also having 3 different approaches on how to solve a problem (S13) has been mentioned.

Having confidence that trainees understand the process which means "that I've spent time with them, and I check, I always validate what they're telling me because sometimes they're not lying intentionally, but they feed me with a load of garbage and then, they often believe whatever is written" (S14). (S14) further explained that he "checks if trainees actually examine the patient; I go to the bedside and walk through this. I watch trainees in how they go through the assessment at the bedside (I step in if they have trouble). It's guiding them through how I would do things. I check what to focus on, e.g., the heart: I ask how to assess hemodynamics, patient examination, ultrasound use, what happened in the preceding period for therapy intervention (understanding the context), understanding the physiology of the heart and disease, come to conclusion about what's going on. I interrogate the trainees that way, often it is about multiple organs", so it is about taking each organ analysis separately, then combining this and asking which interactions between organs is causing what they see on the monitor. (S14)

Providing level of initiative and being ready to provide a rationale in decision making and the nuances (not only guidelines) and individualized care, articulate specific differential diagnosis or management. (S8)

As another sub-theme, curiosity, psychic energy and joy has been pointed out as a contributor to the importance of mental models: "Promote curiosity and provide psychic energy and joy to develop mental models; whether I can see in them a type of joy feelings in what they're doing, when people are doing well, they are enjoying it. Emotions are intertwined with mental models. Feelings are for experiencing and then for thinking about. Thinking is for processing and then doing" (S7).

Performance – theme:
Trainees and staff physicians talked about three main sub-themes of progress indicators as part of performance: progression of a patient state, efficiency and feedback.
Trainees commented that they could tell they progressed when their treatment outcome, i.e., the patient's state, would improve and doesn't worsen, although this might be hard to tell in certain situations (T9, T17).

This has similarly been pointed out by (S16) looking for scholarship and professional attributes, "can they formulate pertinent learning objectives, research and integrate it into their framework".

Efficiency (T16) has been mentioned in various ways, "being able to better and promptly respond" (T8), as well as "speed, conciseness of your notes, and getting straight to the point" (T14). Being concise and precise during handover has been identified as another component (T6).

Multiple trainees considered feedback as another progress indicator. Feedback could be from staff through continuous evaluations and on a daily basis (T4), by asking friends or the team of trust for feedback (T7, T9, T10) to further improve, or the head of the unit who would provide feedback on the knowledge base, skill level, and communication (T18). Also, the acceptance of a decision by the nurse (T2) would be such an indicator.

One staff physician rounded up these comments by explaining there would be various forms (observed, written, verbal) of feedback provided to the trainee on the assessment of their progression, "oftentimes, the doctors are not the best source of evaluation and so we rely upon nursing staff as well as respiratory therapists, pharmacists and in some cases even families to provide feedback for trainees" (S10).

Autonomy – theme:
The sub-themes identified from the trainees and staff physicians' comments were related to independence and confidence.

Becoming more independent (T15) and proficient in asking fewer questions especially at night (T2) have been pointed out.
Staff physicians similarly commented that "knowledge and desire to do it themselves with support from the faculty" (S2), being independent and autonomous in having the discussion (S18) and if they ask less questions and make decisions independently (S6).

Confidence and levels of confidence were noticed (T2, T6, T14), and "how well you sleep at night when you go home, how much you doubt yourself. The experts doubt themselves as well and have many sleepless nights" (T14).
Comfort with a specific task (T16) and having feasible small steps that make you more comfortable with your practice effecting safer and better actions for the patient and their families (T17) has also been identified. One trainee (T15) further added that micromanagement led to the loss of confidence as working in different countries shows varieties in practices: "The program that I joined was a I a little bit like they were micromanaging everything because it was like a VIP hospital. So, they have to be involved in everything. So, I kind of lost confidence during that year."

Reflections – theme:
Trainees mentioned self-reflection and comparison with others (T6) is another progress indicator for them. Self-reflection is a component that includes the awareness of inner biases (T10) and when progressing, reviewing one's own actions after applying knowledge to a new case may not be required anymore (T11). Knowing the unknown has been identified as another way of self-reflection "a good indicator that I'm getting better is when I know, I do not know anything at all, I do not know everything" (T7).

Communication – theme:
Two sub-themes have been identified: teamwork and communication with the patient.
As part of teamwork, there are discussions with peers on how else to solve a problem (T7) or generally the way of interaction with the team (S4). Communication relates to various aspects, i.e., verbal, organized case presentation, handover pertinent/organized manner), collaborative skills (identify what services are required, the ability able to ask pertinent questions/well-worded), bedside manner (family communication, disclosure, bad news) (S16).

The interaction with the patient to the accuracy of the situation (S9) is another key indicator for progress: "[…] the interaction with the patient to the accuracy situation that happened during the night often is a mess […]. When I come in the morning, I have a sense of 80% of the things they're going to be made decision during the day. Unless there are things that are unexpectedly happening now and you're expecting now sort of a decision from the fellow regarding a certain thing. You see them when they are at the beginning, they have certain decisions. And then as they progress you see that they develop this pattern recognition."

3.4.4 Mental models

As stated in literature and in the interview, mental models are one way of differentiating a novice from an expert. Thus, trainees were asked to share how staff physicians convey their strategies and mental models to them and how they develop their own mental models in critical care. Similarly, the perspective of staff physicians has been gained to provide details on how they are conveying their approaches and mental model to the trainees and support trainees to shape their own mental models.

The main themes trainees and staff physicians mentioned were related to:
Knowledge and literature, teaching, reflection and corrections, autonomy, and tools that would help with mental model development.

Knowledge and literature – theme:
Both groups mentioned that the incorporation of the evidence and guidelines are an important factor showing and developing mental models in critical care. Being aware of literature and recent publications are components of being up to date with the evidence and guidelines. Trainees have also mentioned that they follow a trial-and-error approach (T9): "We do take a little bit of what's good from each and what we like from each one to try and build our own way of doing things. […] There's not much evidence in terms of whether it's good or it's bad or it's better or worse. And also, certain things work within the hands of certain people, but they don't work the same with somebody else. It's hard sometimes when we're still learning what to figure out, what works, when we do it."
It was further mentioned that often exceptions are usually represented in the hospitals and not much in publications.

Methods of teaching – theme:
The theme teaching is divided into sub-themes, i.e., Socratic teaching, academic teaching, talk or think aloud at the bedside and replicating staff's actions, as well as structuring and breaking down the complexity. In terms of Socratic teaching, staff physicians often mentioned they would let trainees suggest their approach first, ask questions and guide them (S11): "I use the Socratic method, oftentimes with a sense of humor for the earliest trainees. And then I make it hopefully an easy learning environment, so if there's any

hesitation response, I upscale to the next level of knowledge in training age and then get to the point if nobody knows, then I just launched into teaching. But this I find if you do it with some humor, it allows people to teach what they know and actually gain visibility in the team and trust and all of those things."
Trainees responded similarly, pointing out that working with certain physicians also helps them to better understand their mental model: "the longer you work with attendings the more they know you and give you hints that you otherwise need a lot of time to get there (individualized). The more you know the attending physician, the better you understand what they prefer (e.g., some love sedation, some not and intensify the mechanical ventilation)". (T6)

There would also be academic teaching sessions such as seminars or chop talks to learn from one another.

As teaching and learning is multi-faceted, one component is to follow the talk and think aloud process of the staff. Staff physicians try to use various ways to show their thinking process, e.g., through the use of "physiological points that help to make a decision" (S8). Staff physicians also modify some aspects to challenge trainees to think critically about the cases (S13). Although the process is considered bidirectional (T10), staff share their heuristics (S3) so that trainees develop a better understanding for their recommended steps.
Trainees shape their own mental models through showing their approach to the staff physicians, follow their steps or take parts of each staff physician's approach and apply them. Trainees also seek advice by asking the staff questions.
Structuring and breaking down the complexity of the cases and actions is one major sub-theme brought up by both trainees and staff physicians, whereas staff physicians highlighted many details in this sub-theme. Although trainees recognized that staff convey their mental models by going through a simple and broad step first and then go into depth (T11), staff physicians explained they have a systematic approach (S9), "trainees tend to jump to conclusions and are often heuristic and have a problem synthesizing the knowledge sometimes (S14). My method is to parse things down", so he starts with an individual organ and then merges each organ system together as a full system. Checking the underlying physiology of each system, it almost goes to a cellular level. "Juniors are very good at templating things but not with thinking laterally and abstractly" (S14). Another staff physician highlighted that it's important to see if the trainee is "assessing the situation, understanding and being able to take it to the next level, that one can actually solve a problem, not just repetition of memorization." (S13)
Working with heuristics solely isn't the best way to deal with things: "[…] things happen in medicine and critical care, and often they're in the middle of the night. And if you're just trying to heuristically deal with things, you can do that maybe 70% of the time, but you'll fail if you don't think through the nuances because other factors are happening at the same time. […] So, it's not only a single system, but then integrating the other systems and how they interact, but I tend to dissect it system by system first and then and then consider the interactions between systems. […] I don't react to aberrancies unless I'm really concerned." (S14)
Another staff physician similarly commented about a framework approach outlining the pillars and adjusting the information: "You have 3 things happening at the same time (it's not hierarchical) and this is why the ICU is so stressful: First, stabilize and resuscitate with minimal knowledge, second collect all information, third addressing underlying problem. Then work on specific knowledge points they need to know about the disorder." (S16)

When trainees commented on how they would acquire their own mental models, some mentioned they would look at the broad picture before going into detail (T8, T11) whereas only one trainee mentioned he has a systematic approach of examining the patient from head to toe and listing priorities (T14).

Reflections and corrections - theme:
Reflections and corrections were another theme identified in the comments, mostly mentioned by trainees. Staff physicians provide feedback on trainees' approaches so that they can learn and improve. Some staff physicians would also ask the trainees directly about what they would want to learn (S10) and support them finding the right answer to the problem (S3).
Trainees outlined that they ask “focused questions” (T1) and in case of controversy, they would “go with what I think is the right way” (T8) to form their own mental model.

Autonomy – theme:
The theme about showing autonomy came up multiple times by trainees commenting on how mental models are shaped; “During rounds, it's important that the trainee feels that he is in charge, there's the comfort zone in which you will not learn from where the staff will be taking care of everything and do anything; and the panic zone in which you are left alone and you don't have a backup. It's important to be in the learning zone where you have some stress being at the spot and being critical about what you're saying and hearing, but at the same time there is somebody backing me up if I'm unsure or I'm not sure or somebody can turn around and ask them." (T3)
There are different styles of staff physicians, some would “give a lot of autonomy” but balancing autonomy and micromanagement is considered best (T10). Also introducing oneself to the team is an important step to show one is in charge and can communicate and learn from another (T3).
One staff physician mentioned that developing a “level of certainty and comfort with what’s happening and when to ask someone for help” (S4) is another component of showing autonomy.

Tools – theme:
Various tools to show strategies and support mental model development have been identified. Simulations (T1, T7, T10), visualizations, drawings or lectures have been pointed out as other ways to support this process (S3, S9, S18, S5).

3.5 Discussion

There are many parallels between the literature on expertise and the themes identified in the interview study. Although aspects of expertise described in the literature were not only related to healthcare or critical care expertise, there are many characteristics that apply to ICU expertise and its development as well. The interview detailed out not only aspects of what is already provided in previous literature but added more nuances to the description of expertise. Integrating the perspectives of experts as well as novices shows a great richness and depth of this complex topic and pulls a fuller picture. Although participants were asked to share their insights and thoughts on the previously mentioned questions on characteristics, differences of experts and novices etc., we can see that often similar or the same themes show up in those questions. As part of semi-structured interviews, it was beneficial to have certain flexibility with the questions as some participants shared their thoughts more on characteristics of experts while others commented similar things when they reflected on what they perceive as a difference between novices and experts. This was intentional to identify the nuances and get more insights from the participants. Although one expects that experts may reflect more abstractly than novices and have more experience and gone through the novice-to-expert development process, it is noteworthy that many novices were already aware about a lot of aspects on expertise development and what may be expected of them in the future to become an expert.

3.5.1 Reflections on the study

The interview study shows the depth of expertise perceived by the stakeholders within critical care themselves; as many aspects related to skills, knowledge and performance, study participants also pointed out several aspects about the cognitive side such as developing a certain mentality or mental model in critical care that challenges them. Gaining autonomy and the related confidence in providing a rationale for the decisions made provide a further important facet of expertise.

Reflecting on this study, it is noticeable that expertise as a topic covers a vast range and the limited time of 30 minutes for this can be viewed as a constraint. The demographics show that trainees have various amount of exposure (e.g., year of fellowship or previous experience) and backgrounds, e.g., in emergency medicine and thus might have varying perceptions on progress indicators for example as there might have been a subjective component when they talk about what they want to develop further to become an expert in the field. Literature states that emergency medicine physicians often are “experts at resuscitating critically ill patients” (Johnson et al., 2017 with cross-reference to Somand et al., 2005) which also has an impact on how these trainees would describe expertise coming from various backgrounds and different skill sets.

Further, the staff physicians identified in this study as “experts” are all part of academic hospitals. This does not mean there would not be experts in non-academic centers or staff physicians with non-academic interests, but they might have a different focus in their clinical practice, training the trainees in their field and covering the gaps of research from an academic perspective. In future studies, staff physicians with non-academic interests could also be considered to deepen research around expertise.

3.6 Conclusion

Overall, there are various themes that relate to expert characteristics, differences among experts and novices, the progress indicators, and mental model development. More literature is needed in all of these specific aspects that were mentioned by critical care physicians, especially to understand the complex mental model development acquired in this field.

This study helped to gain a greater perspective on expertise and crystallized some details that could be transferred to the technological aspect of the overall dissertation topic. Not all aspects of expertise can be supported through technology, but there are various parts that could be translated into e.g., grasping a wider range of available information in context.
To summarize the answers to the research questions asked at the beginning of this chapter relating to experts and novices, their skills, and progress indicators, more will be explained in the next sub-sections.

3.6.1 Expertise in critical care

"Expertise is multifactorial and multifaceted" (S13), this is a statement that describes the extent of this term. Expertise cannot be described within a couple of words, it is a phenomenon that captures characteristics and a certain process of acquiring, transferring, and projecting learnings in a dedicated field of interest.
The main themes that came up among both staff and trainee physicians related to certain characteristics that describe an expert in the field, for example, the vast knowledge and knowledge discovery through training and experience (years of practice, volume of patients), the various types of skills they possess (evaluating the information in context, understanding nuances, prioritizing, considering feedback, caring for diverse patients, having a certain mentality) including communication (with various stakeholders e.g., the team, different learners, the patient's family and the ability to convey the message) and a certain performance level (using heuristics, being effective). Certain personal traits were also associated with expert characteristics (not overreacting, confidence). However, this does not mean that an expert has all such characteristics on a high level but may have tendencies in certain characteristics that are stronger in one than in others.

When asked about the differences among novices and experts to get a clearer picture, many of the themes were stated similarly. Knowledge is not only related to textbook learning but also about the ability to come up with a trajectory (experts know the exact steps and can tell the steps), the ability to function in all domains concurrently and not sequentially, to integrate as much information as possible but also synthesize it and use multimodal monitoring.
Experience is not only about the number of patients seen over various years but also learning not to focus on instantaneous values (which trainees often do) and understanding to look for a trend and the waveforms, thereby individualizing patient care.
Skills relate not only to the identification of mistakes but also to conveying the main concepts, thinking outside the box effectively, recognizing deterioration, and developing pattern recognition skills. Filtering the information in a meaningful manner and collecting objective measurements have been noticed as differentiating skills among experts and novices, too. On this note, it was stated that experts like to have a lot of information while novices often feel overwhelmed. "You can't be an expert in everything" (T13), and being an expert is not always a natural consequence of daily practice (T17).
Experts have the ability to understand limitations while trainees often recall guidelines.

The development of mental models has been related to knowledge and literature, learning and applying the evidence and dealing with guidelines, learning about recent publications to be up to date with research and new findings, and applying a trial-and-error approach. Trainees participating in this study were recruited from university hospitals and mentioned that they see the exceptions (or exceptional cases) in such hospitals which are not much represented in publications.

Regarding methods of teaching, it was denoted that knowing the attending physician contributes to a more individualized learning approach and understanding the attendings' approaches or preferences for certain care styles. A talk or think-aloud process is often taken to show the "physiological points that help to make a decision" (S8) and convey heuristics. A systematic approach is often followed: "trainees tend to jump to conclusions and are often heuristic and have a problem synthesizing the knowledge sometimes" (S14). "Juniors are very good at templating things but not with thinking laterally and abstractly" (S14).
It has also been pointed out that often, multiple things happen at the same time (it's not hierarchical) and the trainees need to learn to deal with such situations and create a mental model. Trainees mentioned that they shape their own mental models by sharing their approach with the staff and seeking advice. They try to list the priorities while staff physicians expect them to structure and break down the complexity of the case.

Creating a mental model by knowing one has more autonomy but is still in the learning zone is also important. Staff physicians highlighted that trainees should know when to ask for help. To gain and train such a mentality, different tools are used such as simulations.

3.6.2 Progress indicators: trainees and staff

In terms of knowledge in the trained area, trainees mentioned that they notice progress when they start knowing what others don't know, knowing what the next step is, and transferring knowledge. Gaining comfort about correct knowledge application is another progress indicator. Staff however check if trainees can apply standards of management also for iterated cases and use an approach where trainees suggest their management plan first. With this approach, staff also check how trainees present the case.

For recognizing progress on the mental models, it is often checked if trainees use a framework and are curious and have psychic energy or joy in their actions. It is expected that trainees develop pattern recognition skills, deal with atypical scenarios, figure out what the priorities are, and incorporate all aspects of the patient. Synthesizing the information quickly, identifying what went wrong when something happens, and recognizing when there are life-threatening problems are other factors. Caring about the nuances, individualizing the approach and having multiple approaches in mind to solve the problem are other indicators of progress.

The performance can also be taken into consideration for progress: it is often checked if the patient state is getting better (which is not always the best indicator for assessing progress), the efficiency of trainees' actions, dealing with feedback, and having a precise and concise handover plan is also important.

Becoming more independent has been mentioned as part of developing autonomy.
Reflections play a role too, especially self-reflections; "a good indicator that I'm getting better is when I know, I do not know anything at all, I do not know everything" (T7).

How to work and communicate in teams, with the patient, or having a pertinent handover in an organized manner are other progress factors.

3.6.3 Ideas on expertise development measures for the usability study

The insights of participants can be taken to prepare for a usability study that will be explained in the later stages of this project (Chapter 6). Capturing some key information from this section can be used to a) prepare a usability study, b) measure the points that relate to expertise development of participants and c) prepare the visualizations' design for the interface that will be used in the usability study.

The following aspects can be used to design the study:

- Include scenarios that show diverse patient cases.
- Provide atypical scenario(s).
- Portray cases where the participant has to act independently, thus showing autonomy (single participant running the study instead of a team).
- Participants shall relate to standards of management.
- Participants, i.e., trainees, shall suggest their approach first and present the case to the study conductor.
- Set the expectation that the participants won't have expertise in everything but show facets of expertise.
- Include a talk and think-aloud process.
- The study shall provide context and can be similar to a simulation.
- Provide a situation where participants shall think and plan as much as possible but also consider asking for help in case they don't know how to further proceed.

Measures that can be considered to assess expertise development:

- In terms of performance, the number and severity of errors can be taken into consideration, relating to the correct synthesis of information.
- The think-aloud process can help to gain insights into their reasoning; the reasoning is important to check if trainees understand nuances, see the trend, recognize deterioration and risks, or start seeing some patterns and can tell the next steps. Also, aspects of the mental model will be clearer to see if trainees are able to break down the complexity of the case, consider various aspects of the case, and identify if there are multisystem issues and priorities. Situation Awareness can be taken into consideration in this context too, as it displays three levels (perception, understanding and trajectory) of how participants think in a short-term context.
- A case could be offered where participants "know when to ask for help" that would link to autonomy or communication skills.
- Reflections about limitations or correcting own actions can be another measure.
- Confidence could be measured as a personal indicator for expertise development.

3.6.4 Impact on concepts of visualizations to support expertise development in critical care

Technology and the information that is represented through interfaces play a role in supporting the mental model development over time and thus expertise development in critical care. Interfaces are used in various ways and can be viewed by a single person or a team. Often, it is used as an information retrieval tool but can facilitate learning and convey certain patient states. Gaining autonomy in such a complex field can be supported by increasing the understanding and analysis of in-depth patients' states. The bedside physiologic monitor is one of many tools clinicians use daily. It shows a vast amount of data points, variables, and critical points for assessment. Although there are even intensive workshops that tackle multimodal monitoring in e.g., neurocritical care, these monitors are not limited to portraying single types of information but entail various components of information (such as waveform data, and numeric data) from various sources.

Many aspects have been discussed during the study that could be used to develop visualizations to support critical care novices, especially during their learning process.

Table 3 below shows some aspects that show the potential to impact visualization design.

Table 3: Aspects from the interview that may affect design ideas for visualizations on the bedside physiologic monitor

Theme	**Aspects**	**How visualizations can support**
Skills	• Evaluating the information in context • Understanding of nuances, using a lot of information • Pattern recognition • Filtering information • Collect objective measurements • Receive feedback • Identify mistakes • Treat a lot of high-acuity situations • Recognition of deterioration • Convey the main concepts	• Show variables in context and in relation to another • Portray details of data in various ways • Highlight certain data thresholds or include notifications for certain ranges • Indicate visually the areas of increase or decrease to relate to acuity/alert • Design the visualization in relation to main concepts of critical care
Communication	• Interact with other stakeholders such as the team, different learners, or the family • Convey the message, provide basis for an organized and pertinent handover	• Design a user-friendly display that enables customizability for various end-users of the system • Display trends and important data points in a meaningful way to facilitate handover and communication
Performance and personal traits	• Heuristics • Effective management • Improve performance and not get worse • Concise and precise handover	• Propose a design that facilitates the use of common heuristics of experts • Design representation should provide insights on effectiveness of taken actions (e.g., timing of provided treatment) • The interface may include summary visualizations to enhance handover management
Knowledge	• Trajectory • Integrate, synthesize knowledge • Multimodal monitoring	• Provide data over time to facilitate trajectory • Show various variables and data points of multiple measurements/signals
Experience	• Instantaneous values • Trend and waveforms • Individualize patient care	• Show data over time • Provide different ways of trending the data and waveform patterns in context

Theme	Aspects	How visualizations can support
Mental model	• Physiological points that help to make a decision • Nuances • Don’t dismiss important information • “Juniors are very good with templating things but not with thinking laterally and abstractly” (S14) • Inclusion of expert heuristics • Structuring and breaking down the complexity • You have 3 things happening at the same time (it's not hierarchical • Identifying priorities • When something changes knowing what went wrong • Synthesize quickly • Identifying key and life-threatening issues	• Show various physiologic points over time and in context • Provide zoom in and out versions of the visualizations • Hint towards a certain flow of analyzing the data points • Indicate thresholds for variables and customization to assess when certain actions had an impact on the states • Provide visualizations that are quick to understand and capture a lot of information
Reflections	• Self-reflection	• Providing ways of looking at the data and track previous actions taken (e.g., procedures)

Chapter 4 Developing visualizations to support expertise in neurocritical care

4.1 Introduction

Nowadays, the way of how advanced interfaces are designed is of utmost importance to not only provide all necessary information to take certain actions, but also to show information in a meaningful way and contribute to the end users' mental models when expertise is developed. In healthcare especially, there are different types of end users of such systems; critical care is one of the areas that show complex and highly critical patient situations in which there is a specialized, and often rare accessibility to the area of neurocritical care delivery. This specialization adds a certain jargon to general critical care and deals with neurologically unstable, often worsening situations of patients, where "time is brain", meaning that every second of an indicator for worsening outcome has a tremendous impact on the overall health state with decreasing potential of full recovery. Often, neurocritical care clinicians deal with traumatic brain injury (TBI) patients, who are at risk of secondary brain injury mostly being fatal or having poor health outcome. The equipment and technology used to portray information on the patient state thus play a great role and needs to be considered in detail; the bedside physiologic monitor is one of the key components to think about the trajectory and gathering insights into the patient's health condition.

This chapter first provides insights on the status quo of general display and information design publications, relevant medical information for the area of neurocritical care such as a key variable called Intracranial Pressure (ICP), and the potentials of an ecological interface in this context. Then, the perspectives of critical care physicians will be outlined again on how they are using data on bedside physiologic monitors to make decisions (e.g., trajectory) and what variables are important in context. Further, together with experts in the field, we will explore concepts relevant for neurocritical care that might foster expertise development. Lastly, all insights from the previous chapters will be merged with these findings and ideas for an advanced interface design (i.e., ecological interface design) will be developed.

Overall, the underlying research questions are:

- How can expertise development be supported through interface design in neurocritical care?

To understand the current bedside physiologic monitors and use in daily practice, we want to understand first:

  - How is data on monitors used to develop trajectory for patient care?

Neurocritical care is a specialized area that requires knowledge of specific concepts; Besides focusing on the data and trajectory, we want to know:

  - Which neurocritical care-relevant variables are important to consider?
  - Which neurocritical care-related concepts play a role in developing expertise and why are they difficult?

To include our findings from the CWA models, we would like to explore how these can be combined with the interview findings to help identify design ideas that may support expertise:

  - How do the insights from the CWA models additionally shape the design ideas for the interface?

4.2 Background

Literature in the areas of display and information design, ecological interface design and a key variable tracked in neurocritical care will be summarized in this section.

4.2.1 Display and information designs in neurocritical care

In critical care, various measurements are taken from the patient; a bedside physiologic monitor collects vital signs and waveforms through sensors (invasive or noninvasive). Further, there are ethernet ports that connect the monitor to the multimodal system and hospital network. Connectivity interfaces facilitate the connection between devices that cannot be linked with the vital signs monitor. Usually, these are mechanical ventilators, near-infrared spectroscopy (NIRS), and other devices. Multimodal monitoring provides the opportunity to gain data from different sources and are then organized on a common time scale. This makes it easier for the clinicians to assess the data and come up with a treatment plan. The data can be accessed through the electronic medical record (EMR) system. Physiologic data is usually represented as real-time data, as (a)synchronous data, and trend data. There is also imaging data (CT scans etc.) which need to be rendered. (Dziedzic & Suarez, 2022)

Overall, there is a high volume of data every day associated to each patient in a high-speed environment, making it even more important to represent it meaningfully.

There are studies that take steps towards connecting more devices and accumulating data to portray a greater basis for understanding the health states of the patient. The COSBID M3 monitor is an example of such a multimodal monitoring system that includes BP, ICP, Tissue Oxygenation Pressure (PtiO2), Cerebral Blood Flow (CBF), Electrocorticography (ECoG), Electroencephalogram (EEG) as well as patient videos. (Wilson et al., 2013)

Another research group has developed a remotely accessible and centralized monitor named iSyNCC that includes demographic information, physiological monitoring data as well as treatment-relevant information. Requirements for setting up such a system were gathered through interviews and discussions with clinicians. The intention of this work was to support clinical decision-making and portray computational intelligence that would help send alerts when the health states change, and overall support on decision-making. On the interface, the user can look at events that happened as well as offers the ability to view the history. (Feng et al., 2011)

There are further research studies showing the development of systems that shall provide support to neurocritical care clinicians. One example is the study by Shinde et al. (2017) in which interviews with neurosurgeons and neuro-clinicians helped better understand clinicians' daily needs. Real-time monitoring, accessing data points from the past, recording important events and details, and incorporating data analytics variables were identified to present an effective and visually strong interface. Finding solutions for all these requirements in addition to providing central and remote monitoring, the proposed interface shall improve the efficiency of medical staff.

Also, studies in critical care have been conducted to evaluate integrated displays to let users of the system identify the patient's state faster and more accurately. The display included patient monitoring scales, ventilator, fluid balance, infusion pump, scheduled medication, medication compatibility, and adverse medication effects information. Although this has been a pilot study, the research team proposes that there might be an impact on the reduction of adverse events and cognitive workload too. (Koch et al., 2010)

Another study has focused on shortcomings of existing interfaces and was based on these formulated requirements. Kushniruk et al. (1996) identified that medical users of their interface required a better way to show editable fields, greater consistency (e.g., on data entry procedures and selection methods), faster entry of patient information, templates from which data can be selected, and comprehensibility of graphs and tables.

Different visualization types have been used in critical care interfaces: numeric displays, histogram displays, polygon display, cardiovascular graphical display, pulmonary graphical display, so-called enhanced display, integrated trend display, integrated clock display and many more (as summarized by Andrade et al., 2020).

Currently, there are some interface designs on the market that are considered as more ‘advanced’ displays; Some examples are the:

- ICM+ software, known for brain monitoring, which was developed by the University of Cambridge. Variables are displayed as trend charts and histograms, where variable correlations are highlighted as well. This software provides a Pressure Reactivity index (PRx) related to intact autoregulation monitoring and is becoming more widely used in clinical and research practice. PRx uses MAP in relation to ICP. It further helps to identify what the optimal CPP values for a patient may be.
- CNS monitor is a solution offered by the company Moberg Solutions. This monitor uses a wide time scale and offers the representation of various variables on a single screen. Within the chart, a pie-chart is added for visual support.
- Another advanced interface representation is the T3 Data Aggregation and Visualization software, provided by the company Etiometry. This interface has its roots in the pediatric-cardiac ICU, however, deploys multimodal data in a customizable way. Some additions have been made to the risk analytics: the two indices IDO2 and IVCO2 are provided and got FDA-clearance.

The overall trend nowadays is not only to look at various physiologic variables displayed on a single screen but also to use predictive analytics to diagnose, treat and prognose the patient’s states (Alkhachroum et al., 2022). The examples provided do not all have a strong component related to neurocritical care which this work is contributing towards.

4.2.2 ICP – a key neurophysiologic variable in neurocritical care

Neurological deterioration is often related to changes in brain compliance. ICP is considered a biomarker for many cases such as TBI. It is not only measured numerically but also represented as waveform pattern on many neurophysiologic monitors. Watching the waveform morphology, neurological deterioration can be detected by looking at the waveform pattern. A normal pattern is represented in Figure 6a where the peaks of the wave are placed in a certain order: P1 (percussion)>P2 (tidal)>P3 (dicrotic). When there are issues with brain compliance, the waveform may look like in Figure 6b: Abnormal ICP waveform (P2>P1), (Dai et al., 2020) , showing higher P2>P1. Further, the pulse waveform is linked to the cardiac cycle. Other studies showed that ICP should remain in a certain range to be considered normal. From time to time, the guidelines are updated as more studies are conducted. Currently, the threshold value of ICP is updated to 22 mmHg (earlier 20 mmHg) and should not exceed this value. However, most clinicians still use 20 mmHg as a threshold for high ICP and will thus be used in the studies outlined in the next chapters. Patients are individual and may have different thresholds due to age, treatments, or other parameters (Dai et al., 2020 with reference to Sorrentino et al., 2012). ICP is further correlated with other physiologic variables such as MAP, CPP, oxygenation, and glucose. (Dai et al, 2020)

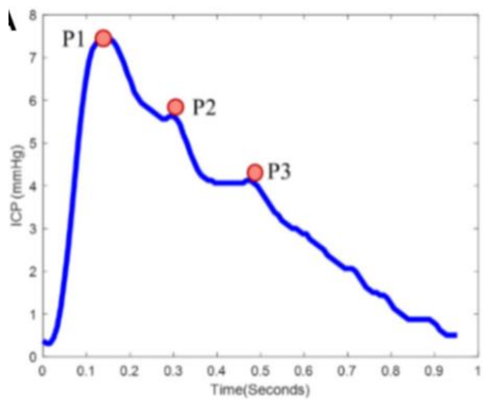


Figure 6a: Normal ICP waveform (P1>P2>P3), (Dai et al., 2020)

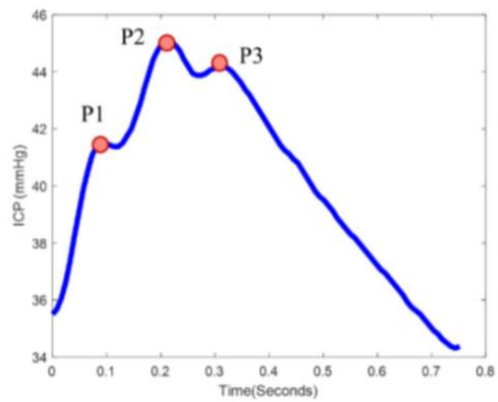


Figure 6b: Abnormal ICP waveform (P2>P1), (Dai et al., 2020)

Other opinions exist on the P2:P1 ratio; a study investigated patients with severe TBI in which they conclude that the "P2 elevation is not a reliable clinical indicator to predict an impending disproportionate increase in intracranial pressure" (Fan et al, 2008).

Further, waveform patterns are classified as Lundberg A, B and C patterns. In the A or plateau waves, ICP remains high over a time period of around 5-20 minutes, together with a larger amplitude. In some cases, the B waves maintain 1-2 minutes and are lower in amplitude compared to A-waves. The C waves are considered as non-significant in terms of pathology but have a smaller amplitude than the A and B waves, while enduring for less than 5 minutes. (Rodriguez-Boto et al., 2015)

ICP is measured both invasively and noninvasively. When ICP is measured through an invasive technique (fluid-filled systems, transducer-tipped catheters, telemetric methods), there may be various constraints such as limited measurement in time and risks of infection. Noninvasive measurements (Impedance Mismatch, Tympanic Membrane Displacement, Transcranial Doppler, Near Infrared Spectroscopy, Optic Nerve Sheath Diameter etc.) reduce such risks but are not yet very accurate. (Kawoos et al., 2015)

Rubiano et al. (2022) further summarize that ICP is often considered a numerical threshold, however, there are other aspects such as the dynamics and mechanics of the fluids that also describe ICP. In the past years, there haven't been many new therapies established or potential ways of controlling ICP been found. Surgical procedures are used to reduce ICP, such as through intracranial decompression.

The duration of high ICP values is crucial to monitor; a study evaluated various cases of high ICP in which a longer duration at high intensities was associated with worse outcome. The authors talk about an ICP time burden that describes the relationship of ICP intensity over time. For adults, a duration of 37 minutes of ICP higher than 20 mmHg was noticed as a poor outcome. (Güiza et al., 2015)
Similarly, another study showed that high ICP time dose was also related to mortality. (Sheth et al., 2013)

Another important concept in this area is cerebral autoregulation.
It is described as the "mechanism responsible for maintaining a relatively constant cerebral blood flow over a wide range of arterial blood pressures". In some cases, cerebral autoregulation may be dysfunctional and poses risks of cerebral swelling. Studies show that dysfunctional autoregulation is often correlated with high ICP levels. (de-Lima-Oliveira et al., 2018).

Methods of data analysis in neuromonitoring are summarized by Kamel and Hemphill (2016, chapter 13) as arithmetic, index-driven, data-driven methods and model-based methods. An arithmetic description is the relationship between MAP-ICP=CPP. Indices exist such as PRx and fever burden. Data-driven methods are e.g., regression analyses, decision trees, neural network and data mining. Model-based methods are dynamic Bayesian networks and dynamical systems models.

ICP is often mesured invasively, however, there are studies where noninvasive assessment is possible too (Rojas, 2021).

4.2.3 Ecological interface design

The EID approach was developed from a nuclear power plant case in the 1960s by Jens Rasmussen. Herein, the interactions and influences of human operators, their used equipment and automation in complex systems has contributed to first findings and the development of the EID framework. Further research has been conducted by Vicente who investigated the effects of such EID onto operator performance. Findings often showed that operators were able to identify issues faster with ecological interfaces. (Burns & Hajdukiewicz, 2004)

A great part of the development of an EID framework has its roots in ecological psychology (Gibson 1979/1986) in which the relationships of humans have been observed within their environment. As explained by Burns and Hajdukiewicz (2004), decisions are limited by the work domain which in return can be analysed to develop understanding of such constraints. Further, cognitive workload in terms of mental calculation or memory can then be decreased by using the right design techniques and representations that relate to the identified constraints. The user shall thus be able to "handle the unexpected" and develop expert skills. (Burns & Hajdukiewicz, 2004)

The main goal is to provide the users with the right (amount of) information to help them develop the skills to prevent critical errors. Overall, EID has been applied in many domains; examples can be found in the healthcare, aviation, energy, and military domains.

4.2.3.1 EID in healthcare

EID studies within healthcare show a great range of applications; research can be found with regards to physiotherapy and rehabilitation, geriatric ambulatory care, hemodynamic monitoring, medication safety, anesthesiology, birthing units, neonatal intensive care, and diagnosis. Some examples are provided below.

A study of EID being applied in an anesthesiology context (Jungk et al., 2000) outlined that commonly vital signs are represented as trends on a timeline. Showing variables only over time is not the only way physicians in that context think and can benefit from a different visual concept for a more in-depth representation. Thus, different visualizations were tested on an ecological interface suggesting that such new representations can improve their decision-making and become more aware of the problem.

Using the WDA, Li et al. (2014) have investigated whether an automatic physiotherapy assistant and rehabilitation system can support the rehabilitation and physiotherapy of patients. With the findings and mapping onto a Part-Whole representation, the key concepts and constraints of the system helped to visually guide a patient with the shown projections. This study showed that the use of automation and EID approach especially in rehabilitation, can be a great support and can benefit different users of such system.

Ecological interfaces have also helped in implementing a new system for geriatric ambulatory care where researchers investigated the needs of clinicians especially in terms of finding patient information. Cranfield et al. (1993) focused on the system being easy to use, clinically useful and providing an open system, client/server and modular design. As participants tested the developed EID, they needed *less time* to go through patient information and perceived *support to make decisions*. However, there were also clinicians who were not very interested or used to using computer systems and thus it caused them more overload as paper records don't display exactly the same information on the EID. From this study, we can see that users may also be discouraged by using a new interface.

Other studies such as those led by Rezai and Burns (2014) show that CWA approaches can inform a system for blood pressure management. Analyzing a work domain and progressing with the control tasks show that blood pressure can be monitored, and the design can help with the patient's *motivation* as well as ability to *take actions* via e.g., communication support.

Lim et al. (2016) examined medication safety problems within care home systems by applying WDA. Seven care homes were subject to the analysis process, where the medication systems have been modeled through an AH. The error analysis has been conducted by looking at different work categories. The AH helped to conceptualize *safety* problems and design interventions to increase medication safety.

4.2.4 EID in expertise research

In this context, CWA helps to develop a rich set of design requirements as well as effective design ideas. These are especially impactful due to their effect on improving performance which has mostly been identified in critical situations when dealing with fault detection and diagnosis. The design ideas influence expert performance, also starting from a novice level. It is thus not only important to understand expertise but also how it can be transferred or developed through design.
Similarities and differences between novices and experts can best be seen in the tasks represented on a decision ladder. Another way to identify similarities and differences in novices and experts can be made

transparent by using Skills-rules-knowledge (SRK) model developed by Rasmussen in the 1980s. According to this model, Skills-based behavior relates to perceptions and actions, the rule-based knowledge helps with evaluating the situation and corresponding actions, and knowledge-based behavior requires higher effort for cognitive processing to analyze the situation. There are two cases: familiar and unfamiliar situations. Experts perform short-cuts in familiar situations, where they can skip certain tasks and directly progress from receiving the information to execution. However, in unfamiliar situations, they would still walk through each step of the decision ladder and assess a case that is less frequent.
Understanding the key components of expert behavior, which can extensively be represented with CWA, a basis for how a novice can reach an expert state faster can be laid. *Heuristic pathways* can be presented to novices that depict expert knowledge as for example shunts illustrated on the decision ladder. Through the Strategies Analysis (StrA), different ways of using the interface can be noted as well. Highlighting possible pathways for the user can support the learning process of a novice user as well as a more advanced user who might have forgotten about prior pathways to reach their goal. (Burns, 2019)

To deliberately and actively support the development of expertise, *persuasive design* can be a way to affect behavior change. Studies showed that fitness behavior or medical data entry showed promising results on how to "train and develop patterns of the best performers" within their system. Showing the operator to consider different or new pathways could be developed by looking closely at the decision ladder or strategies analysis. To nudge operators' behavior, progress towards a set goal can be visualized. (Burns, 2019)

Some examples of research studies use the EID approach in different domains.

Howie (1996) for example investigated the effects of participants using a Dual Reservoir System Simulation (DURESS) II with a traditional vs. an ecological interface in terms of expertise development. DURESS showed a thermal-hydraulic process control that integrated aspects of a complex system as well as variables that are interrelated. The main difference between the traditional and EID was that the traditional interface included only physical information whereas the EID incorporated physical and functional information. This 6-month experiment was designed to examine how participants adapted to the given systems, and findings were compared between the traditional vs. ecological interface. Operators' mental strategies, metacognition, their individual differences (cognitive learning styles), and the way they interacted with the interfaces to reach the task goals was analyzed in this study. Howie's findings show that participants fulfilled their task goals *faster* when developing more expertise and showed more *consistent* values within the tolerance ranges of the controlled system. Other tools she used to analyze these findings crystalized that participants developed more *direct* and *systematic approaches* to operate the system with growing experience over time.

Jamieson and Vicente (2001) explored the application of the EID framework within petrochemical processes. Investigating adaptation to novelty, continuous learning, and distributed collaborative work, they applied the WDA approach to display the system and its constraints in this complex and dynamic field. The authors claim that an ecological interface suits well in this domain because it represents information and their relations to various people working with petrochemical processes, ranging from operators (who would consider looking at information of the Physical Forms level on the AH) up until managers (who monitor overall goals that are represented on the top level of the AH, named the Functional Purpose). The authors thus showed that the EID presents great benefits for various users and supports them within *collaborative procedures*, *learning processes,* and dealing with *unanticipated events*.

A study by Borst et al. (2019) showed short-term training effects between an ecological and instructional group in air traffic control. The study examined the decision-making strategies through think-aloud protocols and findings were mapped on the DL. Novices' performance in conflict detection and resolution was compared to the instructional group. Although this did not show significant differences between the groups, it was noticed that the students in the ecological group showed greater goal-oriented thinking paths and 'control performance beyond the control detection and resolution task'. Borst et al. (2019) highlight the relevance for ecological interfaces to especially train novices "in the early stages of deep knowledge development".

Morineau et al. (2009) investigated the differences between a classical tide display (in table format and Maréegraph format) and an ecological, i.e., tide prediction display. Results showed that novices made more errors compared to experts, the response time was quicker on the ecological interface and showed that they performed better when looking for abstract data. Overall, participants liked the EID better.

Further, in a study on EID in neonatal ICU has been conducted with physicians in different groups (attendings, fellows, and residents, in total 16 participants). The residents especially showed greater differences in the accuracy of diagnosis as percent correctness, when they used the EID vs. existing interface design; the percentage of correctness was visibly higher on the EID compared to the existing interface for residents. (Sharp & Helmicki, 1998)

Another study investigated differences between hemodynamic monitoring and an ecological interface. Changes in the pressure and flow variables were provided related to disease states and given drugs. Nurses and nursing students who used the interfaces showed higher speed and accuracy on the ecological interface. (Effken et al., 1997)

4.3 Method

To further engage with end-users of bedside physiologic monitors and understand their perception and use of monitors, but also in the context of neurocritical care, the following section shows parts of the previously described 60-minute in-depth interviews with the same critical care experts and trainees.

The second part described shows reflections on three key neurocritical care concepts that have potential to lay a basis for initial design ideas for the ecological interface. These reflections are provided through experts during the 60-minute interview.

The third part depicts the relation to the previously explained CWA approach which constitutes the links from the basis to a more sophisticated idea repertoire for the ecological interface.

All these three parts will be described separately first, and then merged into a common context for initial visualizations design for neurocritical care bedside physiologic monitor interfaces.

4.3.1 Interview content

The semi-structured interview provided the opportunity to ask staff and trainee physicians questions around:

- The use and importance of bedside physiologic monitors in the care of patients in neurocritical care.
- The way of perceiving data displayed on the interface and how this helps to understand trajectory.
- The key concepts related to ICP monitoring.
- The variables related to ICP monitoring.
- The difficulty with neurocritical care concepts.

The same data analysis approach as previously mentioned in chapter 3 has been continued in this part.

4.3.2 Expert consultations

The collaboration with our neurocritical care and subject matter expert Dr. Victoria McCredie highlighted three main neurocritical care concepts that constitute the basis of knowledge and practice for neurocritical care physicians.

Keeping these three concepts in mind, the intention was to further get feedback on these concepts from various (neuro)critical care experts in the field to see if the concepts resonate with them in this context as well.

4.3.3 CWA: EID findings

The abstraction hierarchy represented in Chapter 2 shows various abstract levels and connections for neurological systems, treatment and monitoring as well as challenges faced in this area. There are multiple things that can be extracted from this model that would inform certain design decisions for new visualizations to be represented on a bedside physiologic monitor within neurocritical care. Such links will be displayed in the results section.

4.4 Results

The results entail findings from the interviews, the neurocritical care concepts of Dr. McCredie and reflections on them from further experts in the field, and consideration of the CWA approach to delineate core aspects for the creation of potential neuro-visualizations to be displayed on a bedside physiologic monitor.

4.4.1 Interview findings

The interviews captured multiple parts on understanding the importance and use of the bedside physiologic monitor, the use of data for trajectory, key concepts of ICP monitoring, variables related to ICP, and the difficulty of understanding or conveying neurocritical care concepts in general.

4.4.1.1 The importance and use of bedside physiologic monitors

In the interview, trainees were asked about how they use the bedside physiologic monitor and what they think the importance would be of such display.

Often, they mentioned they wouldn't use the monitor solely, as it lacks relevant information for their decision-making process and is limited in its usability. Trainees highlighted having a stronger focus on the flowsheet for trending the patients' states: "The bedside monitor doesn't provide any historical or very limited historical data. It shows the current real time snapshot of various physiological parameters" (T4), also because "paper flowsheets are simpler sometimes, some interfaces are not user friendly and have an information overload" (T5). "Monitors delete data after a certain amount of time, then the flowsheets are used, so that nothing is missed" (T13).

The interface has various purposes, one of them being immediate feedback to treatment:
"Real-time feedback is very important" (T14), as well as checking the reaction to the intervention by looking at curves (T6).

The high cognitive workload is a contributing factor to have a more numerical approach:
"Cognitive power is low when being tired and this impairs the ability to think critically through the situation" (T10). As a result, the trainee mentioned that sometimes they are quick to act on a number because that is the easiest piece of information to comprehend although there is a lot more to patient management than just responding to a number and he tries to limit these "reflexes" so he can think more broadly and critically about each situation (T10).

There is also lacking knowledge about the functionality of the monitors, the conjunction of values and individualizing care:
"People don't even know what our monitors can do, like stroke volume variations or like cardiac output. What are the indications, what are the limitations of each monitoring. So, for me it starts off with you needing to know the technique that you're going to use, how you obtain the numbers or the values and when those values don't work. So, you start by not even applying it incorrectly. And then I've always said, when I teach them [referring to residents or novices], it is never about taking one value, always use a conjunction of values that will be more reliable about where your patient is, because patients don't read books and books can give us an idea, but in reality or care, it ends up being really personalized. So, I trust monitors, but monitors are never going to change mortality or outcomes because that will only come from how you

interpret their results, they give you. So, we have monitors for everything, but we don't use them all the time because for me it is about what do you need in this patient for a particular reason. So I do believe that individualized care based on the knowledge, the theoretical knowledge behind is the best way to interpret monitoring, and then what I mostly do is cardiac and hemodynamic monitoring, but in general, I do believe that the monitoring tool should be used appropriately." (T14)

4.4.1.2 The use of data for trajectory

As trainees learn to gather information through various ways (e.g., flowsheets, monitors, etc.), the staff physicians were asked how they use the past and current information to build trajectory of the patients' states.

The main themes experts talked about when considering past and current values to predict future states evolved around the importance of equipment and response to therapy, as well as the consideration of correlating variables over time.

"The trends over time are important to evaluate the effectiveness of treatment, looking back (at the flowsheet for the past 24 hours) is important too (i.e., historic trajectory). For real-time, you do tests in real-time to help guide your decision-making" (S17)
Getting an understanding of the full picture is highlighted by (S10): "A great emphasis is on physiology and monitoring: 1) clinical exam and context, 2) radiology, 3) neuromonitoring. You need at least 2 out of 3. […] Neuromonitoring alone doesn't give you those answers that one needs, and we always get into problems in medicine when we make decisions off of 1 variable. […]
The clinical examination with diagnostic tests help. It's about the integration of multiple variables and data points to make educated decisions" (S10).

Considering a longer timeline is important, "I try to encourage discussions in rounds 12-24 hours pattern. Often, when nurses change their shifts, it is important to know the 24-hour trend and where we are going. I look at ICP waveforms and patterns (also for EEG). It helps to know where you are coming from, discussing with the physician in charge what has been the trend for the past few days (improving or getting worse) as background information and then building up from own experience when starting a new week" (S13).
"The disease process (how does it affect the organs; I need to know the natural history and actual function of the organs), all of this as a baseline starting point. Consideration of parameters before a patient became ill is important and specifics of the illness. Then the interventions and effects. Looking at trends, if aberrancies happen, I check the monitors if they record properly" (S14).

The amount of time a patient spent in certain conditions (dose) is another indicator for trajectory:
"The consideration of past information and considering conflicting data, how much time they spent in certain conditions and its meaning for the patient" (S4) have been identified as important contributors.

Also, when numbers look atypical or not as expected, one staff physician mentioned he would dive into the details more closely: "If there are unexpected values then I pay extra attention to tracing, reassessing the number after interventions, and do diagnostic tests again (S8). However, it is necessary to keep in mind that therapies may not always be ideal or solve the whole issue; thus, it is important to look at it as a whole:

"every time you introduce a therapy, there's a downside to each therapy. The solution also has a downside, if it doesn't, it's relatively easy to deploy it. I evaluate if the impact of the slope is on the favorable side or the negative side" (S7).

4.4.1.3 Key concepts of ICP monitoring

Both trainees and staff physicians have been asked to share what they think are the key concepts related to neurocritical care. ICP is a key indicator for many neurocritical care cases, and thus, ICP monitoring has been taken as one of the most relevant topics to learn and to develop practical skills as well as theoretical understanding.

Trainees pointed out that familiarity with the equipment is important "it might be scary as you don't know how to fix that; once you know, you just follow your experience and become more confident" (T5). The understanding of the equipment, how it works, where it is placed, what the values mean, and how to act on it and evaluate the patient's condition has been brought up by trainees (T12, 13, 16).
One key aspect in addition to the equipment function is to understand the waveforms and trends which is often seen to be done by attending physicians, checking for abnormalities or patterns of waveforms, whereas trainees look at numbers (T8). Similarly (T4) added that the waveform is transduced occasionally and "some staff teach this, so if there is a patient with a change in the neurological status, the physician would put the waveform up and look at the waveform, what the different components of the waveform look like. Very few staff would do that, […] we would say the waveform looks OK or dampened or inaccurate, or whether the monitor is blocked or not in the right place".
Although the number itself may not tell much or might even look normal, it is crucial to understand the whole pathology (T9). Correlating variables have been pointed out by trainees such as a link between ICP and CPP, neurological state, hemodynamics, imaging and medication.

Further, "the concept of pressure reactivity, I've never actually seen it being done in real practice, but it's a very good concept where if you do something you try and do a certain intervention and then you look at how the ICP responds and see whether you increase the blood pressure and see if the ICP comes down and I was supposed to map the trends over time to see where what an ideal blood pressure would be for the lowest pressure in the brain. […] I think that's something which would be very, very useful, but it's very labor intensive unless there's like an algorithm or something that can do that for us" (T9).

Staff physicians have similarly outlined common aspects such as knowing about EVD management, waveform understanding, physiologic and pathologic changes (S1, S10)), but also the risks of inserting the ICP monitors and learning about the expected responses to treatment and recognition of pitfalls, false readings, limitations of the equipment and the evidence behind it (S12, S16). "The research is evolving; the guidelines are predominantly driven by expert opinions" (S12).
The expectations on the trainees (i.e., fellows) would differ among the year of training, i.e., for first year fellows, it is often expected for them to understand "who should have them, what are counter-indications, what are options for monitors, how do the waveforms look like, what numbers are ok and what's damaging. For second year fellows, we expect how ICP integrates into that, talking about more nuanced trials, integrating harmonics and waveforms and use software (e.g. PRx or CPP optimal)" (S16).
"Certain patterns are at the extreme (if the patient is dying, but someone in the middle is more difficult to appreciate). E.g., when looking at BP changes, we can see how the ICP responds. There are both static and

dynamic components, which are another layer of complexity. Verbalizing and saying it out loud is important. Trainees want exposure“ (S5). At the same time, “each step requires significant knowledge, and the number of steps is very high (for correct interpretation), there are complex interdependencies” (S3).

Certain concepts are relevant in this process too, such as the “Monroe Kellie doctrine, the ICP waveform morphology, pressure-volume curves, autoregulation and how that might be impactful to look for passive perfusing patients, modifying BP goals, altering the drainage of the EVD, as well as the impact to CPP. For people who are interested in more advanced topics, we look at PRx” (S11).

4.4.1.4 The variables related to ICP monitoring

As many trainees and staff physicians mentioned to correlate ICP to other variables in context, a follow-up on this aspect was initiated. When trainees were asked if and in which order they would consider reading ICP in context, most participants mentioned they would look at the numbers or thresholds first, then either the trends or waveforms as a next step.
In context, they would relate also to other variables such as:

- Pupil size
- GCS
- EEG
- Temperature
- PRx
- Ultrasound Doppler
- EVD, CSF drainage
- Fluids
- CT
- Vital signs in general, BP, HR, etCO2, RR, MAP, CPP
- Oxygenation, saturation, CO2, PbO2
- Medication
- pH
- ABG
- Glucose
- Urine

Staff physicians mentioned they have a stronger focus on the waveform and trends. One expert said that the numbers might give an alert, but it is important to trace, trend and correlate the information within the context (S8). Understanding the trend provides information on the ICP burden and what is causing it (e.g., response to medication), whereas the waveform is key because it provides information on intracranial compliance and decompression (S18).
The consideration of other variables correlates strongly with the ones mentioned by the trainees.

4.4.1.5 The difficulty about neurocritical care concepts

“The brain is like a black box” (T4, T9); this statement has been made to underline that neurocritical care is considered very challenging. Participants talked about various aspects that make neurocritical care and underlying neurocritical care concepts difficult. These are related not only to the content, but also the related mentality, skills and data as well as equipment.

Learning the complex content takes time, “it needs time to develop a mentality to be able to act effectively in a specific area. I felt much more confident to use the relevant concepts and evaluate neurocritical care patients at the end compared to the first weeks” (T4).
Also, the transfer of knowledge is deemed challenging “it’s very critical to convey concepts to residents” (T13). In terms of skills, one fellow mentioned that the MRI readings are also important “we are not supposed to read the MRI, but it’s a very important skill for us to know what to look for.” (T3)
Other skills to learn include using the transcranial doppler, EEG and EVD, knowing what to do and what not, considering what the goal is and what would not be appropriate (T3). “It is not a single concept, e.g., ICP and EVD may be complicated at the beginning but the longer you use them, the easier they become” (T6). Another participant mentioned having a good measurement itself is challenging, detecting and predicting before there is permanent damage (T4).
Participants highlighted the challenges with the equipment and data representation: “Some ICP monitors only show the numbers not the graphs which is more difficult to interpret, the graphs provide more information that is important. It’s important to look at the evolution of the waveforms” (T7),” not everybody knows how to read the monitor (e.g., Electrocardiogram (ECG) or ICP), it's not often used and not so easy to explain why it responds in certain ways to certain interventions, and when we would be concerned, what kind of limits we would be concerned with and that's a bit arbitrary and different in every center” (T9), “people don't know much about neuromonitoring, they don't know the values, what it means and how to act on that. We are not completely trained and it's not fully available. Sometimes the equipment is hard to use. There are barriers for adoption.” (T12).
Others have mentioned that “waveforms can be difficult to explain, numbers are more straightforward” (T11), and the relevance of “tracking fluctuations would be much clearer than a number” (T18).

Staff physicians related to similar themes when it comes to difficult neurocritical care concepts and challenges in explaining these to trainees. Common themes related to content, skills, data and equipment but in addition, individualization.
There are also challenging neurocritical care concepts for staff physicians that are not straightforward “the intersection of both areas like acute brain injury and acute lung injury can be challenging, as it is required to balance both” (S1), “airway skills are also challenging” (S17). There is “a lot of information to take in, critical care conditions are new to the residents, how to evaluate trajectory is overwhelming. The education plan focuses on the basics” (S17). As a staff physician, it is further challenging to “being able to explain the different facets to medical learners and respiratory therapists” (S1) and teaching “how to become comfortable with an acutely deteriorating patient and decision-making around triaging, recognizing patterns and anticipate who may deteriorate or not” (S17). Usually, there are “very dynamic measurements” and there is a factor of time related to this; an “increased ICP for short amount of time can be very injurious, and transient changes might be picked up by nurses. When you come to see the patient, they have resolved it and are normal, so you don’t see the change on the ICP monitor. You can’t appreciate the cause for the change if you haven’t seen it. It is important to be able to review the data and what was happening in that moment, for example, was the patient waking up or coughing up.” (S1).
“The waveform is very important, for trends it’s good to zoom out and look back” (S12).

Further comments related to various information and equipment that portray the full picture on the patient state, “balancing medical information coming from the ICP monitor or neurocritical care imaging and your measures for mechanical ventilation, hypoxia, and lung physiology” is challenging. Another participant had a similar opinion that it is required to "not only look at an individual parameter, but at multiple streams. This requires more repetitive experience with different combinations of parameters and takes time to learn" (S5).

Another expert thought about the benefits and limitations with equipment which should be considered as well "we would like to know what to do with those monitors. What are the indications, what are the benefits, how to manage a number or so. So, there's a lot of gray zones, I think in neurocritical care and how those monitors should be used and who should have them and what to do with the number. And I think that creates a lot of insight in which it needs to be discussed with the more senior trainees but where often the recommendations are" (S8). One should know the "limitations of what the monitor shows to you, even if you are at the bedside all the time" (S12). There might be "instrumental problems" and when using the monitor, one needs to check the "right anatomical space. Is it the right place reflected on the ICP monitor?" (S1). The "changes in pupils or the level of consciousness might not be reflected in the measurements, and one needs to be able to express that to other team members too" (S1).

Some comments evolved around the challenges with individualization of care "a personalized approach can be challenging, generally trainees require lots of guidance and supervision" (S2). This statement was also highlighted by another expert "discussing with trainees the nuances of care is important. Recommendations are often vague, so often individualized patient care is discussed" (S8).

4.4.2 Concepts of neurophysiology and expert feedback

The discussions with our collaborator and neurocritical care expert Dr. McCredie led to more in-depth insights into neurophysiologic concepts that are key to integrating into the training of trainees in neurocritical care.

As described by McCredie (2019), there are three neurophysiological concepts (Figure 7) that have an impact on the prevention of secondary brain injury especially: considering a secondary insult dose, the integration of an individualized thresholds approach, and overall trajectory.

The concept of **secondary insult dose** relates to the importance of the time burden of ICP as mentioned in Guiza et al. (2015). It describes the link between the intensity of ICP and the duration of ICP spent in certain ranges. The more time spent in high ICP intensities, the higher the risk for secondary brain injury and worse patient outcome. (McCredie, 2019)

The **individualized thresholds** approach relates to the significance of additional factors such as patient specifics (e.g., age) and pathological differences etc. Although current guidelines indicate a normal ICP range between 5-15 mmHg and a high range above 22 mmHg, all patient information need to be taken into consideration and requires an individualized approach as it is not 'one size fits all'. (McCredie, 2019)

The **trajectory** is a key consideration for understanding the past, current and future direction of the individual patient. An instant change from 9 mmHg to 19 mmHg (from the previous day) has a different meaning than a patient who has been staying at 21 mmHg for the last 2 hours. (McCredie, 2019)

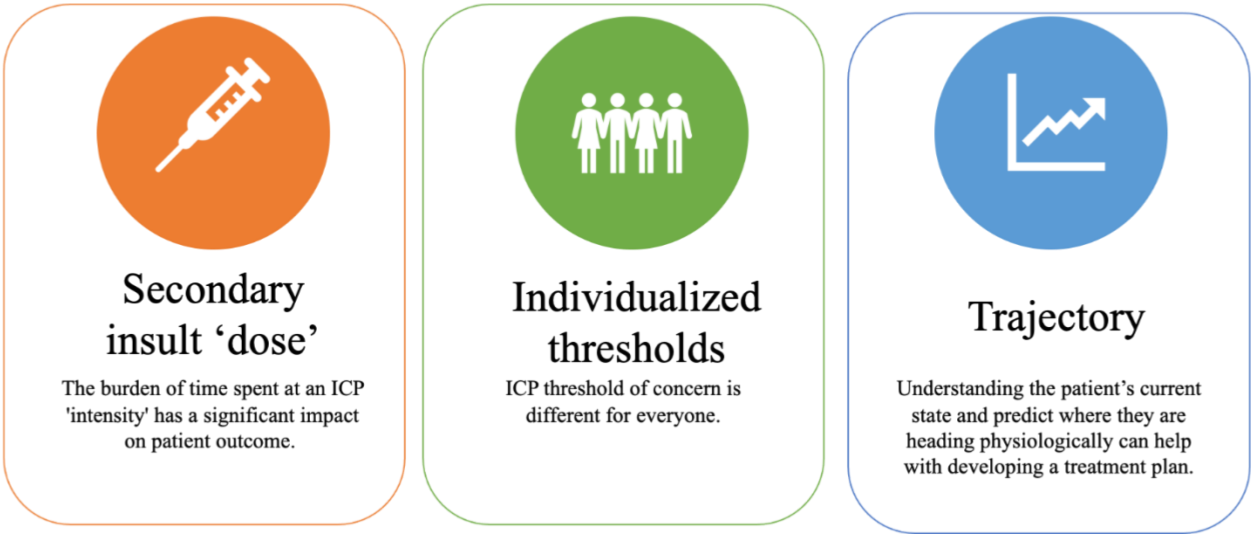


Figure 7: The three key neurophysiologic concepts (McCredie, 2019)

Trainees and staff physicians provided similar inputs about relevant concepts in this domain. However, during the conducted interviews, the three neurophysiologic concepts were displayed and discussed with further experts in the field to gather more insights as well as consider potentially lacking concepts that could ultimately play a key factor in supporting to overcome challenges with design elements on a bedside physiologic monitor.

The main themes captured were:

- Waveform analysis
- ICP in context
- Autoregulation
- Bayesian approach and risks of trajectory
- Causes of ICP elevation
- Different starting points

Two staff physicians mentioned that they would equally be interested in integrating waveform analysis as a neurophysiologic concept. Often, the waveform analysis is related to understanding the peaks of the waveform (i.e., P1, P2, P3) and specifically the ratio of P1:P2 as an indicator for normal pressure compliance in the brain.

Looking at ICP within its context has additionally been highlighted by some staff physicians: "There's a really important issue about what else is happening clinically at the time that there is an ICP elevation" (S13).

One staff physician mentioned that the concept of autoregulation is one of the major topics in understanding neurophysiology, as ICP is used with ABP to determine intact autoregulation of a patient (S10).

Another staff physician added that the concepts could benefit from "Bayesian thinking" (S11) which refers to a more statistical way of looking at the occurrences and probabilities of certain conditions. The risks of the trajectory were mentioned by this physician too, that require critical thinking about potential outcomes and the evolution of the patient states.

Understanding the causes of ICP elevation has been pointed out as another core concept to consider "we don't think enough about the cause of elevated ICP" (S8).

Patients may be in different states of their disease process, which is related to the theme of having different starting points of care.

4.4.3 Consideration of CWA and implications for design

Overall, each previously described section or chapter provides ideas that will be considered with the intention of letting trainee neurocritical care physicians develop expertise. In the context of developing visualizations for the support of expertise development through interface design, I would like to outline aspects covered in Chapter 2 that yet haven't been taken into consideration from a visualizations-perspective. Thus, the following pieces will be further explained in the subsections relating to CWA laying a basis for visualization design too:

- the abstraction hierarchy representing the neurocritical care domain,
- the DL pointing out the differences between expert and novice neurocritical care physicians when treating common cases such as increased ICP patients,
- the gained insights into expertise development from the observations and discussions with the clinicians

4.4.3.1 Design ideas retrieved from the WDA

The AH has shown many links between the aspects represented on each level of abstraction. It can be used to lay a basis for interface design. When looking at the physical forms level for example, we can see that patient specifics are very important to provide the best care for the patient. This implies that individualized approaches of care are necessary. Examples on the AH have shown one part that directly concerns the patient:

- the patient type (e.g., adult or pediatric) does play a role in how treatment is provided,
- the patient condition (obese, elderly, frail etc.,) impacts the thresholds of variables and expectations on treatment reactions, as well as the condition (ICP, HR, BP, lung stiffness) of certain variables which give direction on monitoring and treatment intensity,
- the disease (type, state) which implies the treatment options and severity of condition,
- the damage (type, amount, location) that also shows the severity and options of treatment,
- the relationship between multiple variables,
- the signals (frequency, amplitude, location, waveform).

Another part of the AH shows the challenges faced which are not limited to the location or communication, but foremost to:

- the technology (visualization via screen, computers) that impacts the way information is stored and shown to users,
- the data retrieval (past, current, future) that carries major importance to the clinician to understand where the patient came from and build trajectory,

- and capacity (storage) that is needed in great sizes to capture seconds of data and be able to trend the data.

All these points show current challenges and the need for improved interfaces. Potential design ideas for an ecological interface could address various factors; we can think of visual elements in which e.g., patient information and individualized approaches of variable customizations, as well as visualizations over longer periods of time for each patient and the ability to retrieve this information at any time are reflected. The relationship between various health indicators or variables plays another key role for which different types of graphics can be considered to display interconnected variables (e.g., a summation chart). In this case, one has to identify which variables are dependent on another and how they impact each other under various conditions.

4.4.3.2 Ideas retrieved from the ConTa that can support novices

The previously shown DL was displayed to show a common neurocritical care case when patients suffer from increased ICP. The DL has shown differences between an expert and a trainee taking up such cases and provides insights into the struggles of trainees. At exactly those points of difference, an ecological interface could provide decision support to the trainees, e.g., by providing deeper insights into the aspects they are missing out on. Notifications or providing more context into the data may be considered as such support.

The set of observations for instance, shows differences between experts and novices. At this point, the EID could help to gather all information at a glance or provide accessibility to monitor all relevant variables.

As a next step, the identification and understanding of the system state of the situation seems challenging to novices. The EID could integrate certain ways to display information, in order to let the user become aware of data trends and identify any concerns with the data if they are outside of standard ranges.

The evaluation process of all information is the most difficult part to come up with a diagnosis and treatment approaches. Visualizations could thus support transparent ways of showing all patient information and data trends.

Followed by the interpretation of all the information as well as definition of tasks, the ecological interface could provide visualizations that provide setting up thresholds for variables, implying that certain actions of care should be considered.

The example of a DL is represented below in Figure 8 that shows possible inclusions of support through the EID so that it resembles more the expert DL.

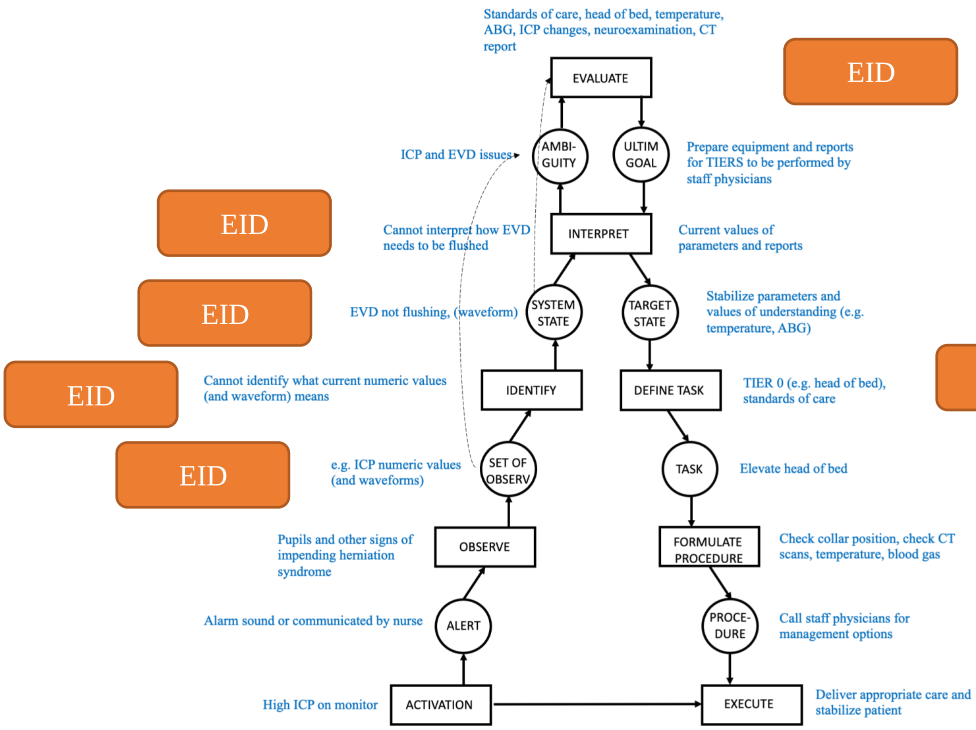


Figure 8: Decision Ladder of novices in neurocritical care with EID

All these aspects have been considered for designing visualizations to be displayed on an ecological interface for bedside physiologic monitors.

4.4.4 Design ideas coming through the observations and discussions with clinicians

Major ideas on how expertise is developed in such a complex area have been mentioned in the above section. Here, we are pointing at certain aspects that have great relevance and the opportunity to be integrated into the bedside physiologic interface design.

One key aspect has been observed during ward rounds. The clinical team gathers around the patient bed to recall the patient's case including all data from various systems, background information, and discussion of the current and expected future states. Rounding is a routine activity that lays the starting point for handovers where all information about the patient is discussed. The data that is represented on vital signs monitors shows near to real-time data as waveforms and numeric values and is limited to showing historic data points on the screen. Multiple data sources and systems exist in the ICU and there is no main screen that merges all data at once. Although some hospitals use digital clinical flowsheets (and others paper flowsheets), trending of data happens rather in the experts' minds by comparing previous data points to the current ones.

We thus see that:

- merging all data on one central screen might be deemed beneficial,
- the integration and accessibility of historic data is crucial to come up with trajectory,
- customizability of the time for any variable is relevant to detect and focus on certain time intervals and the trend of the variable.

Another key aspect can be seen during trainee rounds. Usually, one expert critical care physician takes the time to explain a certain condition to the trainees. Whether it is discussed verbally only, or visualized with pictures on the board, this session aims to support trainees to better understand (rare) clinical cases, with considerations of guidelines or textbook theories for instance. These sessions could be guided with the bedside physiologic monitor that would portray more extensive visualizations of patient cases with such conditions and make it easier to understand the background and full picture of all data points throughout the patient stay in the hospital (or even previous hospital stays if data is available). These visualizations could thus show:

- more data represented over time with potential time customization,
- the ability to select certain patient cases to show the data and look at their individual conditions.

Furthermore, in times when there is absence of guidance and instruction, trainees need to rely on their knowledge, skills and previous experiences. Checking the protocols or guidelines for certain treatments can take time and reduce the time they would dedicate to the patient's case. Although hospitals often have their own protocols, certain strategies of treatment have been manifested (inter-)nationally. This aspect could be integrated in design by:

- Providing certain treatment thresholds for patients that could be set up by clinicians (e.g., the attending physician could enter common ICP thresholds into the interface that would show normal and high ranges for the variable),
- Before exceeding certain thresholds, the interface could visually highlight ranges of data for variables, that show an increasing or decreasing trend.

Receiving feedback does not only have to come from clinicians. It can also be provided through systems that support the process of understanding the data shown. An idea to integrate system feedback or advice could be shown through:

- unobtrusive notifications that draw the user's attention towards certain trends in data. These have to be explained and designed in a transparent way, so that clinicians understand when and how it would show up.

Also, repeated practice has an impact on building routines and recognizing patterns. Having a routine check on the interface (i.e., vital signs, medication etc.) can influence the way clinicians think. Thus, it is important for trainees specially to develop their mental models on how to proceed and recognize correlations between the variables represented on the interface. As most current ICU monitors represent line graphs or numeric and waveform data, a design consideration could be to:

- represent variables not only as line graphs but use different types of visualizations,
- show combinations of graphs that depict a greater context of the ongoing,
- display correlations of variables graphically.

4.5 Discussion

The depicted streams of gathering information and defining more explicit ideas for visualizations that could support especially trainees in neurocritical care, have been merged and summarized in key points within Table 4 below.

4.5.1 Reflections on the study

Although other studies may show only one of the represented approaches to come up with visualizations, we have chosen to combine methods to dive deeper into this complex sociotechnical domain and strive for a more complete representation of the work domain, how task are controlled, how do experts and novices in the field work with difficult concepts and how can these certain biases be mitigated. Not all aspects of the results will be included in the visualizations due to feasibility and time constraints, however, they offer insights that other researchers may benefit from to investigate potential research questions further.

Table 4: Elements for design through CWA, discussions and concepts

Methods	Sub-methods	Resulting elements for design
Modeling: CWA	Observations and discussions	- merging all data on one central screen might be deemed beneficial - the integration and accessibility of historic data is crucial to come up with trajectory - customizability of the time for any variable is relevant to detect and focus on certain time intervals and the trend of the variable - more data represented over time with potential time customization - the ability to select certain patient cases to show the data and look at their individual conditions. - providing certain treatment thresholds for patients that could be set up by clinicians (e.g., the attending physician could enter common ICP thresholds into the interface that would show normal and high ranges for the variable), - before exceeding certain thresholds, the interface could visually highlight ranges of data for variables, that show an increasing or decreasing trend. - unobtrusive notifications that draw the user's attention towards certain trends in data. These have to be explained and designed in a transparent way, so that clinicians understand when and how it would show up. - represent variables not only as line graphs but use different types of visualizations, - show combinations of graphs that depict a greater context of the ongoing, - display correlations of variables graphically.

Methods	**Sub-methods**	**Resulting elements for design**
	WDA	- address patient information and individualized approaches of variable customizations, - as well as visualizations over longer periods of time for each patient and - the ability to retrieve this information at any time - The relationship between various health indicators or variables
	ConTa	- gather all information at a glance or provide accessibility to monitor all relevant variables, - integrate certain ways to display information to let the user become aware of data trends and identify any concerns with the data if they are outside of standard ranges, - support with transparent ways of showing all patient information and data trends, - setting up thresholds for variables, implying that certain actions of care should be considered.
Literature and discussions: Neurophysiologic concepts	Three main neurophysiologic concepts	- Representation of the secondary insult dose, - Visualizing individualized thresholds for variables - Supporting to build trajectory by providing time customization to track past and current data and trends
	Additions through discussions	- Visualizations supporting the waveform analysis, - understanding ICP in context, e.g., by representing multiple variables in correlation - integrating notifications for autoregulation, - adding annotations for potential causes of elevations, - showing time of admittance and disease state, - support with considerations on risks and probabilities of trajectory
Discussions: ICP monitoring and neurocritical care concept difficulties	Key ICP monitoring concepts	- helping with equipment familiarity through similar visualizations in different contexts, - integrating details on waveforms, - portraying a more transparent way of data for developing ability to trends and understand the full pathology, - representing variables in context, - referring to certain relevant concepts as e.g., concept pressure reactivity, - highlighting specific variables and information about the patient for more nuanced care, - showing variables as interdependencies,

		- referring to concepts like the Monroe Kelly doctrine or autoregulation through notifications
	Difficulties with neurocritical care concepts	- enabling the ability to easily demonstrate and convey the representations to others (transfer of knowledge), - integration of various outcomes of e.g., MRI results into context - representing clear signs for appropriate function of the equipment (e.g., if there are gaps of measurements, this should be displayed transparently), - enabling a bigger data storage and accessibility for past and current data sets to facilitate prediction before permanent damage, zooming in and out - show intuitive and basic descriptions of represented graphics - include easy to read waveforms, - provide annotations about intersecting injuries, - provide details about the patient to explain the different facets of the patient state and individualized care - provide graphical representation where it becomes easier to recognize patterns

4.6 Conclusion and visualization ideas development

This chapter merged different streams of methods to investigate potential ways of expertise development through interface design in neurocritical care:

- parts of the interviews were conducted with staff and trainee critical care physicians sharing their experiences and views on bedside physiologic monitoring in neurocritical care,
- discussions with experts in the field were conducted to support the idea generation for important concepts that need to be understood to enable the best care for the patients,
- as well as results and implications from the CWA models were gained that provide relevant aspects to be incorporated into an ecological interface.

Data is used differently by novices and experts; while novices often have a more numerical and threshold approach, experts develop a trajectory for the patient through looking at the big picture (including various states and variables' progression over time) but also considering details that may be signs of indicating deterioration. Thus, the waveform morphology is significant too, which is often neglected by novices.
When participants commented on the importance and use of bedside physiologic monitors in critical care, it was highlighted that there are unmet needs of trainees. The identification of relevant information not being displayed to its full extent is a limitation that should be considered in future design. Aspects about the limited usability and not even knowing what is all possible to view on the interface show that there are technical usability limitations but also a component of lacking training on the display functionality or customizability. Participants deem the monitor very important as it provides immediate feedback to treatment that is relevant to see. However, the high cognitive workload is often the reason for a rather

numerical approach (i.e., acting on numbers). Individualized care does not only relate to (theoretical) knowledge but also to how the data is being interpreted. Here, it is already possible to see challenging aspects that should be considered for future design and will be taken into consideration for the visualizations created in this work.

The use of data for trajectory purposes showed considerations of staff physicians regarding the importance of equipment being relevant for response to therapy but also to consider the appropriate functionality of the monitors. The correlation of variables, waveforms, and patterns show distinct insights to be merged into a full picture. The understanding of data is important to identify disease processes and also the amount of time spent in certain conditions. Conflicting data must be understood, and in case of atypical or unexpected values, critical care physicians try to trace it back, reassess the number of interventions, and run diagnostic tests again. The therapy provided may not always be ideal, but they try to assess and provide the best possible way to support the patients.

While trainees shared their perspectives on key concepts of ICP monitoring being familiarity with equipment, understanding waveforms and trends, looking at numbers, understanding the whole pathology, correlating variables, and the concept of pressure reactivity, staff physicians considered further aspects. The management of EVD, risks of brain insertion, limitations of equipment, and the evidence behind it show links to the technical side. False readings happen and to be able to correctly interpret information, it requires a high number of steps. There is also importance in understanding the waveform morphology being represented on such interface, the consideration of autoregulation of the brain, indications for that on the pressure-volume curves, and the overall response to treatment and recognition of pitfalls. Another concept pointed out was the Monroe-Kelly doctrine. Staff physicians also mentioned that the expectations on the trainees differ among years of fellowship training. For example, first-year trainees are expected to learn who should have an ICP monitor, what are counter-indications, what are options for monitors, what do the waveforms look like, what numbers are ok, and what's damaging. The second-year trainees shall take a step beyond that and look into more nuanced trials using various sources (software, waveform data, harmonics).

As ICP is related to many variables and needs to be considered in context, trainees showed a different approach to looking at ICP compared to staff physicians; trainees look at the numbers or thresholds first and then the trend or waveform. Staff physicians however care more about the waveform and trends, while numbers may be considered as alerts.

The variables correlating in context related to vital signs, medication, fluids and other neurological assessments (e.g., pupil size).

Trainees referred to the difficulty of neurocritical care concepts by tackling the vast content, skills and tools required, the mentality needed for it, transferring knowledge, detecting and predicting before (further) damage, data, equipment, and the consideration of the evolution of waveforms. The statement "the brain is like a black box" shows the overall complexity of this specific domain. Staff physicians added that individualization, pattern recognition, understanding the full picture using multiple streams, and the ability to express the findings to the team members play a key role. The ability to zoom out and look back, as well as taking limitations of equipment into account are further difficulties of neurocritical care concepts.

When we discussed key neurophysiologic concepts with our clinical collaborator, the three concepts of secondary insult dose, individualized thresholds, and trajectory were outlined. As this constitutes one expert perspective, further opinions and additions to these concepts were exchanged with more (neuro)critical care

experts. Additional concepts were pointed out such as taking a stronger focus on the waveform analysis (e.g., p1:p2 ratio), understanding ICP in context (also with other clinical indicators), thinking about autoregulation (relationship between ABP and ICP), trying to find the root or cause of the actual ICP elevation rather than just acting on it, being at different states of the disease process for all patients, and incorporating Bayesian thinking (e.g. probabilities and risks of trajectory).
These concepts are often perceived challenging to learn as novices require more exposure and connections from theoretical concepts to practical skills and sense-making of that in context.

CWA can be used to constitute visualization requirements from analyzing the complex work domain on multiple levels of abstractions and differences in how novices and experts behave differently in control tasks. Through the comparison between novices and experts, we have proposed ideas at which level can the EID support the novice to behave or start developing more in the direction of an expert. All these insights helped to gain distinct thoughts of physicians, combine them with observations and CWA models, as well as with existing neurophysiologic concepts. Initial ideas are outlined in the following subsections.
Visualization elements retrieved from the CWA models and observations relate strongly to the findings through the discussions; the ability to customize time settings, having a central screen showing different graphics, having notifications, showing various variables in context (correlation), and supporting the identification of thresholds and individualized approaches are major themes that came up.

Discussing technical feasibility is key and led to the development of static sketches.

4.6.1 Static sketches

As represented in the above table, there are many aspects that have shown up multiple times in different sections. Although the intention was to integrate as much relevant information as possible into the design of new visualizations for the ecological interface, there were parts with a bigger focus on, and technical limitations for the implementation of a potential prototype. When looking at the neurophysiologic concepts by Dr. McCredie and merging these with the feedback from the other critical care experts, some initial static visualizations were developed.

4.6.1.1 Preliminary visualizations

The initial ideas on the design of visualizations have been developed together with the co-op student Cathleen Grace Leone at the Advanced Interface Design Lab at the University of Waterloo. These ideas incorporated aspects of the previously described elements and led to four visualizations:

- an ICP waveform trend chart (Figure 9)
- visualization of connected waveform peaks to detect abnormal ICP (Figure 10)
- a summation trend chart and (Figure 11)
- an ICP dose histogram (Figure 12).

All visualizations will be described and shown in Figures 9 – 12 in the following.

ICP waveform trend chart: The results showed that it is important to understand and assess the waveform of certain variables such as ICP. The ICP waveform morphology can be described as pulse cycles having three peaks showing the percussion wave with its peak named P1, the tidal wave and the related peak described as P2 and the dicrotic wave representing the third peak P3. In a healthy state, the first peak would be the highest: P1>P2>P3. However, when intracranial compliance decreases, it is often seen that especially P2>P1. Thus, a lot of waveform assessment is related to comparing the first two peaks as a ratio. The left y-axis would represent mmHg and the x-axis time. The preliminary visualization shows trendlines that show each wave and peaks connected solely and represent an alert when lines cross another, such as in p2 becoming greater than p1.

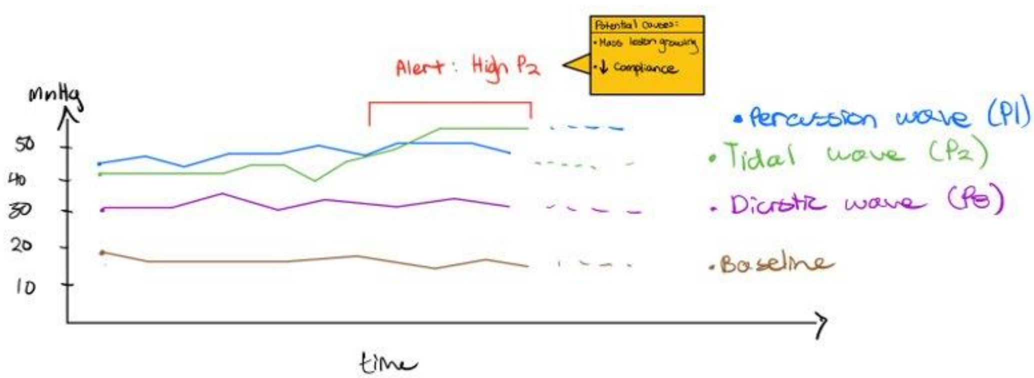


Figure 9: Preliminary visualization 1 - ICP waveform trend chart

Connected waveform peaks to detect abnormal ICP: Relating to the waveform representation again, this visualization would focus on one peak sequence solely. In a normal, i.e., healthy, condition, the bars connected on the left representation show the sequence of p1>p2>p3, whereas in an abnormal state, the line connecting the bars would especially highlight a non-conform sequence. The light grey area between the bars shows the expected values for each connected peak for ICP. The line highlights the abnormal state and visually alerts the user of the visualization that something is not going right.

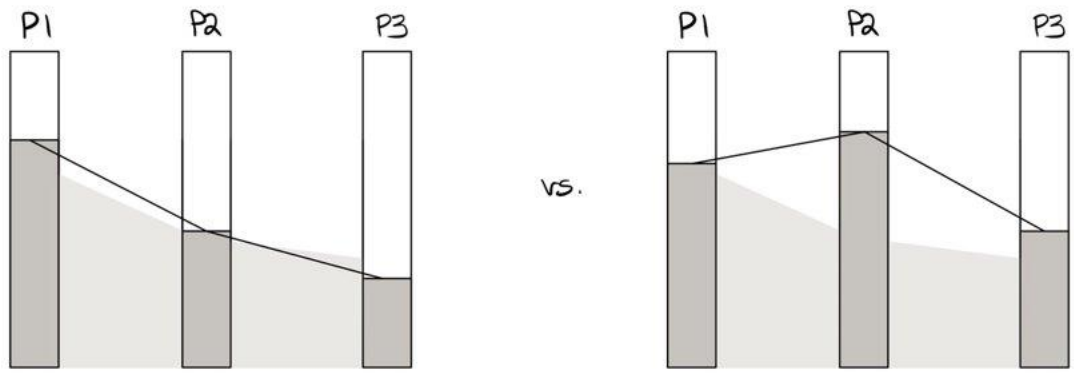


Figure 10: Preliminary visualization 2 - Connected waveform peaks to detect abnormal ICP

Summation trend chart: The third visualization represents a summation trend chart that shows three variables in context: ICP at the bottom, CPP on top and ABP as a line above those two variables. The left y-axis would represent mmHg and the x-axis time. As a neurophysiological connection between the three variables, it is known that ICP+CPP=MAP (mean arterial pressure or arterial blood pressure ABP). The intention of this graph is to show a correlation and dependency of variables in context and nudge the user to think about autoregulation. Often, autoregulation is determined by the correlation of these variables. This type of graphic is an example for multivariable displays often used in ecological interface design (Burns & Hajdukiewicz., 2004)

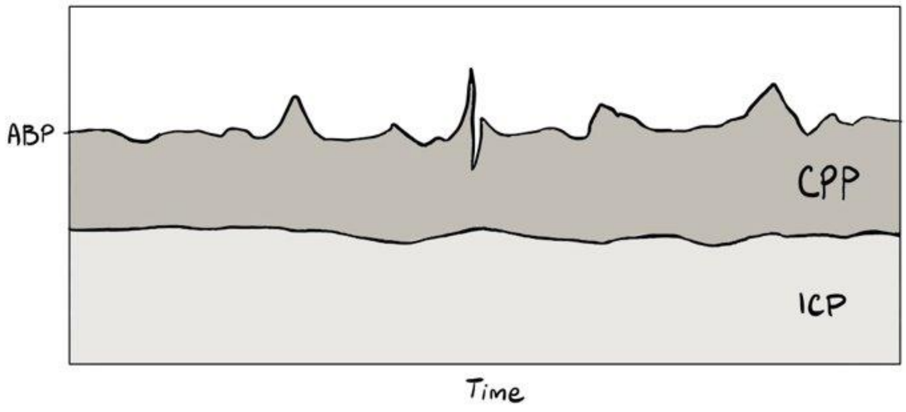


Figure 11: Preliminary visualization 3 - Summation trend chart

ICP dose histogram: In this graph, the axes are described in two ways: in the upper graph the time is represented over the past hour and the x-axis shows ICP in mmHg, whereas the lower graph shows the opposite.

Here, the intention is to represent bars in different intensities and colours, i.e., a darker red colour for higher values that would be alarming. This representation shows one point in time and is not stretched over multiple hours. It reflects the ICP intensity for a certain duration of time and might represent a current ICP value.

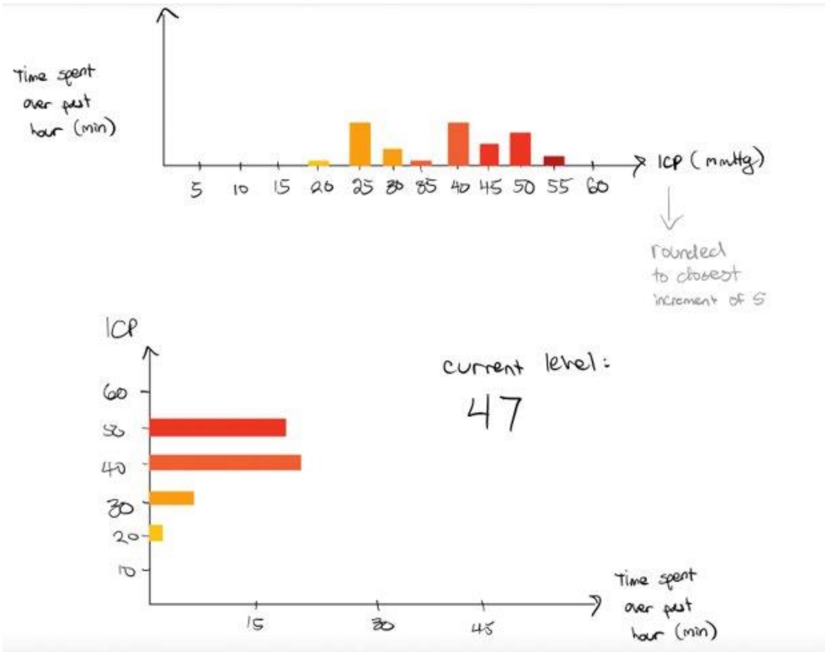


Figure 12: Preliminary visualization 4 - ICP dose histogram

Discussions about the feasibility of the graphics to be incorporated on a prototype interface highlighted some constraints and ideas for iterations or discontinuation.

Examples for discontinuation were the waveforms graphs for instance because the signal processing of such ICP measurements for each pulse incorporates challenging tasks to work around artifacts, characterizing the signals appropriately and coming up with an algorithm (Dai et al., 2020). This also applies to preliminary visualization 2 showing each peak as a bar for a single pulse sequence. The temporal representation over time would be missing and a model would need to be deployed to calculate various aspects such as the grey area or the lines and highlighted parts within each bar. For the scope of this work, it has been eliminated. However, the overall idea of showing a line graph with multiple variables and different colours has been maintained in the next iterations.

Preliminary graph 3 has been taken to the next level of static representation.

The ICP dose histogram (preliminary visualization 4) has been iterated into a more in-depth dose representation of ICP burden over time. Providing annotations on the same visualization (e.g., an indication in time when a medical intervention happened) might help understand potential causes for ICP elevation and may give hints on the different states of the disease process.

Further considerations have been taken towards the individualized thresholds approach and correlating annotations to track causes for ICP elevations and disease state as well. The individualized thresholds may also provide insights on the risks of trajectory; for example, if the ICP has passed the high threshold for a certain time or move in a certain pattern, it might lead to the considerations of risks for the next time interval.

Another idea coming up from the previous findings is to integrate a notification that might alert the user about dysfunctional autoregulation. This part has not been integrated on any visualization at this point but is considered for the full interface development view.

The next generation of the iteration process after reflecting on feasibility and applicability of more aspects from the above table, the static visualizations have been deepened and discussed with Dr. McCredie, and are represented as 4 major visualization concepts below.

Visualization 1: Secondary Insult Dose

This visualization (Figure 13) captures the ICP intensity levels represented over time. The left y-axis thus shows the ICP Time Burden as a percentage from 0-100%.

Each bar distributed over the x-axis shows a specific instance in time. The right y-axis describes an average value of ICP over the timeline as the black line in front of the bars. The different ICP levels are assigned to a colour: if ICP is below 15 mm Hg, it is represented as green. For a transitional range, i.e., 15-20 mmHg, the bar indicates a yellow part. For any ICP value above 20 mmHg, it is depicted as red on the bar. The user shall be able to recognize the different ICP values represented as normal, transitional and high ranges and understand the trajectory of the ICP values.

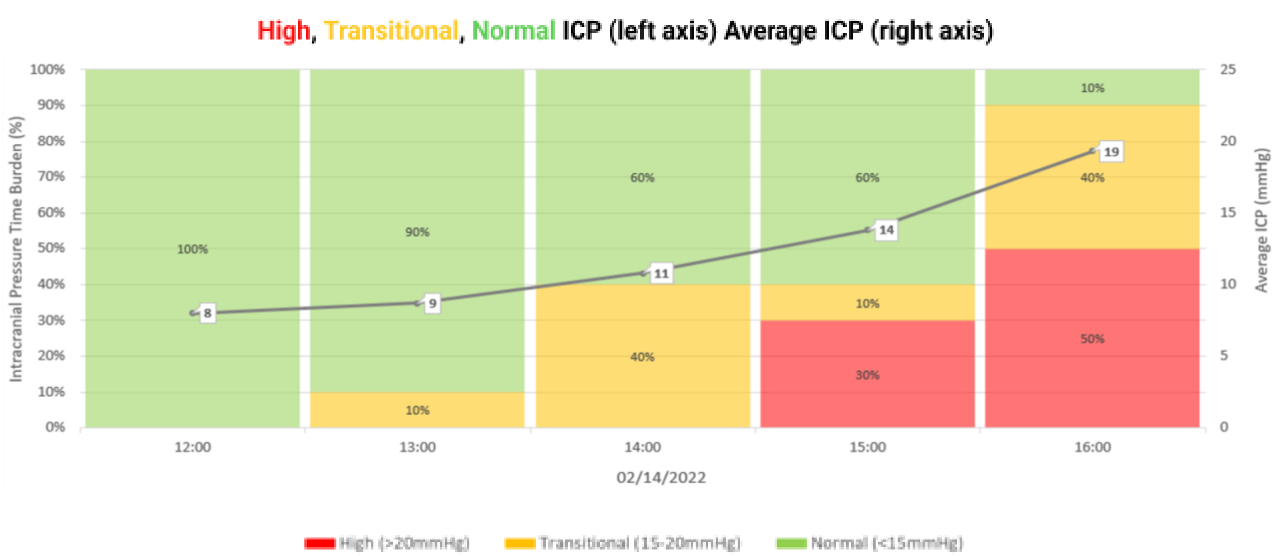


Figure 13: Secondary Insult Dose visualization

A slightly different representation of the above graphic is in the form of a donut visualization. The donut visualization has a similar concept of the normal, transitional and high ICP ranges; however, it shows a summary of last 24 hours and represents the average ICP in the middle of the donut. In its dynamic version, it would enable the selection of a timeframe. For the static sketch in Figure 14, this was explained to the interview participants.

Figure 14: Donut chart

Visualization 2: Individualized Thresholds

The next visualization in Figure 15 captures the measured ICP values on the y-axis and a timeframe on the x-axis. The dotted lines on top of the visualization represent the targeted thresholds for the ranges. A transitional range for ICP between 15-20 mmHg is indicated in yellow and a range above 20 mmHg is represented in red. The modification of the threshold targets is a useful indicator for keeping track of the patient’s trajectory. In some cases, these thresholds are required to be adjusted and the visualization will indicate when the threshold targets have been changed.

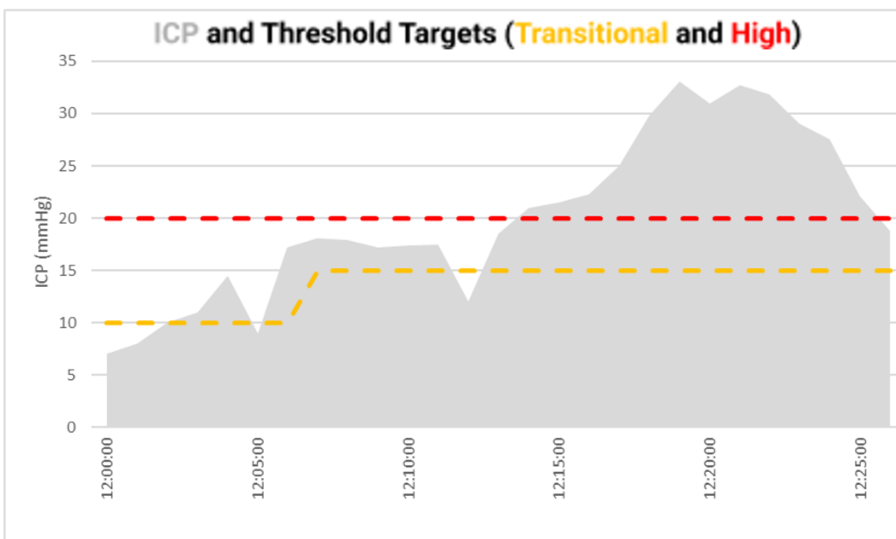


Figure 15: Individualized Thresholds visualization – version 1

Different variations of this concept have been visualized in Figure 15 and Figure 16. The ICP line itself can be shown as a black line with numbers associated to the peak and low values. The ranges may be shown in the background to quickly associate the trend in specific ranges.

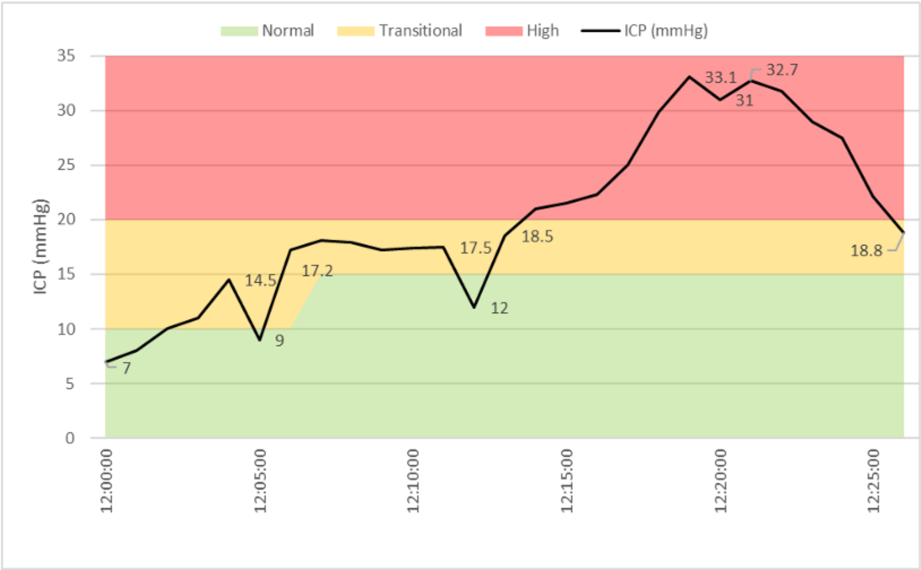


Figure 16: Individualized Thresholds visualization – version 2

Visualization 3: Trajectory

This visualization captures the concept of trajectory in different forms. The upper visualization as a gauge style shows a green and red area where pointing arrows can be seen. The arrow indicates a percent change rather than an ICP value. This approach shall visualize that, if the ICP value has increased by a certain percentage, the arrow shows the percentage in the red area, whereas a decrease would be shown in the green area. With a percentage change representation, a more individualized perspective is created. The lower representation as a trendline indicates an hourly and 24-hour ICP change with a triangle showing the tip up for an increase, or the tip down for a decrease. The visualizations in Figure 17 shall support the clinician to be more aware of the trajectory of an individual patient.

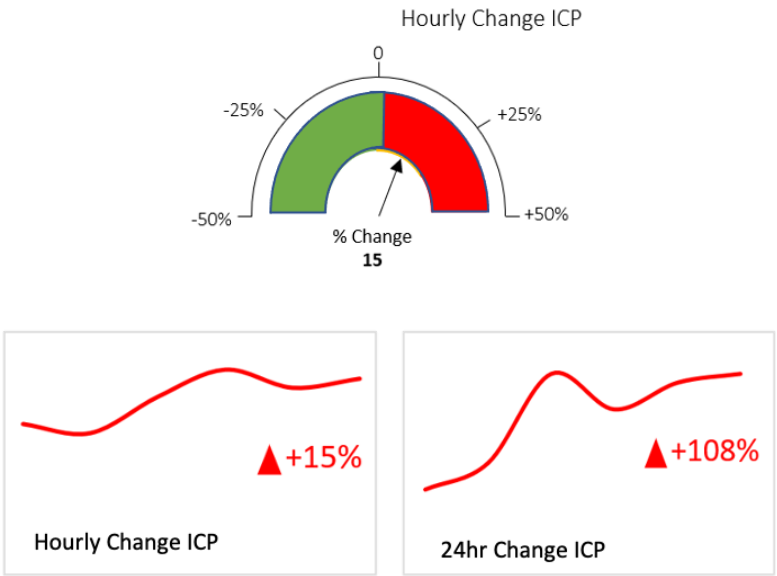


Figure 17: Trajectory visualization

Visualization 4: Summation graph of ICP, CPP and MAP

A core idea of EID is to represent variables that are interrelated with each other. The visualization below shows one of the core physiologic relationships between the variables ICP, CPP (cerebral perfusion pressure) and MAP (mean arterial pressure) in critical care: ICP+CPP=MAP. The summation graph in Figure 18 shows CPP as an area at the bottom, ICP stacked on top and MAP as a black line above both.

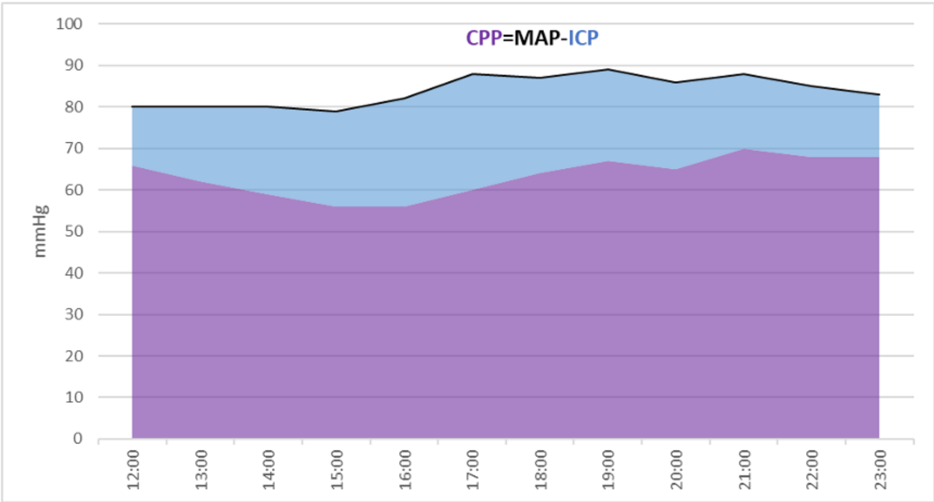


Figure 18: Summation graph visualization

Different variations of the graph are represented below, e.g., in Figure 18 and Figure 19. The line graph below shows each variable as a separate line within the same graph in different colours. In this representation, the increase and decrease of each line is distinctly shown.

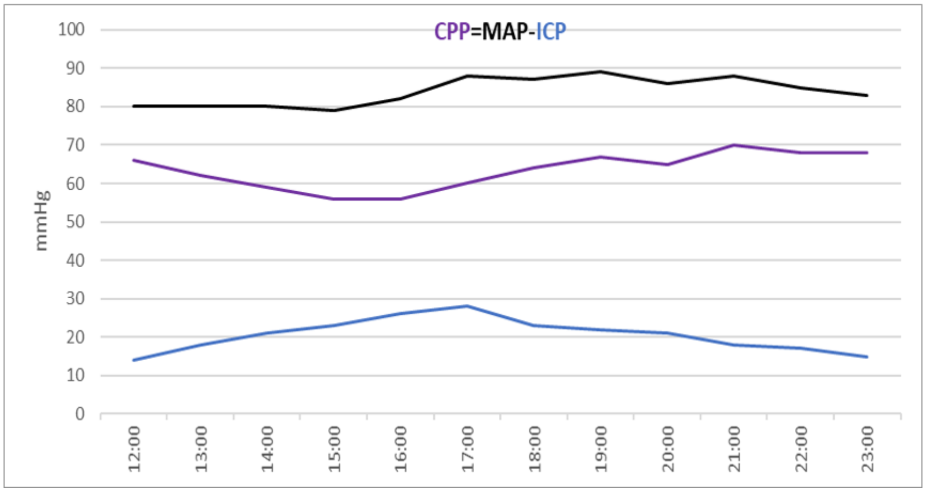


Figure 19: ICP, CPP, MAP line graph

Another version below (Figure 20) shows the differential of ICP in an additional colour such as red for an increase in ICP and green for a decrease in ICP. This way, the idea of the summation graph is still present and the differential of ICP is explicitly shown.

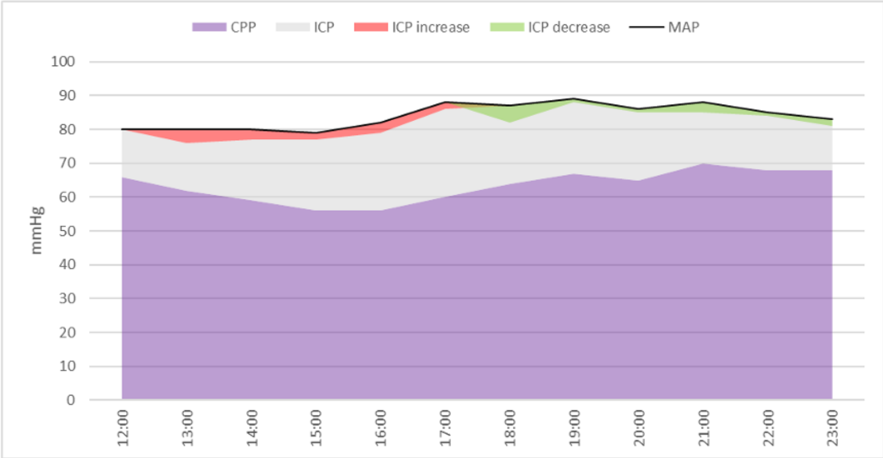


Figure 20: Summation graph with an increase and decrease indication of ICP

All the above-described visualizations shall foster awareness for the main neurophysiologic concepts outlined above to support expertise development in neurocritical care. The next chapter aims to further explore its use, possible limitations and areas for improvement with trainees' and staff physicians' feedback.

Chapter 5 Perceptions of novices and experts on interface design ideas for bedside physiologic monitors used in neurocritical care

5.1 Introduction

Little is known about visualizations' impact on clinicians' decision-making process in critical care.
This chapter aims to show the perceptions of trainees and staff physicians on the developed static visualizations described in the previous chapter. Both participant groups have been identified to provide valuable feedback; the experts' (i.e., staff physicians) view contributes to what is needed on visualizations to support expertise development in terms of their reasoning and shaping of mental models in neurocritical care. The trainees' (i.e., novices) perceptions are important to better understand how they are thinking and interpreting the new visualizations, which ultimately affects their plan of action.

The research questions we try to answer in this chapter are as follows:

- How are novices and experts perceiving the static visualizations with regards to usability aspects, and how do experts think about the visual representations of the neurocritical care concepts on an interface?

To detail out both perspectives, we keep these subquestions in mind:

  - From the perspective of novices, to which extent are the visualizations easy to understand and useful in context? What can be improved or shown differently?
  - How are experts thinking about the relation of the concepts to the presented static visualizations? What major concerns do they perceive that require improvement and what aspects can stay in its shown representation?

Gathering all input from the physicians,

  - how can we improve the design of the visualizations overall?

These insights may show us a potential flow of the visualizations represented together on an interface. The flow is important to consider as the mental model development will be supported through the various links of the shown visualizations. Early ideas will be explored with expert critical care physicians.

All these steps shall help to constitute a basis for the ecological interface and incorporate ideas for the next planned usability study. The end of this chapter will thus demonstrate the final design for the visualizations as well as the preparation process for the usability study to assess the ecological interface.

5.2 Methods

A semi-structured interview study with an inductive analysis approach was conducted to get an in-depth understanding of how physicians think in neurocritical care. This study is part of the 60-minute interview that was outlined in previous chapters, whereas the feedback elicitation part took around 30 minutes per participant. Due to time limitations, some visualizations were skipped with some participants, but the focus was to get main feedback points that could be related to the overall idea of the visualization. Usually, the order of the visualizations presented for feedback were:

1) The ICP dose visualization (bar and donut chart)
2) Individualized thresholds visualization
3) Trajectory visualization
4) Summation trend chart visualization

The approach to start the feedback phase varied amongst trainees and staff physicians. Trainees were asked to explain the visualizations to both interviewers (the author of this dissertation and Kathleen Schaef) in as much detail as possible. In case of missing descriptions, the interviewers probed the participants by pointing out towards the missing information to detect whether they forgot about sharing the meaning or whether it is related to lack of understanding of those aspects. After trainees' explanations, the interviewers shared the purpose and explanation of the visualizations for completeness and providing rationale on how we came up with such visualizations.

For staff physicians, the interviewers provided a short explanation directly (i.e., they were not asked to describe the visualizations to the interviewers). Experts then had the opportunity to (dis)agree on the way of how the concept is represented through a visualization. They could also ask questions if anything was unclear.

The discussion about the visualization with both trainees and staff physicians often related to preferences of information representation such as numeric values, timelines, colours etc., but also confusing aspects or potential ideas for improvement. Notes have been taken simultaneously to keep track of the key findings.

These discussions took place through the virtual platform Microsoft Teams and have been recorded and transcribed through Microsoft Teams.

All participants were offered remuneration or gift cards if that was a preference. The data was de-identified and stored on a password-protected OneDrive account.

After transcription and relistening to the interviews with the participants, short summaries have been created and sent back to the participants via email. This offered the opportunity for the interviewers to ask about any unclear parts in the transcripts, and changes or additions to anything to their statements.

Further, feedback and clarifications on misinterpretations could thus be provided.

An inductive thematic analysis (as described in the previous chapters) has been performed by both Kathleen Schaef and the author of this dissertation. Most trainee participants had run the study before the staff physicians were scheduled. Common themes and sub-themes that were first denoted through the responses of trainees were later brought up and discussed with staff physicians to identify improvements in design.

This study received ethics clearance from the University of Waterloo Research Ethics Board (#42892).

5.2.1 Recruitment of participants

Critical care physicians have been recruited for this study. Trainees have been contacted through an email that Dr. McCredie forwarded to the University Health Network (UHN) program directors within Canada. In some cases, Dr. McCredie reached out to experts in the field via email directly to support the recruitment process. All participants received an invitation to sign up for the study through our online calendar on Calendly® including all relevant information about the background and intentions of the study.

A total of 36 participants (18 trainees and 18 staff physicians) were recruited for the interviews. The details of the demographic information can be found in section 2.4.4 within Table 1 and Table 2., This part represents a continuation of the previously described interview with the same participants.

5.3 Results

This part shows the feedback from both groups (trainees and staff physicians) on the presented static visualizations. The major themes and subthemes identified for each visualization are summarized in Tables 5-8.

5.3.1 ICP dose

Table 5 portrays a summary of the themes and subthemes created through the codes of trainees and staff physicians. The bolded descriptions came up in multiple visualizations and the starred subthemes represent that staff mentioned the same subthemes as trainees.

Table 5: ICP dose themes and subthemes

Visualization	Theme	Subtheme- Trainees	Subtheme-Staff
ICP Dose	User Confusion	- Interpretation of ICP burdens and percentages - Wording of y-axis - Two axes in context with graph description	
	Visualization Application	- Trend - Complimentary	- Trend* - Complimentary* - **Different Users**
	Distribution of ICP Values	- Average ICP - Fluctuations in ICP	- Average* - Fluctuations in ICP* - Duration of ICP Events
	Impact on Management	- **Transitional Threshold**	
	Usability	- Intuitiveness - **Desire for Simplicity** - **Collation of Information**	- Desire for simplicity* - Differentiation of higher ICP values - Collation of Information* - **Intuitiveness***

			- Two axes in context with graph description* - Wording of y-axis*
	Visualization Interaction	- **Time Scaling** - **Target Customization**	- **Time Scaling*** - **Target Customization***

User confusion – theme:
As trainees have been asked about their understanding of the visualization, there were some subthemes within user confusion that related to confusing aspects about how to interpret the ICP burden, the portrayed percentages on the bars, the wording of the y-axis, and the axes in context with the graph description.

- Interpretation of ICP burden and percentages. One participant was confused about the representation of a percentage on the bar: "100% of the hour from 12 to 1 (12:00-13:00) they were just at 8 mmHg" (T18).
- Wording of y-axis. "The % of time spent at ICP intensity, what does that mean? […] OK so I think it means that, when the ICP is in the normal ranges, we look at it less" (T13).
- Two axes in context with graph description. "That kind of threw me a little bit because on the right-hand side on the y-axis, the higher the number, the higher the ICP so you are kind of processing two directions at the same time. So, I think I kind of expected the high to be mapping with the high." (T4)

Visualization application – theme:
In the visualization application theme, trainees commented on the relevance of trends:
"In the ICU we are so used to a temporal flowsheet, everything is hour by hour. […] These type of things (referring to the donut visualization) that take away the temporal x axis are interesting, but they don't add to my clinical decision making. It is almost like a factoid that is just there to look at for data visualization, but it is not something I would miss if it wasn't there. […] My brain is trained to think in a temporal axis for all physiological parameters" (T4).
Staff physicians commented that the bar chart especially gives indication of the trend and deterioration of the patient, whereas the donut chart it is difficult to appreciate when the patient was high in the last 24 hours (S4).

Staff physicians also added that different users may use the visualization: "Different level and types of providers may benefit from different types of presented information" (S2)

As both bar graph and donut chart were shown, trainees as well as staff physicians mentioned they might serve different purposes but could also be complimentary:
"The first one (bar chart) would be more acute changes over a shorter period of time and then you can see whether a certain intervention worked or didn't work with that, whereas this one (donut) would be how the patient was doing over the last few days." (T9)
One staff physician mentioned that the donut is applicable for a global summary but not useful for specific management, they both are valuable and complementary (S18).

Distribution of ICP values – theme:

The theme distribution of ICP values related to comments about the average ICP value but also about representing fluctuations: “A mean ICP may not necessarily convey all of the nuances that may need to be addressed.” (T1)

Other comments related to the interest about spikes (T14, T9, T1), and the frequency of spikes (T8).

Staff physicians had various perspectives about the average ICP value either being a useful parameter (S11) but also stating that it wouldn’t show all the nuances (S7), and highlighting min/max values or the variability would also play an important role (S8).

Also, staff physicians added their opinions on the duration of ICP events; the peak ICP is not as important without knowing the duration because often, there will be very high peaks for a short period of time (S18), and it is important to know whether it was sustained because there might be different reasons for ICP peaks (e.g., coughing) (S8).

Impact on Management – theme:

Some trainees talked about the impact on management, relating to the transitional threshold: seeing 40% yellow would not impact their management (T8).

Usability – theme:

Usability has come up as a theme in various contexts, especially when trainees talked about intuitiveness (i.e. easy to understand and use) of the visualization: “I love the colors, and I love the percentage how it is presented because it is very visual, and you can see the moment when it begins to be dangerous” (T7), however, some trainees also verbalized insecurity of how to use the graph: "I am still struggling a bit with understanding how to translate it and use it appropriately." (T10)

Staff physicians mentioned that the colours are intuitive for both novices and experts (S16) and is international because it reminds of the traffic light colours (S10).

One participant also highlighted that it would be more beneficial if it was simpler with less color or less numbers (T5). A staff physician added that novices often don’t have the pattern recognition required to make use of more complex visualizations and that statistical indicators (e.g. showing a range band) might not be useful: “I think in terms of visualizations at the bedside, the more expertise it takes to interpret a visualization, which is inherently something you are trying to build to simplify things, is a mistake." (S10) (desire for simplicity-subtheme). A good balance between the quantity of data and simplicity of the interface is important (S5). Adding more colours could be perceived more confusing for trainees (S16).

Also "It is very useful. It triggers the alarm in your head in multiple ways. Not just looking at the ICP number the amount of time spent which is a huge concept" (T4) (collation of information-subtheme).

Some staff physicians added that the representation of a patient with high ICP has a wide range: "The patients I am concerned about are above 15mmHg the whole time. […] I don't know if this graph is going to be very discriminatory" (S7). Others suggested having another colour category of a darker red for ICP values greater than 25 or 30 mmHg to highlight the ranges more clearly (Differentiation of high ICP-subtheme).

One staff physician also noted that this visualization helps show that it is not only about the number but for how long ICP remained in certain ranges, saying it is better to have an ICP of 44 mmHg for 5 minutes than an ICP of 28 mmHg for 2 hours (S6) (collation of information-subtheme).

Some also commented on the two axes in context with graph description especially about the order of representation of the colour-stacking. High ICP could be presented on the bottom instead of normal ICP (S10).

Feedback was also provided on the y-axis wording. For some staff physicians, it led to confusion "not clear what secondary insult dose means and what intracranial pressure time burden is" (S17) while others showed preference for ICP time-burden (S11, S3, S10) or written out as Intracranial Pressure time burden (S9, S4). One staff physician commented that novices may not know what burden means (S2).

Visualization interaction – theme:
Another main theme coming up for the ICP dose visualizations is visualization interaction. A comment about the time-scaling aspect is: "If a patient's condition changes, it usually changes on an hour or minute by minute basis and so I would want to be able to zoom in even closer." (T4)
Some staff physicians commented that it would be useful to graduate to different times (S14, S1) but might also risk data overload (S1).
Target customization was deemed important: "[…] maybe a certain ICP in one patient to the next may physiologically mean different things depending on, what other information you have about the patient" (T1).
Also, staff physicians agreed that customization of the targets would be useful (S18) but it would be helpful if it can be done only by few people on the team (S2), "for example, for certain patients we talked to the neurosurgeons, and they tell us we don't want the ICP above 12 in this patient. For example, in any other patient, we wouldn't even bat an eyelash over, but for this particular patient there might be some pathology that's particular that we need to consider" (S1).

5.3.1.1 Impact on design
To incorporate as much of the feedback as possible, different approaches have been considered to overcome potential challenges and preparations for the next version of the interface as well as for the usability study. Aspects of user confusion can be addressed by providing appropriate training with detailed explanations of potential confusion. During the interviews with trainees, the explanations provided after their initial descriptions to the interviewers, a detailed explanation was nevertheless given to the trainees that did not raise any further questions.

After the discussions with the staff physicians about the y-axis wording, the decision was made to note "ICP Time Burden" as the appropriate description. To prevent any misunderstandings about burden, this word is also explained in the training.
The two axes on the bar graph together with the stacking order of ICP intensities was another point of confusion and discussion; the decision was made to keep the high ICP values at the bottom as it is easier to appreciate a rising (and thus more concerning) trend of high values in ICP. An important point for the training is to clearly describe the ranges of ICP with the related colour meanings. To customize the graph and provide more variations in users' interest in looking at certain ranges or axes only, a turn on/off option was included for all colours and the axis on the right-hand side shows the average ICP value.

The comments about including information on the peak values made us decide to include a hover interaction on each bar that would show the max values.

The duration of ICP events was considered important too, so a click-function on each hour of interest was provided which would lead the user to the individualized threshold and summation trend chart visualizations for more detailed views.
The fact that some patients may have ICP values in high ranges and very high ranges suggested adding another intensity level, i.e., dark red, for values above 30 mmHg.

Although feedback was given about complimentary aspects of the ICP dose bar vs. donut chart representation, we decided to add the idea of donut type of chart into the very initial view of a bedside physiologic interface, i.e., the vital signs representation. As the donut chart provides a certain alert to the user, we have been considering a more consistent representation with the bar chart, i.e., a single bar showing the different intensities for the selected time including percentages of how much time the patient spent in those ranges.

In the training we further mention that only certain physicians may be able to customize the targets.

5.3.2 Individualized thresholds

The main themes and subthemes are summarized in Table 6 below.

Table 6: Individualized thresholds

Visualization	Theme	Subtheme- Trainees	Subtheme-Staff
Individualized Thresholds	User Confusion	- Detection of threshold change - Interpretation of the colours of ICP intensity ranges	
	Content Feedback	- Annotations	- **Annotations***
	Design Preferences	- Familiarity - Ease of Understanding - Salience	- Consider what is being displayed together - Area under the curve - **Salience*** - Consistency - Coloured line
	Impact on Management	- **Transitional Threshold**	
	Usability	- **Desire for simplicity** - **Collation of Information**	
	Visualization Interaction	- **Target Customization** - **Exact Values**	- **Target Customization***

			- **Exact Values***
	Visualization Application		- **Different users** - **Integration with other visualizations***

User Confusion – theme:
Detecting the threshold change was perceived as difficult (T17). Another confusing aspect was mentioned about the interpretation of the colours of ICP intensity ranges: "I don't know what the coloured part is" (T15).

Content Feedback – theme:
One subtheme was identified referring to annotations.
Trainees commented that having annotations would be helpful especially if a patient is stimulated or given a treatment (T2).
Staff physicians mentioned multiple times it would be useful to have e.g., interventions or medication denoted within the system "important to help visualize the effect of our intervention and confirm the impression which sometimes may be wrong" (S7). This would further be useful "especially in training" to see the feedback of what the impact of various treatments are (S3). A filter for the events by category could be an additional interaction of value (S10).

Design Preferences – theme:
One trainee mentioned that the visualization is familiar to what they are used to seeing in the hospital (T18) and similarly, that the dashed lines would be familiar too (T5) (familiarity-subtheme).
In terms of ease of understanding, two trainees expressed their preference for the colour blocked visualization as more intuitive and easier to understand and interpret (T1, T2), while another trainee found dashed lines easier (T5).

Seven trainees preferred the colored blocked visualization as it is "more obvious from the first look" (T3) (salience-subtheme).

Some comments by staff physicians related to the representation of the lines (consider what is being displayed together – subtheme). One comment related to showing the dashed line with targets which would be less likely to be confused with the bar graph (S8).

Few comments related to the area under the curve. (S3) commented that his preference would be to see the area under the curve visualization where the total area colour would correspond to a specific ICP range.

Another comment referred to salience aspects; it would be better to keep the visualization quiet until there is a problem (S3). One idea provided by the same staff physician was to color the line for transitional and then color the entire area for high ICP. He mentioned it would be good to draw the attention but not overload the user.

Keeping colours consistent (consistency-subtheme) has been noted as an important aspect; it would be useful to have the same colours for the bar chart (S16).

Two staff physicians also mentioned their preference for coloured line representations (S5, S9) (coloured line-subtheme).

Impact on Management – theme:
Trainees highlighted different aspects about the transitional threshold: "[…] green or yellow it's the same for me. For 18, 15, 10mmHg (ICP) it is the same treatment." (T8). Whereas another trainee mentioned he doesn't care about ICP being 17 or 18 mmHg but if they change to different ranges (T6).

Usability – theme:
Trainees had different impressions about the usability of the visualization and commented on the desire for simplicity "this graph is very nice, informative and simple." (T3)
Others, however, mentioned they don't need the colours: "I can see the curve, why would I have a need for the green and the red and the yellow. I know what the thresholds are so why should I have these colors" (T6).

Another comment by a trainee provided valuable insight into how collation of information plays a role in design: "The current way of doing it is very active meaning that I would have to look at the paper (order sheet) and I would have to look at the patients ICP and then I would have to synthesize the two. This takes that work away and all I have to think about is whether they are in the zone or not so it takes some of that mental load away." (T4)

Visualization Interaction – theme:
Several comments were given about target customizations: "as somebody who is a proponent of personalized medicine, I would say yes, because we all know that individual patients do differ, and their physiology may differ. And so, these things may be important if tailored to the patient. So, I would say yes, that would be helpful. Because maybe a certain ICP in one patient to the next may physiologically mean different things depending on what other information you have about the patient." (T1)
However, another trainee mentioned "if you adjust it, it needs to be shown somewhere and everybody needs to know the rationale behind why we adjust it." (T9)

Staff physicians also saw value in having the thresholds customizable (S11).

Several opinions came around including the exact values on the graph; some preferred having a min and max indication on the graph (T8), others preferred a hover over function (T18), while another commented it is not necessary to include the number on the graph (T3).
Staff physicians gave feedback that the ability to read the exact values on the ICP graph would be helpful (S8, S16, S7), while another staff physician mentioned that it is possible to read the value directly from the y-axis (S12).

Visualization Application – theme:
The users of the interface vary; thus, having different users plays a role in how they see and interpret the data: "Doctors are primarily interested in diagnosis and nurses are primarily interested in care. So, a nurse might like one sort of way of looking at data because they are interested in minute-to-minute care, where doctors are going to be interested in pattern recognition for diagnosis." (S11)

Also, in conjunction with the dose visualization, the individualized thresholds graph seems to address some concerns raised with the dose graph (S11) (integration with other visualizations-subtheme).

5.3.2.1 Impact on design

Although confusion appeared around the colour representation, this was easily solved when explaining the visualization to the trainee.

Many participants commented that individualizing the care of patients is important and thus, it is key to appreciate the nuances of care and information presented on the interface. Having annotations would help this process of adding the ability to see the response to interventions and care. A filter option about the type of interventions will be included in the design.

Feedback was provided on preferred colour and shape especially for making the connection between various aspects that might appear across multiple visualizations. We thus decided to keep the same colours but keep the ranges quiet, while colouring the lines for different ranges and having the line dotted to not overload the user of the system. The ICP dose visualization does not portray the potentially customized threshold targets and thus, these individualized thresholds graph adds value and consistency. The effectiveness of showing various versions of this graph can be tested in future research as there is potential to further investigate its use with various stakeholders, different experience levels, in different combinations of visualizations etc. The ability to customize the thresholds would be provided to certain staff physicians who are main decision-makers for the patient treatment. This would be included in the training for the usability study.

5.3.3 Trajectory

For the trajectory visualization, four themes have been discovered through the discussions with trainees and staff physicians as seen in Table 7. Trainees commented more on user confusion aspects, visualization application and interaction, while staff physicians shared their thoughts on usability.

Table 7: Trajectory themes and subthemes

Visualization	Theme	Subtheme- Trainees	Subtheme-Staff
Trajectory	User Confusion	- Interpretation of percent change - Implication of percent change - Trendline Visualization - Gauge Visualization	
	Visualization Application	- Training - **Integration with other visualizations** - Alert	- **Alert*** - Presence of absence of evidence - Changes above threshold - **Integration with other visualizations***

			- Promotes looking beyond threshold
	Visualization Interaction	- **Time Scaling** - **Exact Values**	- **Exact Values*** - **Time Scaling***
	Usability		- Interpretation of percent change* - Presence or absence of evidence - Implication of percent change* - Redundant

User confusion – theme:
One trainee faced challenges with the percent change interpretation; "it probably means that ICP has increased around 15% every hour. " (T2)
Also, when it comes to the implication of the percent change, the trainee commented: "I find this visualization relative, so I don't know the actual number of ICP. I don't know what to do with the data. I know it's probably pretty bad, but I mean if someone comes from the ICP of 5 and then the other day is 10 (100%), it's probably not that bad though, but maybe had to keep it mind why it increased so high, but it's still not reached the threshold...the relative change is a big more difficult to understand for me." (T2)
There were multiple comments by trainees stating they would be unsure how to use the percent change but might provide research value (T12).
Further confusion was raised about what the gauge style and trendline visualizations represent (T4, T15).

Visualization Application – theme:
One trainee found use in the visualization's application from a training perspective (T7).

Further, trainees thought that this visualization could be used integrated with other visualizations such as the individualized thresholds visualization (T18, S8, S3, S16).

Concerns were also raised about how this visualization would not guide intervention: "useful as a snapshot, useful as an overview, not useful to guide intervention." (T18)
However, there was value in using this visualization as an alert:
"To me this seems analogous to the donut in that it is an alert and indicator for me to be primed that on two different time scales things are heading in the wrong direction." (T4)
Also, when relating to novices in critical care, one trainee highlighted that less experienced physicians tend to only look at the number and having this type of visualization would act as a cue to pay attention to the trend. He believed that a larger time frame of anywhere from 6 hours to 25 hours would be more valuable for this percent change than hourly (T11). Others commented as they gain more experience and have a better idea of what the percentage of concern may be (threshold percentage), that this will be even more beneficial. He also said that this is especially helpful when a patient is below a threshold but has a significant percentage change that this could help him be more active and monitor the patient more closely (T10). "It is giving me a lot of information of what's happening now and showing me the trend of what's been happening in a very short period of time." (T17)

Staff physicians also added that it would help to get the attention for patients in an acute stage (S16) and the aim is to "increase our ability to identify trends and synthesize information and decrease our cognitive load and everything is flashing, it's a bit of a problem." (S3)

Staff physicians related to the changes above threshold since "a lot of neuro is threshold driven, not that we always know what if the threshold for that specific patient is" (S7) but also "If somebody went from an average of 4 to an average of 10, I don't know whether I would think the same if somebody went from an average of 18 to an average of 24. So, there is an absolute amount which is the same and it's a smaller percentage." He further explained how the patient that went from 18-24 is more worrisome even though the percent change is less. (S11)

However, there were others that saw the impact of this visualization on promoting to look beyond the threshold: "When I think about it conceptually… I'll pull back and look at the 24-hour time scale and I do sort of look at an overall trajectory but it's very dependent on the context." (S18)
He also mentioned that many circadian events happen in the ICU such as morning exams, bathroom, etc. and that the percent change could be useful to check from the exact same hour the prior day, "this might help with some of the more routine context that we end up having to deal with." (S18)
(S2) added that usually, novices often think about absolute numbers and are resistant to consider management without a numerical and threshold approach.

Also from an application perspective, many commented about the presence or absence of evidence and opinions varied;
"So, I think because there's not a lot of empirical evidence about what those numbers mean, all I'm really caring about is when it's crossing the threshold and how many times it has crossed the threshold, how much time you spend above that threshold which you've done already with the others (referencing the other visualizations)" (S10).
Another staff physician commented on quantification:
"I am not sure the benefit of quantifying it, unless the quantification is associated with specific thresholds. […] If percent change is anchored or associated with a diagnostic state, or if the percent change is associated with a sign of treatment response or that treatment is failing, then in that case it's important to quantify." (S17)

However, others noticed value in having a different way of thinking through this visualization:
"I can understand why a trainee may not care (referencing percent change visualization), but for me, I care. I may not know what to do with it but I am going to put pressure on science to tell me what to do with it." (S8)
Also, "as people become accustomed to looking at this information, they may start to make sense of it and use the information to help inform decisions. […] the fact is you do need it, you just don't realize you need it so some of this is sneaking in user-friendly, intuitive interfaces that are some type of analytics into people's consciences of what they get used to seeing, and then they realize it's valuable." (S2)

Visualization Interaction – theme:
Several trainees mentioned they would like to see exact values such as starting and ending values on the trendline; "I don't know if the waveforms on the bottom really help me because there is no scale." (T4)

There were further comments about the time scaling, one trainee said that a 6-hour view would be most useful (T6) or a 24-hour view (T17).

Staff physicians provided similar feedback about having start and end points with a potential addition of a max value over the time frame on the trendline (S12).
Time scaling was similarly recommended as 6 hours (S2), 4 hour (S5) or a customizable time frame (S1).

Usability – theme:
Staff physicians highlighted various aspects about the usability of this type of visualization. There were comments about the interpretation of percent change that relate to the way of thinking:
"I think percent change does provide more information, but it's a matter of thinking differently" (S12), or "I really struggle with percent change because the phenomenon is not linear and there is an implicit sort of implication in the concept that your denominator is either constant or is really relevant because it is a fraction." (S11)
(S6) also added about the implication of the percent change: "there are physiological changes in ICP that can be related to many physiological reasons that don't necessarily have a pathological meaning".

Some concerns were raised about the presence or absence of evidence with a percent change indicator: "I am not going to make any management decisions over this." (S10)
It could also be used for a reflective approach: "Often because the percent change may be reflective of something else and not in and of itself important." (S18)
Multiple staff physicians commented that it is possible to get the main information through other visualizations and were thus summarized in the subtheme redundant.

5.3.3.1 Impact on Design
Aspects of the trajectory visualization were perceived as a useful addition to other visualizations. Often, feedback was given about it supporting as an alert for helping the patient in an acute phase, even if the percent change is still below a concerning threshold. Some staff physicians saw more value in the use of a percent change indicator when the patient has crossed a concerning threshold, while others found it useful even before passing a certain threshold. Minimum, maximum and average ICP values will be portrayed next to the percent change so a quick link can be made at one glance.
Although the visualization may not appear very intuitive at the start, especially to trainees, some confusions were verbalized about the interpretation which can be overcome through an accurate explanation during the training.

Different opinions were shared on gauge or trendline preferences to represent the percent change but the trendline in context has been outlined as more useful since it is easily linked to the overall trend of the patient trajectory.

5.3.4 Summation trend chart
Similar themes and subthemes have been identified in the discussion with participants. Staff physicians additionally commented on usability and design preference aspects as seen in Table 8 below.

Table 8: Summation trend chart themes and subthemes

Visualization	Theme	Subtheme- Trainees	Subtheme-Staff
Summation Trend Chart	User Confusion	- Challenges identifying individual variables - Legend	
	Content Feedback	- Importance of CPP targets	- Importance of CPP Targets*
	Visualization Application	- Relationship of variables - **Integration with other visualizations**	
	Visualization Interaction	- **Time Scaling** - **Exact values**	- Autoregulation - Relationship of variables* - Exact values
	Usability		- Challenges identifying individual variables*
	Design preferences		- Line graph - Advanced Summation trend chart - Summation trend chart

User confusion - theme:
Multiple trainees faced challenges identifying individual variables on the summation graph stating, “It is pretty hard to tell if the ICP has changed over time or not with this graph” (T2). Feedback was provided on the representation of the stacked variables, such as having three lines instead for each variable which would be easier to appreciate for increases and decreases. One trainee further added that it is “less intuitive if you are not picking up on the colours” (T1) as there is no legend, referring that it is not very intuitive to link the colours of the equation to the graphical representation.

Content feedback – theme:
Both trainees and staff physicians shared their thoughts on the importance of CPP targets; some physicians stated that they prefer to see CPP more accurately: “CPP is what I am looking for and what I am targeting” (T3).

Staff physicians discussed having a target for CPP (S10).

Visualization application – theme:
The relationship of variables is a core indicator for providing patient care: "we may do things to modulate CPP and MAP as opposed to just ICP." (T1)

Another trainee commented that she would look at ICP and CPP in different situations “If ICP is normal, I don’t worry much about CPP but if ICP is abnormal then I would look at CPP to see how high I need to push the MAP up to get a good CPP.” (T9)

It was also noted that this visualization would be complimentary to other visualizations (T10).

Visualization interaction – theme:

To appreciate changes in the variables, feedback was given on having different time scaling of the x-axis (T17).

Further, having a hover over feature (T5) would be useful to see exact values: “I would like to have the ability to put the pointer and show me the values as I moved through the graphic.” (T7)

Staff physicians mentioned similar functionality that would help to better detect changes of ICP and CPP.

Usability – theme:

Staff physicians commented on challenges identifying individual variables and changes in the context of usability; (S3) highlighted that it would be especially difficult to see what is going on during emergency situations as the changes in the variables are hard to appreciate.

Design preferences – theme:

In terms of design, some staff physicians preferred to have a line graph instead: "if it gets too complicated it is not going to be used." (S15)

It was also seen to be easier for early learners as it is more common and configurable more easily (S12).

When comparing different versions of the summation trend chart, two staff physicians said they prefer the summation graph (S5, S9). In more advanced design iterations of this visualization, an increase and decrease representation was added: “The bottom one [referring to visualization with increase and decrease] has a lot more information, even if it is a little more complex.” (S18)

This staff physician recommended to also include the increase and decrease representation for CPP. It was further mentioned that this representation provides information on autoregulation "ICP is increasing and CPP is dropping which would suggest to me that I don't have good autoregulation because CPP is not maintaining where it should be." (S12)

Considering this advanced trend chart, it was commented also from an expertise development perspective "I think it would be a little harder to grasp and to teach because it is not the way that we've done things in the past, but it doesn't mean it isn't what we need to think of for the future. […] I think the straight lines are easier and more basic, but I really like the next one (summation with red/green area) for some kind of more in depth knowledge" (S12)

Similarly, another staff physician added "When you start doing things like that and forcing yourself to look at your instrument (referring to more complex visualizations), you can get more out of it" (S11).

5.3.4.1 Impact on design

Major challenges faced related to the differentiation of the single variable changes and the ability to detect this. Trainees and staff physicians found the representation of the hemodynamic equation (CPP=MAP-ICP) useful. Staff physicians especially commented on the ability to better detect autoregulation, being a major concept in neurocritical care. Although mixed comments were provided about the more advanced summation trend including the increase and decrease aspect, discussions with our clinical collaborator Dr. McCredie led us to keep the simpler version (without the increase/decrease indication) for the first prototype version of the interface to be used in the usability study. There is still potential in exploring variations of this visualization that would especially help develop expertise in the long run.

In terms of design decisions, the hover over interaction sounded like a feasible and good consideration to display the exact values of the different variables. The ability to toggle variables on/off would be integrated to allow users look at one variable of preference at their preferred time customization. To support the user of the visualization and prevent any confusion about the individual variable changes (especially for ICP detection), we decided to show the individualized thresholds visualization below the summation trend chart to directly link the time instances and exact values across two visualizations.

5.4 Discussion

The interviews provided deep insights from both trainees' and staff physicians' perspectives. Both perceptions provide value as they are the main users of such displays (together with other clinicians), and it is important to identify the potential support for the end users for better care of patients.

Although both groups' comments were not always aligned, the final decisions on the designs were discussed within the research group to build an effective final version as a prototype to be used in a usability study with trainees.

5.4.1 Reflections on themes, subthemes and design impact summary

Both trainees and staff physicians provided general feedback about the visualizations as well. This feedback is captured and related to themes and subthemes from the interviews. Design impact ideas are listed next to them in Table 9 below. Themes with a * indicate presence in both groups (trainees and staff physicians).

Table 9: Themes, subthemes and related design impact for each visualization

Visualization	Theme-Subtheme from interviews	Design Impact
General	- Usability-Accessibility	Test visualizations against various colour deficiencies.
General	- Visualization Application-Different Users	Include a combination of simple and more complex visualizations and allow users customization options (i.e., time scaling, toggling of variables).
General	- Visualization Interaction-Time Scaling*	Allow users to zoom in and out of the timeline and allow users various time scaling options for the trajectory, and ICP dose summary.
General	- Visualization Interaction- Target Customization*	Allow customization of ICP targets but have default values for ranges.
General	- Visualization Interaction-Exact Values*	Hover over interaction to show the exact values of variables on the visualizations.
ICP Dose	- User Confusion-Interpretation of ICP burdens and percentages - User Confusion/Usability-Wording of y-axis* - Usability-Intuitiveness*	Explain visualization in the interface training video.
ICP Dose	- User Confusion/Usability-Wording of y-axis *	Label the left y-axis ICP Time Burden.
ICP Dose	- User Confusion/Usability-Two axes in context with graph*	Optionality to turn variables off and visualize one y-axis at a time (customization).
ICP Dose	- Visualization Application-Trend*	Include the trendline next to the ICP dose summary visualization as the dose summary (donut) does not show the trend.
ICP Dose	- Distribution of ICP values-Average*	Keep the average line to show the trend and optionality to turn on/off.

ICP Dose	- Distribution of ICP values-Fluctuations in ICP*	Hover over interaction to show the max value of ICP.
ICP Dose	- Usability-Differentiation of higher ICP values	Include a fourth ICP intensity category for very high ICP (greater than 30mmHg).
ICP Dose	- Distribution of ICP-Duration of ICP events	Anchoring the visualization on the dose visualization and allowing users to click on the hour to see the more detailed visualizations and what the burdens are comprised of.
Individualized Thresholds	- User Confusion-Interpretations of the colours of ICP intensity ranges	Explain visualization in the interface training video.
Individualized Thresholds	- User Confusion-Detection of threshold change	Include ICP target ranges alongside the visualization.
Individualized Thresholds	- Content Feedback-Annotations*	Include annotations on the interface.
Individualized Thresholds	- Design Preferences-All Subthemes	Consider what is being displayed together on the interface and display visualization that is consistent, however, distinct from the dose visualization to avoid confusion.
Trajectory	- User Confusion -Interpretation of Percent Change - User Confusion-Trendline Visualization	Explain visualization in the interface training video.
Trajectory	- Visualization Application-Alert - Visualization Application-Presence or Absence of Evidence - Visualization Application-Promotes looking beyond a threshold	Display visualization on the first screen of the interface.
Trajectory	- Visualization Application-Changes above Threshold - Visualization Interaction-Exact Values	Display visualization alongside the min/max ICP value.
Summation Trend Chart	- User confusion/Usability-Challenges identifying individual variables* - User confusion-Legend	Explain visualization in the interface training video.
Summation Trend Chart	- User confusion-Legend	Include a standard legend in addition to the equation above the visualization.

Summation Trend Chart	- User confusion/Usability-Challenges identifying individual variables*	Display alongside individualized threshold visualization (only shows ICP) and keep ICP on the top of the summation graph (hardest variable to detect).
Summation Trend Chart	- User confusion/Usability-Challenges identifying individual variables*	Allow users to toggle variables on/off to view one variable at a time.
Summation Trend Chart	- User confusion/Usability-Challenges identifying individual variables* - Visualization Interaction-Exact Values*	Hover over interaction for exact values of variables.
Summation Trend Chart	- Visualization Application/Physician Training-Autoregulation	Include cues on the interface to check for autoregulation.
Summation Trend Chart	- Visualization Application-Relationship of Variables - Design Preferences-All Subthemes	Display the simple summation graph version of the visualization as it shows the relationship of the variables and address usability/user confusion concerns around identifying the individual variables.

Some themes haven't been addressed directly and thus, ideas for future considerations are outlined in the next section.

5.4.2 Reflections on future considerations of design

Firstly, it is noteworthy to highlight that the end users of a bedside physiologic monitor are not restricted to physicians only; thus, one recommendation is to run further investigations on the visualizations with different users such as nurses for instance. Early research findings have been published for this exploration (Üreten et al., 2022).

The effectiveness of one visualization with various iterations could be tested too. In this study, the main focus was to identify preliminary aspects on usability and the concepts linkage through the visualization design.

Further, it is important to note that the visualizations shown relate to continuous ICP monitoring; often in the clinical environment, there is not always the chance to get a clean display of ICP signals (there is often noise or issues with the measurement itself). Thus, it can be investigated to look at non-continuous ICP monitoring too and how that would change the understanding of the visualizations.

In previous sections, it has been outlined that the ICP waveforms play a significant role too in understanding the data. Although we discontinued the idea of providing more details on the P1:P2 ratio or waveform analysis, we recommend further studying the data processing and extraction of waveform peaks that could be relevant to novices in the field especially.

Some specific details about the concept of providing a percent change indication has been discussed by staff and trainee physicians, and it seems that this concept can be further examined as this offers a different perspective or mindset to think about ICP in a more relative form. There could be hesitancy for adding this indicator to the communication and discussion of ICP among clinicians without having evidence yet of its usefulness, however, it provides a new way of discussing the analysis.

The overall comments on CPP have often related to an equally relevant variable as ICP; thus, it was discussed to have visualizations representing CPP targets or optimum CPP and CPP dose. Other neurophysiologic parameters have been discussed earlier and open the venue to further exploration of showing variables in context or keeping the visualizations' concepts but customizing them for different neurophysiologic variables.

Autoregulation is another core concept that trainees often learn through textbooks and might not consider directly at the patient bedside when new to neurocritical care. Although some staff physicians mentioned they could understand intact autoregulation from seeing data points across the summation trend chart, we have included a notification that would show up on the interface when a certain correlation of variables raise concern for impaired autoregulation. Further studies could investigate the usefulness and effectiveness of such notification under different clinical circumstances.

5.4.3 Reflections on the interview study

The conduction of this study took place during the Covid-19 pandemic where there were many challenges within the ICUs across the world. In such times of worldwide emergency, it has been a great challenge to get in touch with any critical care physician or personnel who may have time to share thoughts on the study. Having a greater sample size (we targeted around 30 participants in each group) could have shown more individual insights but would also underline further our current findings as similar themes were showing up among all visualizations.

Overall limitations related to time constraints faced for the study and participants' availability. Discussing visualizations within 30 minutes was a great challenge as there were multiple aspects to be discussed and it was not always possible to get feedback on each visualization with each participant. Some ideas for iterations have also been brought up during later discussions with participants to get more directions into potential final designs. However, the discussion on such possible iterations was stronger for the later-scheduled participants as we were able to hear patterns of major concerns from participants at the beginning and addressed them once they came up multiple times. Again, the effectiveness of each iteration on a single visualization should be evaluated in the future.

Also, sharing the challenges some trainees faced with certain parts of the visualizations were sometimes brought up with the staff physicians to get further ideas from the staff. This could also be viewed as creating bias for themes that came up and should be differentiated in an ideal case.

The visualizations shown were static and did not provide interactivity at the point of the interview conduction. The visualizations have been displayed separately, which in the ideal case would be shown as potential combinations to better link the value of each visualization in context.
Some ideas on a potential flow were raised towards the end of the planned interviews and provided some insights into user flows. This could benefit from multiple opinions in future studies.

As pointed out earlier, the visualizations presented showed ideal measurements and signal processing of ICP (continuous) data which in reality might not be the case always. Having stops in the signals or measurement interruptions can create different views of the visualizations and could be tested in future work.

The colours provided on the visualizations can further be improved and tested in various combinations for various levels and types of colour blindness.

Also, the size of the visualizations represented were in a PowerPoint slide and might differ from the exact interface sizes used in the clinical environment. When presenting the visualizations, no patient context was provided which could help to get the participant more into the clinical thinking process. Some assumptions were outlined such as having automation or digital inclusion for annotations which represents a challenge still for clinical personnel. Although many hospitals nowadays use digital flowsheets, the annotations are often not as precisely entered on digital platforms and physicians refer to the communication among nurses or other clinical personnel.

Another potential limitation relates to the demographics of participants in this study:

Not all experts were specialized in neurocritical care but had significantly relevant knowledge in critical care that would still provide great value to perception of the visualizations.

Also, for the trainees, it is often challenging to incorporate all their feedback as novices often don't know what they don't know or expected to still develop those mental models of an expert.

The participant pool was from North America (Canada, USA) and could benefit from a more international recruitment for future studies.

There is also the aspect of having access to technology, interfaces, digitalization and availability of certain type of equipment (e.g., ICP monitor) that varies among private and public hospitals but also on a level of rural and urban locations of the hospitals.

5.5 Conclusion

This chapter summarizes trainees' and staff physicians' perceptions and feedback on each presented static visualization. The main themes captured aspects of usability challenges or confusions, the relationship between variables, visualization application and content feedback, threshold detection, visualization interaction and customization, as well as the accessibility or usability by different end-users.

Overall, novices were confused about aspects such as the term 'dose' in this context as it is a term not often used in their training. This can be addressed by providing explanations during the training. Similar aspects of the quick identification of the different thresholds or ranges can be tackled in the training. When explaining these things to the trainees, they had full understanding of the visualizations and found them useful. We were able to identify more details on what specific aspects of ICP physicians watch out for, e.g., maximum values. By adding certain interaction opportunities (toggle or hover over features on the interface), more details on ICP can be presented and customized. The inclusion of annotations was deemed useful to better grasp the context such as interventions that previously happened.

Staff physicians found the visualizations clear and highlighted potential ways of reducing confusing aspects of the visualizations. They further commented that although certain terminology is not yet being used much in clinical practice, the consideration and use of different terminology or visualization representation trains trainees to think about the presented information in different ways and would impact their mental model development. Instead of focusing on a purely numerical and threshold approach, trainees need to think more in-depth about the trends and relationships of multiple variables and individual patient conditions. The understanding of autoregulation plays a key role in neurocritical care, and the visualizations provide hints towards the assessment if autoregulation is of concern. All in all, there was a strong link to the neurocritical care concepts described before, and the visualizations are representing these well.

The final design of the interface and the visualizations are based on the feedback provided by the trainee and staff physicians, as well as discussions with the research team. Overall, it was decided to have three major tabs:

- The vitals screen, showing the real-time representation of core variables as well as the ICP summary (ICP burden summary bar, ICP trend, and average/min/max ICP values) and the "impaired autoregulation?" notification.
- The ICP burden tab, displaying the ICP dose chart over time (1-hour interval with customizable scale) and annotations on the right-hand side.
- The pressures tab, including the summation trend chart at the top and the individualized thresholds visualization below, and the annotations on the side.

These visualizations were created and implemented as interface representation on a web-browser format with our team member Rayyan Quraishi, joining us for an internship in the Advanced Interface Design Lab.

5.5.1 Vitals screen

The initial view on the interface is the vital signs screen that also entails the ICP summary and impaired autoregulation notification at the top right-hand corner (Figure 21).

The vital signs represented would be HR, ABP, ICP, SpO2, RR, CPP and etCO2 whereas the latter three are only numeric representations and the others waveform + numeric values. The big number indicates the value measured in that second while the values next to it provide the normal range that is usually customizable. This screen is based on the Philips IntelleVue model.

For the study, we have included an "i" button next to the patient's name and sex on the top left corner, showing an additional window upon click. This window shows the current scenario text as well as a photocopied flowsheet.

Further, at the top middle, the vitals screen, ICP burden and pressures tab are clickable areas to navigate to the other screens. The date and time are represented at the top right corner.

The buttons at the bottom will be non-clickable for the planned usability study, except for the patient info button showing the core information on the patient such as age, sex, date of admission etc.

The ICP summary at the right-hand side shows a single bar with the percentages represented inside each stacked area which in total show a 100% view of ICP in different intensity ranges spent in the selected time frame. For example, the time customization is indicated below, with 1 hour, which would show that the patient has spent 87% of their time in a transitional range and 13% in a high range within that hour. Below the bar, the ICP trend is outlined as a line that shows the ICP trend for that hour with a declination of 25%. The average ICP within that hour is represented on top of the min and max values.

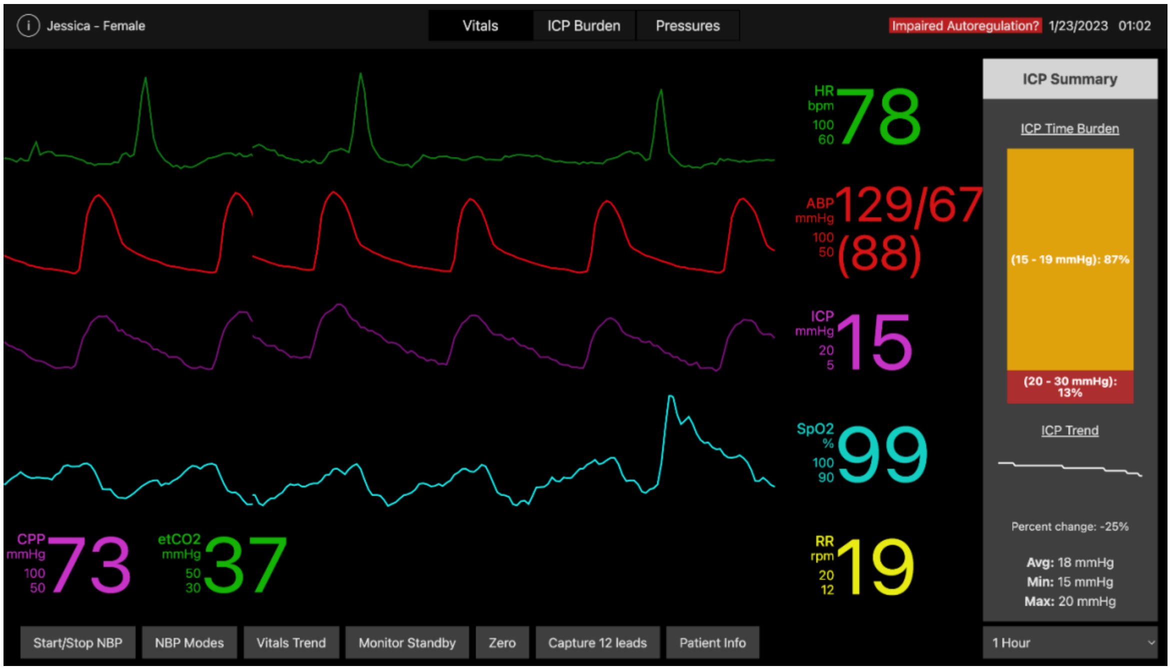


Figure 21: Vital signs screen of the prototype

5.5.2 ICP burden tab

The ICP burden tab (Figure 22) focusses on the ICP burden with clickable functions for the different ICP ranges of normal, transitional, high and very high. The specific indications are denoted on the top right corner showing the ranges for the colours. The average ICP is also a clickable option and shows the white line going through the graph. Annotations would be denoted at the right-hand side of the graph and will be displayed as blue circles on the time axis within the ICP burden visualization. It is also possible to filter the annotations by type (e.g. medication, procedures, patient position). This visualization enables the hover over function that would show the exact percentages for the specific hour including the average and max ICP. The time slider below the graph would enable additional interaction as well as the date and time selection areas on the right bottom. The background of the time slider shows only high and very high ICP bars to catch the attention of the user and let them easily navigate to concerning areas. The possible time selections are 24 hours, 12 hours, and 6 hours, which would show one hour per bar.
Clicking on a specific time on the x-axis leads the user to the Pressures tab or by directly clicking at the Pressures button at the top middle.

Figure 22: ICP burden tab

5.5.3 Pressures tab

The third tab shown in Figure 23 represents the summation graph at the top, the individualized thresholds graph synced on time below, and annotations as well as time and date selection on the right-hand side, consistent with the previous view. The variables MAP, ICP and CPP are clickable. The hemodynamic equation portrays a reminder of the relationship between the variables.

The individualized threshold graph captures three dotted lines that imply the threshold for each range; yellow for the transitional, red for high ICP and dark red for very high ICP range.

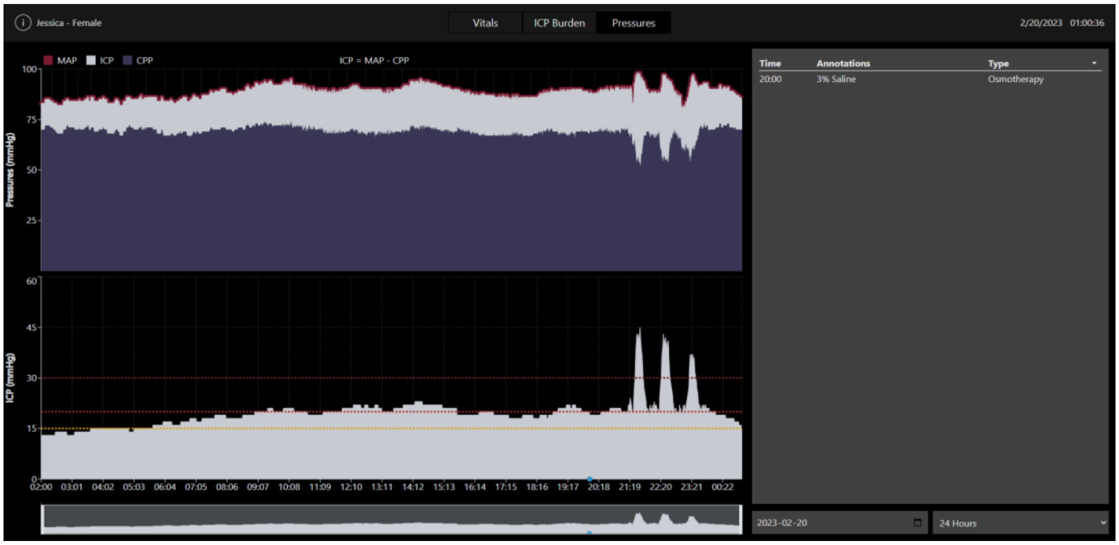


Figure 23: Pressures tab

The interactions are similar to the previous screen. The hover-over function would show the thresholds as well as the variable values for each graph as seen in Figure 24 and Figure 25.

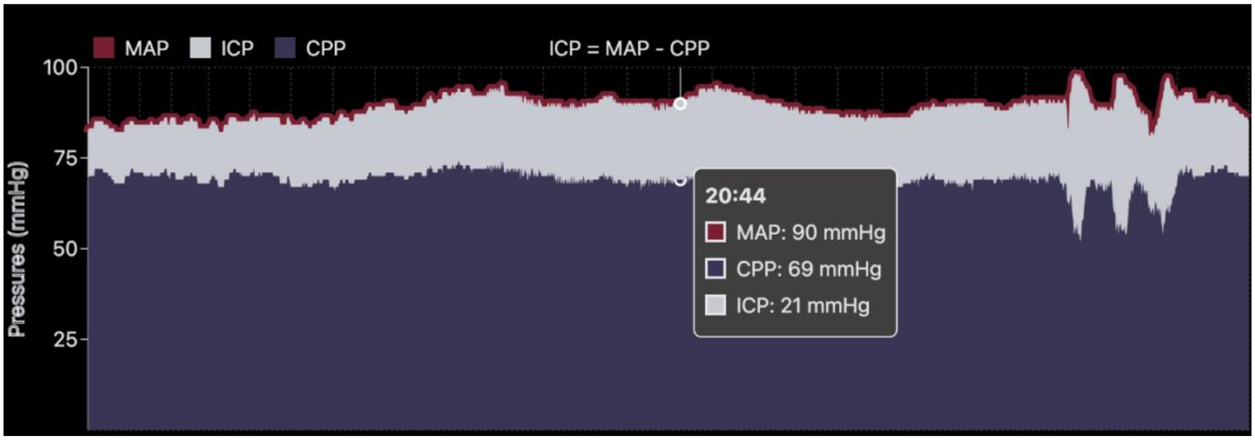


Figure 24: Summation graph - zoomed in

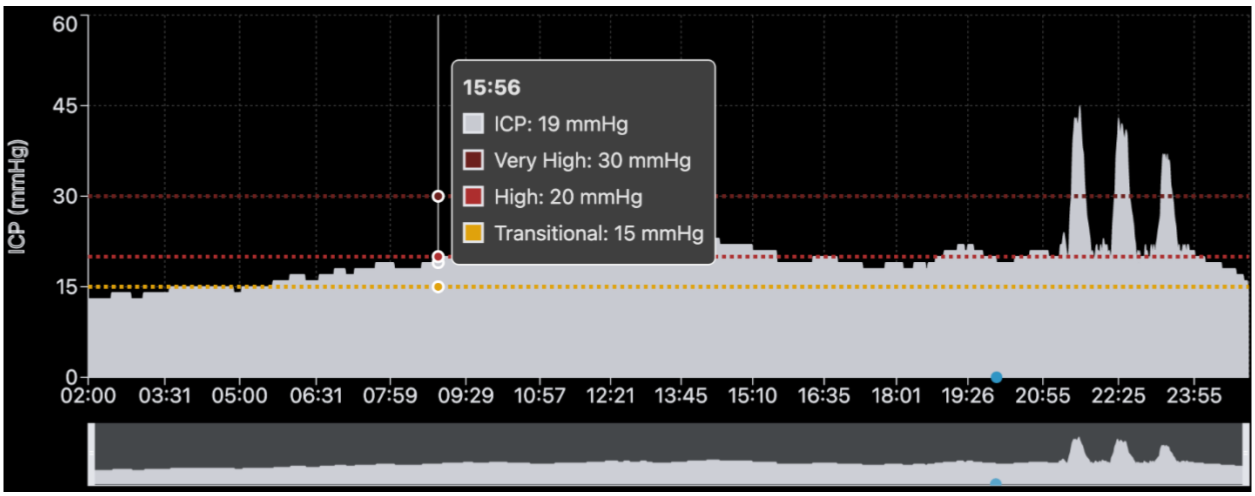


Figure 25: Individualized threshold graph - zoomed in

When the visualizations were created, some accessibility test were run for colour blindness simulation as mentioned on the website colour-blindness.com (Coblinder, 2000)

According to the National Eye Institute, there are four main red-green colour deficiencies deuteranomaly, protanomaly, protanopia, and deuteranopia (National Eye Institute, 2023) which were tested for the individual visualizations.

The pressures tab passed the test, i.e., the colours were distinguishable for all four colour deficiencies, whereas the ICP burden tab and the ICP summary showed limitations for protanopia and deuteranopia (red-green differentiation concerns). Major concerns arise for the differentiation of normal and high ICP. This implies that participants will be screened for colour deficiency in the usability study. Future work can investigate using various colour combinations, different patterns and shapes that would make it easier to distinguish major differences.

Chapter 6 The impact of EID on neurocritical care novices

6.1 Introduction

Considering the complexity and dynamics of neurocritical care units, it is important to investigate how the technology and the data represented on interfaces are being used. The gaps in literature show that there is a need for investigating in more depth how bedside physiologic monitor data representations are perceived by key decision makers, such as physicians, in neurocritical care. Interface design plays a key role in collecting and understanding data and can benefit from better and domain-relevant designs in the future.

This chapter is based on the research findings of the previous chapters. It aims to show the use of an ecological interface in comparison to a standard bedside physiologic monitor interface used in neurocritical care (mostly, the Philips bedside physiologic monitor is used in Ontario, Canada, and will be used for the control group). The represented study in this chapter has been designed to gain insights into how the EID supports aspects of expertise development in neurocritical care.

The high-level research question is

- To which extent does the ecological interface support expertise development (here only investigated in the short-term) when being compared to a standard (currently used) bedside physiologic monitor interface?

This question will be examined in the context of neurocritical care patient cases.

As discussed earlier, the term expertise involves various aspects and thus, we have focused on certain measures to include in this study that would show steps towards developing expertise. The first question listed below will help us get insights into how experts tackle such patient cases in general:

- What are expected strategies (commonly used by experts) for clinical cases in neurocritical care?
- Does the group using the ecological interface (experimental group) make fewer errors, as well as less harmful (or severe) errors, when compared to the group using the standard interface (control group)?
- To which extent does the experimental group provide a more expert-like reasoning and Situation Awareness (SA, e.g., level 3: trajectory), compared to the control group?
- Comparing both groups again, to which extent are the experimental groups' reflections on strategies outlined in their reasoning, the overall study and interface design, richer than those of the control group?
- To which extent are the confidence and performance perceptions of the experimental group higher than the control group (if at all)?
- To which extent are usability aspects on the ecological interface perceived higher than the standard interface?

The previous chapters showed various measures of expertise and considerations for interface design. However, we have identified and selected only certain measures for the usability study that are included in the research questions above.

Based on the CWA framework, the strategies analysis will help understand the overall assessment expectations of experts on novices in neurocritical care. This will be discussed with a neurocritical care expert.

We hypothesize that the experimental group using the ecological interface will make less errors as well as less severe or harmful errors. The harm level or grade of severity of errors will be evaluated by an external neurocritical expert.

Deep insights into participants' perceptions of the interface in a clinical scenario context will be more visible through listening to their reasoning. The reasoning will be linked to SA levels 1-3.
For this part, we hypothesize that the reasoning pattern of the experimental group will be more focused on tracing back the causes for ICP elevations, perceiving the trends immediately and having greater SA for levels 2 (understanding) and 3 (trajectory), compared to the control group.

In the previous chapters, reflections were identified as a further expertise development indicator; thus, we want to see if participants' reflections in the experimental group are stronger when they think about their strategies and have the opportunity to change aspects they mentioned during the scenario (i.e., self-correction or critical thinking). As the ecological interface offers various ways of thinking about the data in its visual context, we expect them to think more deeply about potential risks and causes for ICP elevations. Reflections on the overall study and interface design elements will help improve the study and interface for future research. These descriptions are related to what we claim of reflections being 'rich'.

Further, another hypothesis tackles the confidence and performance ratings to be higher for the experimental group compared to the control group. By providing more details and ways of visually understanding the patient cases in more depth through the ecological interface, we expect that participants will be more confident and perceive their subjective performance on the cases higher than the control group.

In the previous chapter, usability showed up as a key component to understand information better in context which is especially important for novices to learn. Thus, we hypothesize that usability aspects among the experimental group are higher or better than the control group. This may also be a potential explanation for participants' reflections on strategies, performance and confidence. However, we will not examine the correlation between the reflections and performance or confidence directly but take the usability feedback to validate that the ecological interface is not less useful and more difficult to use than the standard interface or can even be considered more useful or easier to use than the standard interface. The usability questionnaire will also be related to the integration of neurocritical care relevant concepts such as the ICP burden, understanding of the relationship between ICP and other vital signs, trajectory, individualization of treatment, and overall, being an improvement of information displays used in the ICU. We thus hypothesize that the ecological interface will be rated better on these usability aspects compared to the standard interface.

6.2 Methods

This section introduces the experiment design and the steps required to run the study. This study received ethics clearance from the University of Waterloo Research Ethics Board (#43798).

6.2.1 Experiment design

Participants were split into two groups: the control group using the standard interface prototype and the experimental group using the ecological interface. This study overall shows a between-subject design approach. However, as seen in Figure 26 below, the control group was shown the EID training and was given the opportunity to comment and reflect on the EID as well. This step was included to collect more impressions on the EID in general. The overall study progression for both groups is represented in Figure 26. In both groups, participants went through two scenarios. The duration of the study is approximately 60-90 minutes and is run in a single session, thus we hypothesize that the considered measures show short-term effects only and can be further explored on a long-term basis in future studies.

Remuneration was offered to all participants.

Overall, both groups were trained on the interfaces they were assigned to use for the study. Then, scenario 1 was provided where participants were asked to think aloud and answer probing questions by the researcher that relate to SA levels 1-3 and showed insights into their reasoning. This method has parallels to other EID studies where SA was measured (e.g., in Burns et al., 2007), and followed a qualitative approach.

After the scenario, participants were asked to rate their subjective workload (in this chapter, only performance on a scale from 0-100 is considered and neglects the other workload indicators that were assessed through the NASA TLX questionnaire).

Confidence in handover (scale from 0-100) and reflections on the interface and their own strategies were further assessed qualitatively. The same procedure was repeated for the second scenario. After both scenarios, participants were asked to provide ratings to a usability questionnaire. In the control group, participants were additionally shown the EID training at the end of the study and asked to rate the usability questionnaire again but this time for the ecological interface, as well as share their thoughts on what they would have done differently if they were to work on the scenarios with the ecological instead of the standard interface. In both control and experimental groups, participants were given the chance to share final comments, feedback, and reflections on the study. Errors and their severity would later be evaluated by a neurocritical care expert and is part of the measures.

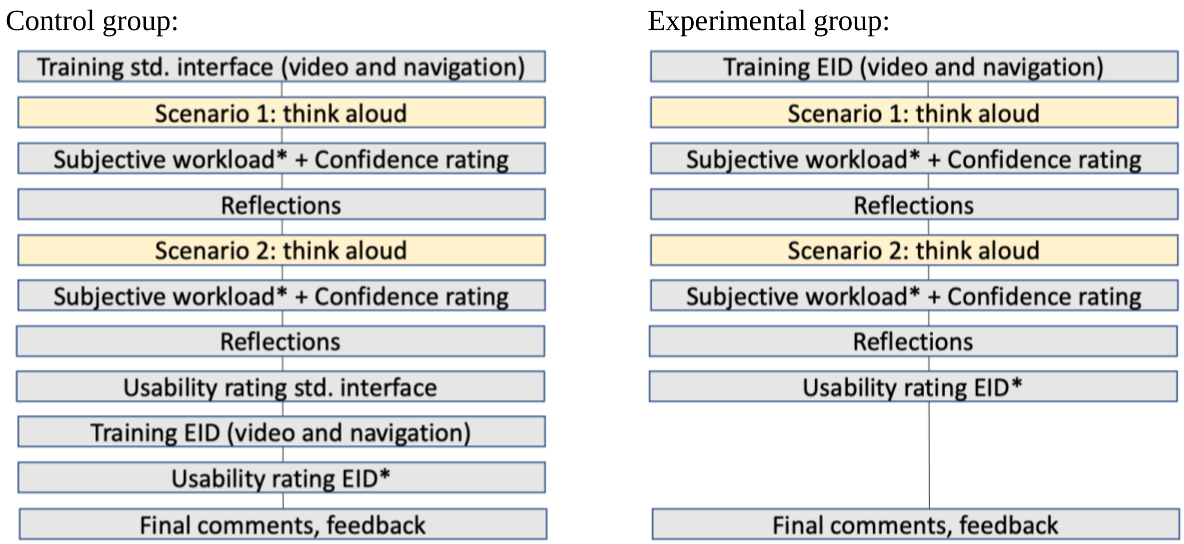


Figure 26: Study progression

(The standard interface is abbreviated as "std. interface")

The prototype interface development is briefly summarized in Appendix N.

6.2.2 Scenarios

When developing scenarios for the usability study, there were multiple parts involved; first, the author of this dissertation conducted online and literature searches for neurocritical care cases. Second, trainees were asked during the interviews what they would consider as a critical case in neurocritical care, where ICP was commonly measured. Third, two neurocritical care experts were asked to discuss and formulate the scenarios together with the author of this dissertation.
Common cases represented in neurocritical care where ICP is increased often relate to traumatic brain injury (TBI) and Subarachnoid hemorrhage (SAH, i.e., bleeding in the brain).

When trainees were asked in the interviews what specifics would make a patient case with high ICP critical, they responded that there could be:

- an acute change in the neurological status where it is required to have a CT or call neurosurgery (T4),
- neurological function deterioration, hemodynamic changes or intracranial bleed (T4, T10, T11, T13)
- threshold values, e.g., ICP 22 mmHg (T11)
- no further treatment options and ICP not going down (T12, T13)
- that the "monitor doesn't tell you much, but the clinical picture does" (T12)

With all these inputs, the intention initially was to create around 6 scenarios which is common for EID studies, however due to time and resource limitations and challenges with physicians' availability, two scenarios were developed together with Dr. Victoria McCredie and Dr. Alberto Goffi, both neurocritical

care experts, to represent a clinical context for the interface use. In the first scenario, the intention was to show a TBI case with a slow uptrend of the ICP variable in context reaching threshold values.
The second scenario represents a case where ICP hits high values in a certain pattern. This requires recognition of multiple relationships of variables in context to realize that the patient is not autoregulating and is in a life-threatening situation representing a so-called Lundberg A-wave pattern. In the second case, the scenario text will highlight that the nurse is calling the physician because something has changed (acuity, hemodynamic or neurological deterioration of an SAH case). The previously provided medical support and data will be depicted in the flowsheet that act as constraints or directions for the participant's options of treatment.
Certain relevant information will be represented in the flowsheet and scenario description so there is no need to do the neurological exam in that moment by themselves.

The data for both scenarios were created by Dr. McCredie and the scenario text in context with the data was formulated together with Dr. Goffi. The two scenario scripts are represented in Appendix L.

6.2.3 Training

Two training videos have been recorded through the platform Loom © to standardize the amount of information and duration on the interface introductions. One training video covers an overview on the navigation and explanation of the standard interface including the additions and limitations on the interface created. This video took around 2 minutes describing the vital signs' standard view in addition to the availability in the interface of a digitalized photocopy of a commonly used flowsheet.

The other training video explains all details on the ecological interface including descriptions of the new visualizations added, customizability, interactions, and limitations. The recording took approximately 12 minutes describing 3 tabs in total, namely vitals including the flowsheet, ICP burden, and pressures. The full script can be found in Appendix M.

6.2.4 Participants

Nine participants in each group took part in the study (in total 18 participants). They were randomly assigned to either group. Participants were first asked to provide demographic information which is represented in Table 10 below.
Inclusion criteria for the recruitment were:

- (neuro)ICU trainees, i.e., currently fellows
- Recent (neuro)ICU trainee graduates, i.e., early-career ICU staff
- Neurology residents with significant (neuro)ICU experience

Although the aim was to recruit a higher sample size, the recruitment time was extended, the inclusion criteria altered, and international recruitment (not only North America) took place. Thus, the study language had to additionally change from English to also include Turkish. The recruitment process took place through multiple ways:

- approaching fellows at a critical care conference directly and asking for interest in participation, sharing flyers and making announcements during the conference,
- emailing previous study participants who had agreed to being re-contacted for a follow up study,

- an invitation email was sent out to trainees through our clinical collaborator Dr. McCredie and previous experts who participated in the interview study,
- invitation emails sent out to critical care organizations,
- contacting experts in the field directly, who may forward the invitation to the study,
- professors and personal network from my own network and initiative in visiting hospitals internationally.

Participants filled out their demographic information prior to the study conduction. Their familiarity with ICP monitoring was also asked to better understand their frequency of use of the ICP variables in context, which helped to identify if they should be taking part in the study.

Table 10 below shows that most participants signed up from Canada, one person from the USA, two participants from Turkey and one participant who finished their critical care fellowship in Canada but at the time of the study conduction works in Thailand. The graduation years from medical school range from 2007-2018. Participants indicated that their graduation from medical school is from Canada, Thailand, Lebanon, USA, Spain, Brazil, Argentina, Costa Rica, Turkey, Mexico, and the UK. Further, participants' year of fellowship program varied between 1-5 years. Few participants indicated that they had previously specialized in Critical Care Medicine or pediatric ICU, and two participants mentioned they are currently neurology residents with 4 years of neurocritical care experience.

The specialty certifications of participants included:

- Neurology
- Internal Medicine
- General Surgery
- Emergency Medicine
- Anesthesiology
- (pediatric) Critical Care Medicine.

Although many participants indicated their subspecialty certification in General Critical Care, there were some who indicated Respirology, General Internal Medicine, Neurocritical Care and Neurology as well.

Overall, 8 male and 10 female participants signed up. The average age of participants was 33 years. The familiarity with ICP monitoring on average was 7 on a scale from 1 being unfamiliar to 10 being very familiar.

Table 10: Demographic information collected for the usability study

	Are you an ICU trainee physician (fellow) in Canada or U.S.?	When did you graduate from med school?	In what country did you graduate from med school?	Which year of fellowship program are you in currently?	What specialty certification(s) do you hold?	Which subspecialty certification(s) do you hold?	Which gender do you identify with?	What is your age?	How familiar would you rate your familiarity with ICP monitoring on a scale from 1 to 10?
Participant 1: EID	CA	2017	Canada	1	Neurology	General critical care	Male	29	8
Participant 2: std	CA	2011	Thailand	3	Internal Med	Respirology	Female	35	8
Participant 3: EID	CA	2017	Canada	1	General Surgery	NA	Male	31	4
Participant 4: std	CA	2010	Thailand	2	Emergency Medicine	Critical care	Female	36	9
Participant 5: EID	CA	2016	Canada	1	Internal medicine	general internal medicine	Male	30	6
Participant 6: std	CA	2017	Canada	2	Internal medicine	General critical care	Female	34	9
Participant 7: EID	CA	2015	Lebanon	3	Anesthesiology	General critical care	Female	32	6
Participant 8: EID	USA	2018	USA	1	Neurology	Neurocritical care	Female	29	10
Participant 9: EID	CA	2016	Spain	1	Critical care medicine	General critical care	Female	31	7
Participant 10: std	CA	2016	Brazil	2 with prior Critical Care experience	Internal medicine	General critical care (2 year)	Male	30	10
Participant 11: EID	CA	2009	Argentina	2	Critical care medicine	General critical care	Male	38	8
Participant 12: EID	CA	2010	Costa Rica	5	Anesthesiology	General critical care	Female	36	5
Participant 13: std	TR	2018	Turkey	4th year in neurology resident with 4 years neurocritical care experience	Neurology	General critical care	Male	28	7
Participant 14: std	CA	2013	Spain	1 year neurocritical care, previously 2 years PICU	Pediatrics, ICU	Neuro- critical care	Female	35	8
Participant 15: std	TR	2018	Turkey	4	Neurology	Neurology	Male	29	4
Participant 16: EID	CA	2007	Mexico	2 with prior Critical Care experience	Internal medicine	Respirology	Female	40	7
Participant 17: std	CA	2015	UK	2	Anesthesiology	General critical care	Male	32	2
Participant 18: std	Thailand	graduated 2022 July	Thailand	1 year staff	General surgery	General critical care	Female	37	8

Abbreviations used in the country of the current trainee column: CA-Canada, USA-United States of America, TR-Turkey, UK-United Kingdom. Participants were numbered as P1-P18. The assigned interface is indicated next to the participant numeration with std=standard interface, and EID=ecological interface. PICU stands for pediatric ICU.

6.2.5 External evaluation and preparation for evaluation

The study was conducted through Microsoft Teams (as before). All recordings have been transcribed and summarized into key statements. The statements for both groups have been compared with each other and organized by themes (inductively). The statements from participants of both groups were discussed and compared for each scenario by a neurocritical care expert. The strategies the expert would consider important were discussed upfront and will be explained in the findings section. The neurocritical care expert evaluated errors and their severity based on the context provided.

For the purpose of the study and to relate to the hypotheses and research questions stated earlier, the findings were analyzed based on selected measures of expertise development indications, portrayed in previous chapters resulting from literature as well as key findings of the interview findings. As further indicators were identified overall, the selection of measures was based on possible measurements for assessing an interface within a usability study. As mentioned in the previous chapters, qualitative approaches (e.g., through an inductive thematic analysis) were used to gain understanding of the depth of participants' thinking process in addition to some ratings.
In future studies, other indicators for expertise development can be further used and assessed.

6.3 Results

The results and analysis relate to various aspects; first, the strategies of an expert for such neurocritical care patients are considered as a basis and provide insights into notable differences to novices' approaches. The results with regard to the specific measures selected for this study will then be examined.

6.3.1 Strategies Analysis

To model the strategies of an expert, the CWA framework has been taken into consideration and shows an information flow map representation.

The overall strategies of an ICU expert are depicted in Figure 27 below when taking care of an admitted patient. The expert starts with a brief and general check of the environment to perceive any signs of distress and danger. The vital signs represented will be assessed for any acute signs and a physical exam will be done quickly. The primary survey deals with an assessment of airway (A) to rule out obstructions, breathing (B) to rule out respiratory arrest or failure, and circulation (C) to assess whether the patient has any cardiac arrest or shock. The primary survey incorporates an overall assessment, treatment, and monitoring. At any time, a request for help or assistance can be made. Once the initial control and support for the patient state is finalized, the second stage begins. If any information is missing, the initial assessment is restarted without progressing to the second stage.

The second stage presents a secondary survey and reassessment of the patient. This step might take longer and requires the incorporation of various reviews such as medical records etc.

A care plan is to be developed at this stage and the mental model of the expert is shared with the team for consensus and making sure nothing has been missed out. Any information required for handover is discussed too. Communication is also required outside of the ICU team, i.e., with the patient in case they are conscious and, or with the family or caregiver. Again, at any time the expert may request help or assistance.

The major component of the strategies applied here that shows the expert's way of handling a case is concurrent management, whereas the novice in the ICU often goes for a sequential management pattern.

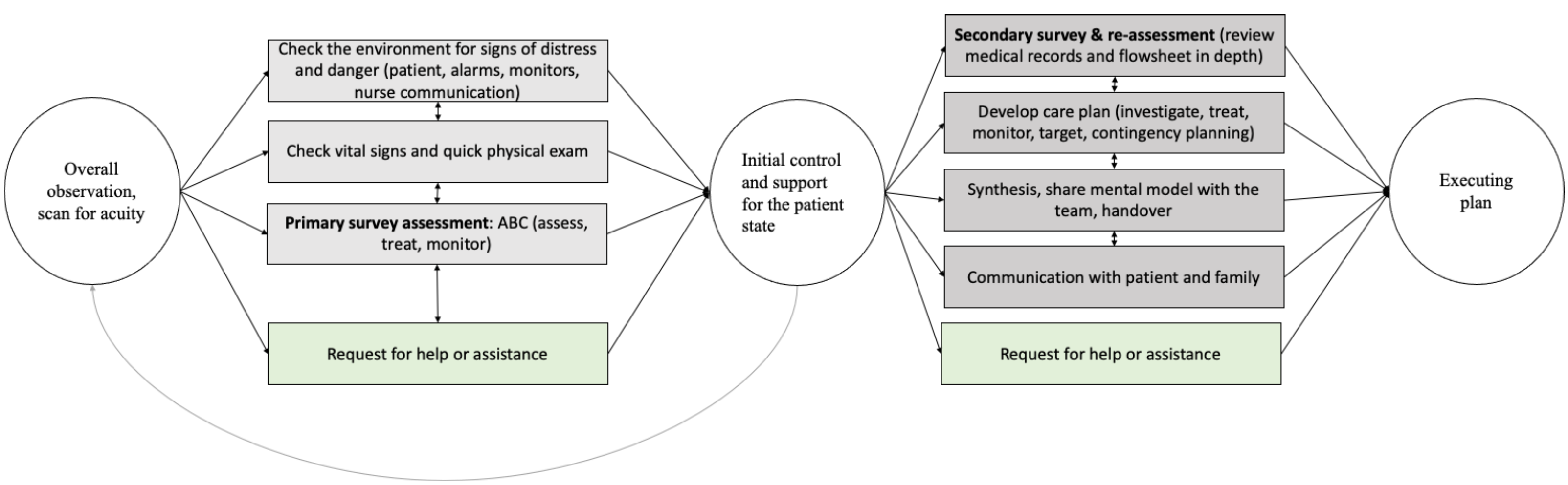


Figure 27: Information flow map of expert physicians in the ICU

Taking a closer look at the strategies that are pursued by expert critical care physicians, the models below portray between-patient and within-patient strategies, applied by the decision-makers. These two cases are important to understand how critical care physicians prioritize patient care and act towards different levels of risks and severity e.g., brain damage.

In neurocritical care, clinicians need to carefully understand patients' health states throughout their entire stay in the unit, and if available, assess past patient data to understand their situation fully. When physicians are alarmed by the team or through notifications by any device due to health deteriorations, they apply strategies to detect and assess abnormalities as seen in Figure 28 and Figure 29, to best support the patient with highest priority and improve their health condition. Although physicians and the team have certain levels of situation awareness of current health states of the patient, they need to predict the possible causes of abnormalities as well as treatment outcomes with the help of different data streams such as the bedside physiologic monitor. To come up with trajectory, it is important to use past and current data, and understand that each patient has individual thresholds for data ranges.

At the beginning of the critical care training, novices often struggle with the application of standard procedures and applying individualized care for patients, as well as zooming out from the current state to a broader picture of the health state progression. Expert physicians however have built their mental models and know how to look and interpret small changes in data; they include the aspect of time in their decision-making pattern and critically look at small changes of ICP for instance, to prevent secondary brain injury.

As shown on the DL, novices often know to apply early stages of the guidelines or so-called TIERs and rely on the standard procedures and ranges for such treatment plans. Expert physicians have gained intuition for quick assessment of abnormalities and easily know when and how to apply all possible treatment options including the adjustment of individual patient thresholds, possible future outcomes of applied therapy and have awareness for checking potential outcomes in specific time ranges.

With the StrA, we want to draw attention to the ability of switching strategies within and between patients in a critical care setting. Although this has not directly been part of the usability study (switching between patients), it is an extended model to the overall strategies.

In Figure 28, we have identified different strategies that would apply in a case where physicians need to prioritize the patient support **between various patients** in a similar state (e.g., neurocritical care, patients with TBI). In this case, they would evaluate distracting factors, assess the triage scale of a patient, evaluate the level of concern or acuity, consider the triangulation of comorbidities a patient might have or act proactively especially for acutely incoming patients.

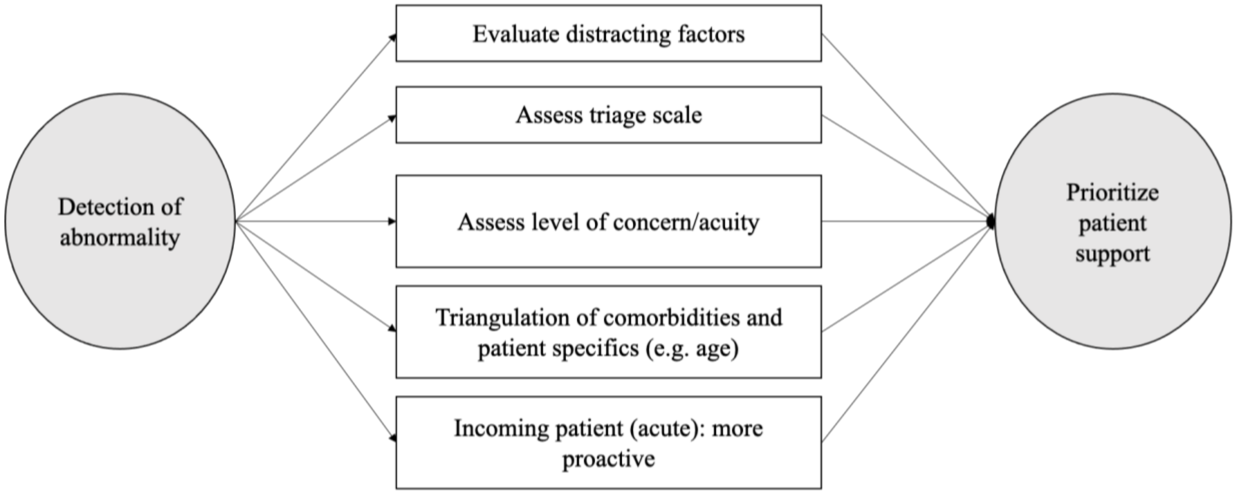


Figure 28: Between-patient information flow map of neurocritical care

Within-patient strategies involve the application of different procedures named TIERs in critical care. The different levels of TIERs in neurocritical care relate to the level of brain damage severity in the patient. Physicians would thus apply TIER 0 (such as raising the head of the bead and checking the collar position) when they detect that the patient's brain is at risk. TIER 1 (e.g., applying hyperosmolar therapy) is applied if the brain is damaged and TIER 2 (e.g., changing ventilation strategies) if the brain is about to die (see Figure 29). The greater the concern is perceived by the physicians, the more aggressive they would be about the treatment actions.

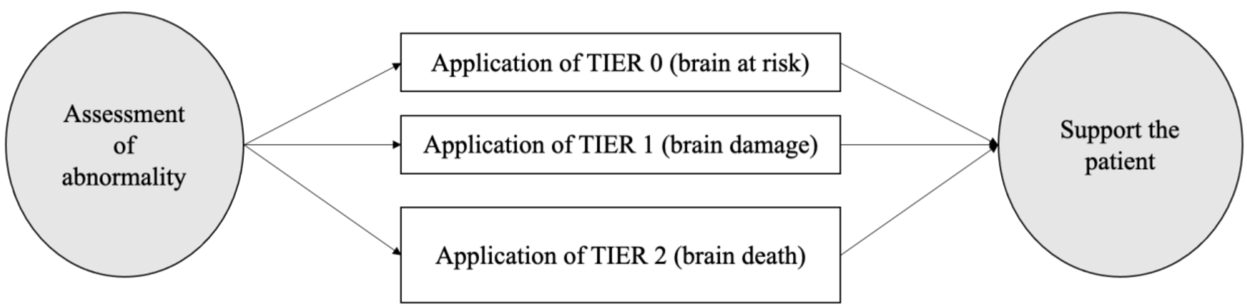


Figure 29: Within-patient information flow map of neurocritical care

Both figures show that physicians can switch strategies depending on the focus of one patient at a time (or being called in) and considering strategies to support and prioritize between multiple patients. This dynamic process is constantly evaluated in experts' minds and is aimed to transfer to trainee physicians' developing their mental models.

In the usability study, participants are confronted with two patient cases that direct them to take a more within-patient approach, as both scenarios show a single patient unrelated to other co-existing patients in the unit. The expectation of the scenarios is to perceive and assess the abnormalities represented through the interfaces guided by the scenario descriptions, come up with suggestions for procedures and treatment

of the patient while considering risks, thus provide trajectory and handover summaries of the patient to the next ICU team.
For the assessment of participants' responses and think-aloud, an expert neurocritical care physician shared his rationale and expectations for each step and evaluated errors and the harm level of participants' responses.

6.3.2 Number of errors and completeness of steps

With the neurocritical care expert Dr. Goffi, we have reviewed participants' statements and actions (visible through the think aloud) for both groups and scenarios and identified the number of errors. Parts of the relating StrA were thus a means of comparison for correctness and completeness. Comparing the control and experimental group, there are differences in the error count seen in Table 11:

Table 11: Number of errors and omissions of both groups and scenarios

		Standard interface	EID
Scenario 1	Number of errors	12	9
	Number of omissions	45	43
Scenario 2	Number of errors	8	4
	Number of omissions	35	36

In this table, the number of errors is higher for the control group as expected. However, no extreme differences are recognizable.

When considering the completeness of the steps, we encountered the number of omissions compared to the expert's strategies. The number of omissions was slightly higher in the control group for scenario 1 but lower for scenario 2.

6.3.3 Harm or severity of errors

The severity of errors was classified based on the 6-point Harm Scale (as in: Walsh et al., 2017).
The scale reached from 1 to 6:

- 1=unknown harm,
- 2=no harm,
- 3=mild harm,
- 4=moderate harm,
- 5=severe harm,
- 6=death.

However, when going through the evaluation with Dr. Goffi, we added scales in between such as 2-3, 3-4, or 4-5 since some errors could be viewed between the fixed categories too, depending on the expert physician's assessment.
Frequency distributions have been used to portray the results in this and the following sections, similarly as in St-Maurice & Burns' (2014) paper.

Among both groups, the severity of errors was rated and displayed through frequency distributions (Figure 30- Figure 33) for scenario 1 first and then for scenario 2.
The frequency graph (Figure 30) below shows how often both groups' participants made errors based on the error scale classification. The y-axis represents the frequency while the x-axis shows the error scale including the extended version (as described before). The control group using the standard interface is represented in blue, while the experimental group using the ecological interface is shown in orange bars.

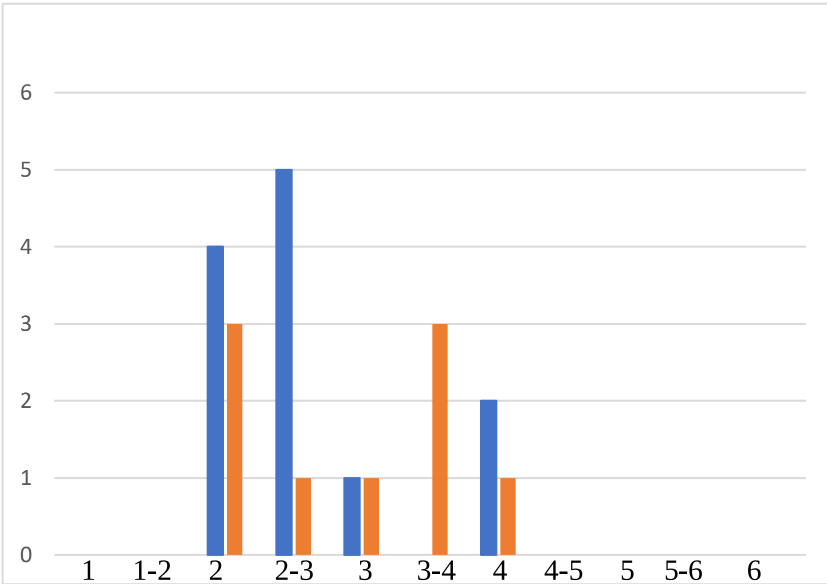


Figure 30: Scenario 1 Error severity frequency distribution

The control group has a higher number of errors, which is more significantly visible in the 2-3 severity category compared to the experimental group. The experimental group however shows a higher frequency of errors in the 3-4 harm scale while the control group presented slightly higher frequency on the moderate harm error scale (4). No errors were found on level 1, 1-2, 4-5 and upward.

The omissions are also represented below in Figure 31 on a frequency distribution graph. The experimental group omitted more relevant information that was related to the harm scale 3-4 and 4-5, while the control group omitted more details on harm scales 2 and 4.

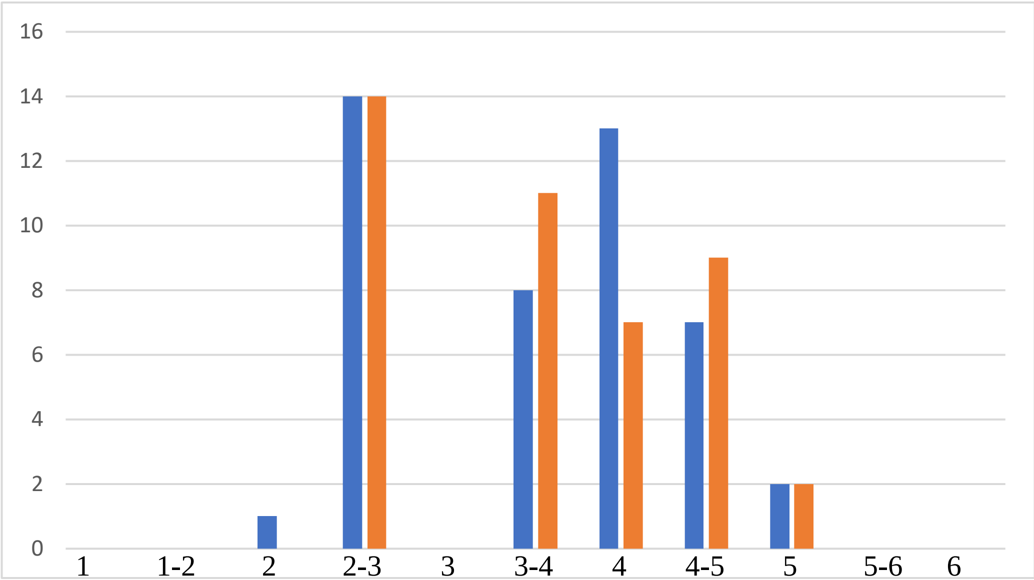


Figure 31: Scenario 1 Omission frequency distribution

In scenario 2, Figure 32 shows that the number of errors in each harm category was higher in the control group (represented in blue), except harm-level 4.

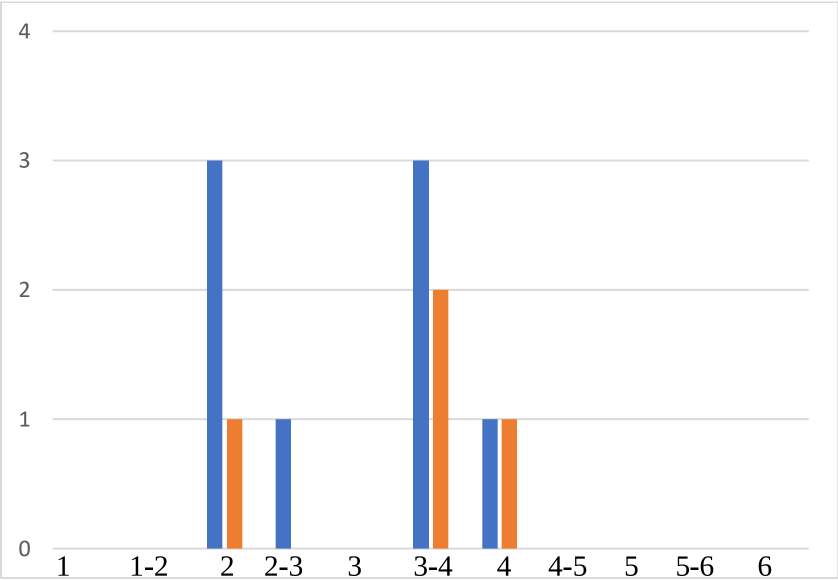


Figure 32: Scenario 2 Error frequency distribution

While there were more omissions in the harm scales 2, 3-4 and 4 in the control group, the experimental group showed more omissions in the 4-5 harm category (as shown in Figure 33). The frequencies of relevant omitted information were equally high in the harm scale 2-3 and 5 for both groups. Nothing was detected for harm scales 1, 1-2, 3, 5-6 and 6.

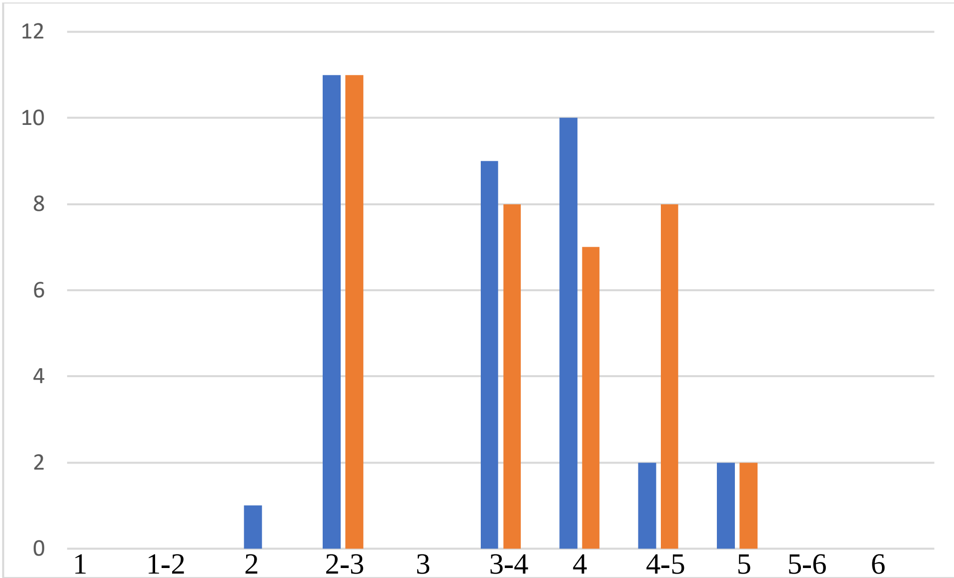


Figure 33: Scenario 2 Omission frequency distribution

In terms of severe errors, (level 5) there are none represented among both groups in scenario 1. There seems to be a tendency for the control group to show a higher frequency of errors around no harm, no to mild harm, and moderate harm, while the experimental group's frequencies of errors are slightly higher around mild to moderate harm in scenario 1. Further, the omissions in this scenario relate highly to no harm to mild harm in both groups, and there is a bigger tendency around the harm scale 4 visible in both groups. The omitted aspects in both groups relate to severe harm with the same frequencies in both groups.

For scenario 2, the error frequency of the control group was higher, especially on the no harm scale and mild-to-moderate harm scale. The frequency of omissions seems to be higher for the experimental group around moderate to severe harm; however, both groups have similar frequencies in the scales of 3-4 (mild-to-moderate harm) up to severe harm due to the omissions.

6.3.4 Reasoning and Situation Awareness

Letting participants think out loud provides ways into how they think when looking at the interface. This was also related to SA levels, and participants were probed by the researcher during the scenarios if they did not share details during their think-out-loud process. The questions asked related to the three SA levels as follows and were not quantitatively evaluated but rather qualitatively to understand participants' directions of thoughts and upcoming themes on those three levels:

Level 1: Perception. Initial assessment, describing what they see.

- Can you share what you see is going on?

Level 2: Understanding. Concerns verbalized for the patient.

- How would you describe your understanding of the patient's situation?
- Are you concerned about anything?

Level 3: Trajectory. Short and long-term risks perceived, treatment plan and priorities, communication, and handover.

- What are your thoughts on the risks for this patient? What might happen to the patient in the next 5-10 minutes/couple of hours, 12 hours or in the long term (e.g., next 3-10 days, or weeks)?
- Is there anyone you would like to communicate with about this patient?
- What things can be done now, and what can wait?
- How would you handover this patient to the next team, what are key aspects the next team should be aware of?

Details on the themes and codes can be found in Tables 12 – T17 for scenario 1 and Tables 18 - 23 for scenario 2 in Appendix O.

6.3.4.1 Scenario 1

Specific descriptions and themes on each level and think-aloud statements are provided for both scenarios. The expectations on these levels were discussed with Dr. Goffi and tendencies for both groups were looked out for during the evaluation, also relating to the error and omission aspect.

Level 1: Initial impression

One main theme was captured when the control and experimental group participants shared their thoughts on initial impressions: they all directly checked the vital signs (see Table , Appendix O).

Both groups commented on the numeric values of the variables they saw, initial concerns related to medication, herniation, autoregulation, hypertension, and settings of the equipment.

Some participants in the control group noticed that the ICP waveform was non-compliant whereas others did not recognize this. Comments from the experimental group related later to this (level 2: concerns). The experimental group, however, shared their observations on the ICP uptrend and initial statements on certain correlations of variables regarding the response to therapy.

This part shows that the uptrend of ICP is directly visible on the ecological interface and the correlation of that to provided therapy is noted at first glance. Some control group participants seemed to focus on the waveform quicker, while this was raised by the experimental group right after making the initial observations.

Level 2: Concerns

The main themes raised by participants of both groups relate to the physiologic variables, the effect of variables on the health state, the patient state in general, procedures, and the correlation of response to therapy (Table , Appendix O).

When control group participants raised concerns, they added more details about variables such as Heart Rate (HR) and Mean Arterial Pressure (MAP) compared to the experimental group. Both groups commented on the increasing trend of ICP and other neurologic assessments such as pupils.

Also, when sharing their thoughts on the effects of physiologic variables on the health state, the control group mentioned hypotension, bradycardia with Cushing reflex, and prolonged hydrocephalus as differences from the experimental group. Both groups talked about rebleed and herniation as potential effects of such a state but highlighted there would yet be no concerns about vasospasms. A participant in the control group mentioned not being concerned yet about Delayed Cerebral Ischemia (DCI), whereas this was perceived as a potential risk in the experimental group.

The control group suggested hyperventilating the patient to bridge the patient to the aneurysm intervention, while the experimental group highlighted that a CT and blood work check might be required. Comments in the experimental group also related to not having the need for any other intervention for now. Both groups commented that the Extraventricular Drainage (EVD) and drainage target should be checked.

The experimental group commented further on the correlation of the response to therapy; reasons for why ICP might be high could be due to sedation, coughing, pain, and waking up. It was noticed that the medication delivery was not yet maximized and could have an impact on the current state, but the response to therapy with the EVD was well.

In level 2 SA, the experimental group also shared more thoughts on the correlation of the response to therapy which the control group did not talk about much even in the later stages of their analysis.

Level 3: Short and long-term risks

The themes raised in this category were: physiologic variables, proposed procedures, and risks (Table , Appendix O).

More details were provided within the experimental group when looking at the proposed procedures; some suggestions were to wake up the patient because of low and decreased propofol levels, check the pain scale and medication, repeat CT, provide more medication in general (hypertonics, saline, blood pressure (BP) medication) and consult the interventional radiology (INR) and neurosurgery to assess the EVDs in case of no effects seen after hypertonics were given.

In terms of risks, the control group mentioned hyponatremia, a syndrome of inappropriate secretion of antidiuretic hormone (SIADH9), worsening levels of consciousness, bradycardia, infections long-term disabilities, and no hypertension or hypoxia for now. The experimental group however mentioned risks of edema and blocked EVD potentially as risks that differ from the control group. Common risks highlighted by both groups were e.g., rebleed and vasospasms.

When talking about short and long-term risks in level 3 SA, participants in the control group listed a wider range of risks compared to the experimental group. The experimental group, however, provided more insights into the proposed procedures combined with some risks they perceived.

Level 3: Priorities and treatment plan

Common themes among both groups when discussing priorities and treatment plans relate to physiologic variables, medication, procedures, and risks (Table , Appendix O).

Talking about physiologic variables, the control group listed many variables they initially mentioned at the beginning of their observations again, stating that ICP is high, but the participant wouldn’t adjust it for now,

keep the patient normoxic, exclude hyperthermia, check lytes, and control other variables listed in Table 15. The experimental group talked about a more focused description of treating and observing BP, controlling other variables, and raising the respiratory rate (RR).
The control group brought up a greater variety of medication they would provide such patients, such as nimodipine, lassix, antiepileptic prophylactic, or anti-seizure medicine. There were also comments about increasing or decreasing propofol to reduce risks. The experimental group mentioned increasing sedation and analgesia to reduce pain and provide hypertonic saline.

Procedures also differed among both groups; although many were mentioned earlier already, the experimental group suggested providing TIER 1 and TIER 2 therapy, to check electrolytes and blood work while keeping the patient asleep.
Risks associated by the control group related to highlighting hypertension, hypotension, fever, hyperglycemia, and swelling. Rebleed as a risk came up among both groups. The experimental group mentioned that vasodilatory effects are also a risk factor for the patient.

When talking about priorities and treatment plans for the patient, the control group came up with many considerations of how the physiologic variables should look and how the different types of medication can be used. The experimental group listed fewer options for medication needs and suggested TIER 1 and 2 therapy in general.

Level 3: Communication
When both groups commented on clinical personnel they would communicate about this patient, they mentioned neurosurgery, neurovascular, and nurses (Table , Appendix O). The control group also found it important to contact anesthesiologists and endovascular clinicians. Radiology and the patient's family, while the experimental group mentioned the neuro-interventionalist and attending fellow or resident.

In terms of communication, different clinicians would be considered among both groups whereas the control group also considered talking to the patient's family.

Level 3: Summary and handover
As a final remark on participants' observations and considerations for the patient shown in scenario 1, they were asked to come up with a summary and a plan for handover to the next team (Table , Appendix O). The themes all related to the ICP uptrend, response to treatment, communication with other clinical personnel, suggested procedures, previous procedures made, suggested therapy and medication, as well as risks.

Both groups noticed the ICP increasing pattern, whereas the experimental group also highlighted that CPP dropped while ICP increased, the waveform compliance was poor, and the heart rate (HR) increased.
While the control group mentioned that there was an effect of the sedation and drainage, the experimental group said there was a good response to the EVD placed.

Consultation with neurosurgery has been brought up by both groups. The control group added neurology, the endovascular team, and their own ICU team, whereas the experimental group noted that a radiology consultation would be relevant.

Suggested procedures by both groups are similar. While the control group participants mentioned that coiling should be considered, the experimental group said that surgery should be considered if ICP crises are acute and consistently over 20 while ICP control is no longer a possibility.

With regards to previous procedures, the control group mentioned that the EVD was placed, while the experimental group listed more details of what has been done, such as hypertonics were given, RR and CO2 was increased, propofol was provided, and the patient came in after the OR with an unsecured aneurysm.

Risks were similarly tackled among both groups with a few differences. Rebleed, vasospasms, hypertension, and hydrocephalus were mentioned in common, whereas the control group highlighted DCI, transtentorial herniation, bradycardia, and agitation. The experimental group listed propofol as a risk factor.

Suggested (medical) therapies brought up by the control group were antiepileptic treatment, letting the patient sleep, and increase ventilation. The experimental group noted that they would consider analgesia, neroperiphine, having a low threshold for more hypertonic saline, consider coiling and titrating down propofol to wake up the patient slowly. Common aspects were to increase sedation and provide Nimodipine.

For the summary and handover, both groups noticed the ICP uptrend but only participants of the experimental group mentioned a correlation between a decreasing CPP and increasing HR level, while ICP shows a poor waveform compliance. The control group had mentioned the waveform compliance issue during their first observations of the interface; however, they did not bring it up again during the summary part. Both groups revealed that there had been a response to the therapy that was provided (e.g., sedation). Further, both groups commented on risks, but the control group added more details. Previous procedures were highlighted more strongly by the experimental group; suggested therapies were to a certain extent similar among both groups, whereas the control group considers coiling, and the experimental group considers coiling if there is no option left anymore to control ICP if it keeps in the high ranges.

Some differences in the type of medication are also visible in both groups.

6.3.4.2 Scenario 2

Statements are summarized among all three SA levels and can be found in detail in Tables 18 – 23 of the Appendix O.

Level 1: Initial impression

Themes derived from the initial impressions level 1 SA provided insights on the vital signs check, suggested procedures, ruled out therapies or diagnoses, communication, previous procedures, risks and suggested medical therapy (Table , Appendix O).

When participants checked the vital signs, both groups noticed the ICP uptrend. At points, they were talking about an appropriate look of ICP or visible spikes. Participants in the control group talked about the ICP waveform that seemed to look appropriate. Other variables were checked in context, whereas the control group was describing their observations on more variables than the experimental group.

Both groups came up with some initial suggestions for procedures they deem relevant for this case; while the experimental group mentioned checking the CT, EVD, positioning of the patient, and noticed no surgical treatment yet, the control group was suggesting to additionally check ventilator settings, start with the typical high ICP management TIER, and consider various other factors such as the neurological state and waveforms.

The control group ruled out some conditions the patient is not suffering, such as a Cushing effect, focal deficits, respiratory distress or sedation effects. Some participants of the experimental group mentioned there would be no emergency at the moment but recognized the Lundberg A-waves of the ICP pattern.

Participants of both groups denoted that they would like to communicate with the nurse to understand why they were called or what concerns were perceived.

Initial comments were raised about previous procedures. The experimental group mentioned multiple aspects about the response to previous therapy and clarified that they would have done something earlier to react.

Participants in the control group mentioned at this point that a risk they perceive is blunt cardiac trauma, whereas more risks were verbalized by the experimental group within their initial impressions on this case such as bleeding, fractures, edema, mass lesions.

Participants in the control group immediately suggested some medical therapy ideas while the experimental group highlighted such later.

From the first impressions both groups shared, it is possible to see certain tendencies in their way of perceiving the ongoing. While the control group is considering many different physiologic variables, the experimental group comments more about ICP.
The control group seems to take a more general viewpoint when considering procedures and immediately talk about medical therapy. The experimental group however shows greater tendencies with regards to the response to therapy and risks they perceive.

Level 2: Concerns

When the groups commented on the concerns they perceive, the main themes were the ICP uptrend, risks, suggestions for therapy and medical treatment, previous procedures or therapies, diagnoses, and communication. Details can be found in Table of the Appendix O.

While the control group noticed the high ICP values, the experimental group started to question the spikes and acknowledged a correlation of ICP with MAP and CPP.

To get a better picture of the concerns, both groups suggest similar therapies such as guiding with the TIERs of therapy, neurological and temperature check. This also refers to similar suggestions for the suggested medical treatment.

While rebleed, herniation and brain death are risk considerations of the control group, the experimental group has concerns about ischemia especially when seeing the second Lundberg wave. Comments also related to an exclusion of herniation and that a decompressive craniectomy might be needed.

In terms of diagnosis, the control group mentioned hypertension as their main concern while the experimental group stated a decreased brain compliance and existence of Lundberg waves.

Several comments were provided from the experimental group on previous procedures and therapies; the EVD was checked, rebleeding, scans, monitor insertion, concerns about sedation and understanding of the thinking process of the previous physician who had been in charge of the patient. The experimental group further added that a consultation with neurosurgery would be beneficial if the situation worsens.

The concerns raised overall related to more aspects of ICP and its correlation with other variables for the experimental group when compared to the control group. While the control group thought more about hypertension as diagnosis, the experimental group associated the pattern they recognize for ICP with the Lundberg waves and correlation to previous therapy and procedure in that context.

Level 3: Short/long-term risks

When participants talked about risks, they mentioned that the risks depend on the ICP trend, but there would also be long-term and short-term risks associated with such patient cases. They proposed procedures and added other comments that are summarized in the theme "general risks". Details are represented in Table of the Appendix O.

When commenting on the risks related to the ICP trend, the control group mentioned that it is not easy to see from the flowsheet what happened before the increasing trend. The experimental group, however, started hypothesizing that if the autoregulation would go back once the ICP is controlled and CPP would be possible to further increase.

When relating to long-term risks, the control group highlighted general ICU infections that might be an additional risk factor, whereas the experimental group focused more on the brain function.

Although both groups commented on similar aspects for the short-term risks, the control group mentioned a greater variety of potential risks.

One participant in the control group mentioned it would be too soon to prognosticate and information is missing for that at this point. The experimental group, however, talked about multiple general risks they perceive and mentioned that the young age of the patient is an indicator for quick swelling.

The proposed procedures were similar overall, while the control group additionally brought up to use the transcranial doppler for further examination and check relationships between variables ($PaCO_2$ and $etCO_2$). The experimental group participants considered maximizing medical therapy, getting more (neuro) monitors but also preparing the family for the risks.

Although many of the perceived risks aligned among both groups, it is notable that the control group again showed a more general approach on potentially existing risks (e.g., line infections etc.) and the limitation of the flowsheet that makes it difficult to prognosticate. The experimental group had a stronger focus on the control of ICP, autoregulation and consideration of age as an additional factor for a shorter time of brain swelling.
Also, the procedures varied among both groups; while general aspects on treatment were similar, the control group considered the transcranial doppler and checking relationships between some variables important, while the experimental group verbalized that the maximization of therapy should be focused on and be ready to share the ongoing with the patient's family.

Level 3: Treatment plan and priorities
When both groups were asked to share their thoughts on the priorities and treatment plan they would go for, the major themes related to the consideration of physiologic variables, medication, procedures, and risks. Details are outlined in Table of the Appendix O.

While the control group commented on a variety of physiologic variables they would like to control and keep within a certain range, the experimental group commented only about BP control. They shared more thoughts about the procedures in general and prepare their plan for treating the potential next wave in a more aggressive way.

The medication comments among both groups were similar whereas the control group commented more precisely on the exact medication.

In terms of risks, both groups had similar impressions whereas the control group commented further on hypothermia, hypoxemia and hypotension risks and the experimental group reflected on the potential correlation between coughing, discomfort or secretions resulting in the current situation.

At this point, participants of the control group shared further risk factors while the experimental group tried to connect the reasons for such ICP trend. Major differences also relate to the idea of treating more aggressively than before (experimental group).

Level 3: Communication
When participants mentioned themselves or were asked by the researcher to share whom they would communicate to about this patient, the control group related to neurosurgery, nurses, radiology and the patient's family while the experimental group considered only neurosurgery at this point. Table shows details in the Appendix O.

When asking specifically about who the participants would communicate to in such patient case, neurosurgery was the most common response. Even if participants in the control group mentioned nurses, radiology or the patient's family when being asked, these also came up during the overall think aloud of the experimental group.

Level 3: Summary and handover

During the final remarks on the scenario, all participants commented on the ICP uptrend, response to therapy, communication with others, suggestions for procedures, thoughts about previous procedures, diagnoses, risks they perceive and the medical therapy they would suggest for such patient. Table summarizes details and can be found in the Appendix O.

About the ICP uptrend, the control group's participants also mentioned at this point that the case presents a young patient and describes observations on many physiologic variables, especially ICP being high. The experimental group focused more on their observations on periods of high ICP and CPP.

Both groups talked differently about the response to treatment; while the control group mentioned it seems easy to ventilate and oxygenate the patient in addition to the EVD drainage and relationship to ICP, the experimental group thought the spikes could relate to the nursing care.

The experimental group considers communication important, especially with radiology, neurosurgery and the nurse if ICP is over a certain threshold for a few minutes. The control group mentioned (neuro)surgery and TTL (trauma team lead).

Some participants mentioned details about their diagnoses such as the patient being tachycardic and hypertensive, or the pupil reactions being ok (control group), or comments on the noncompliant waveform that shows the autoregulation being off but having normothermia initially (experimental group).

The risks both groups highlighted at the handover stage were to a certain extent similar. However, the control group considered to keep an eye on HR for risks of bradycardia, while the experimental group also included complications with the monitor insertion as a risk factor.

The procedures they suggested considered appropriateness of the EVD, double checking the CT, potential neurosurgery and control of the neuroexam. The control group additionally stated they would check for multiple injuries, temperature, specific indicators in the blood work as well as the HOB or C-spine collar. The experimental group highlighted they would be more aggressive and try all strategies as a difference to the control group.

The medical therapy both groups suggested revealed similar suggestions around saline, mannitol, sedation in general, osmotherapy, and pain control.

When both groups described the previously given procedures, they both mentioned the EVDs while the control group raised insecurities if the EVD was clamped or not working. The experimental group said that the EVD is currently open for drainage but not much is drained.
Different details were brought up when sharing what has been done before, such as the waiting for the scan results, sedation, ventilation (control group) and some insights on saline (experimental group).

The summary or handover part gives an overview of what participants in both groups deemed important to share with the next team. While the control group shows a general overview and overall viewpoint about the patient (considering other injuries, many physiologic variables etc.,), the experimental group again highlighted to try all possible strategies and treat more aggressively overall.

6.3.5 Reflections

This part incorporated multiple aspects:

- Reflections on taken strategies during the scenarios,
- Reflections on the whole study,
- Reflections on the interface in general and in use during the study scenarios as well as usability evaluations.

Reflections on the interface and strategies were provided after each scenario. The overall reflections on the study were discussed at the end of the study.

6.3.5.1 Reflections on taken strategies during the scenarios

Firstly, participants reflected on their approach to detecting, understanding, and predicting the patients' states, as well as on potential alternative approaches and reasoning for choosing the strategy they talked about during the scenario think-aloud. Some participants also reflected on the limitations of the chosen strategy.

The themes are summarized for each group.

Control group

The main themes identified in this group were the following:

- Targets provided by experts
- Focus on numbers
- Lack of information
- Use of systematic approaches
- Considerations of alternative medication
- Consideration of future steps
- Consideration of overseen aspects
- Consideration of additional measurements

Targets provided by experts – theme:

One participant mentioned "[…] usually, we ask for the targets of systolic BP, ICP and CPP" (P4) which related to their rationale of the patient's well-being and settings needed. Expert physicians have the expectation that the trainees (fellows) come up with the targets themselves and are able to progress with the assumptions.

Focus on numbers - theme:

Another participant pointed towards the use of a more numeric approach "I wasn't looking very closely at the ICP waveform, at our hospital, we look more at the numbers" (P6). This statement shows that fellows do not always incorporate the integration of multiple data representations that may provide details on the patient's state.

Lack of information - theme:
This participant (P6) also mentioned that if they knew the problem, they would have changed the approach but would otherwise stay with the approach mentioned during the scenario.
Although multiple participants in this study mentioned that they would be interested in seeing the CT scan for the patient to make decisions, one participant mentioned that this was actually not needed because the information on ‘mFisher4’ was given in the scenario description which entails the necessary information from a scan (P13).

(P13) also mentioned that neurosurgery usually provides other medication such as cortisone steroids for antiedema treatment which was validated to be wrong by the expert physician. Further, the participant said that he doesn’t know the patient’s response currently so it is hard to assess this patient, though he could have thought more about the ICP and drainage.
(P14) added, that seeing the scan would have helped to see if the patient doesn’t have hydrocephalus. This would mean that the ICP parenchymal monitor would suffice and there would be no need for EVDs.

Use of systematic approaches - theme:
Further comments focused on systematic approaches as an alternative, such as the ABCDE approach to go through each step one by one, as an alternative to quick assessments as in the scenario. If there was more time given at the bedside, this participant would double check his own approach by systematically checking the standard ABCDE procedure and look for aspects he might have missed out on. (P10)

(P17) reflected on scenario 1 that he would like to do a “clinical exam and assess neurologic states, airway breathing, circulation and check other injuries.” He further would have wanted to review the imaging and use that to inform management and family discussion. He added that “people usually intervene aneurysms pretty early”, and he doesn't know why it hasn't been done yet, “maybe there was a decision made by the surgical team”. In terms of the management plan, he mentioned he “could have focused on monitoring ICP and CPP more, maybe allowing for more sedation in the 24-48 hours to give the patient e.g. normal saline (if they hadn't that before) as well as sodium 140-145”. After scenario 2, this participant reflected on own limitations and mentioned he focused mostly on the brain but “this is a trauma patient that could have trauma anywhere”.

Considerations of alternative medication - theme:
(P13) reflected about other medication that neurosurgery often uses (i.e., cortisone styreoid for antiedema treatment).
Further, (P18) said she would give more sedation “plus minus mannitol” and notify neurosurgery while checking the EVD and pupil reaction. (P18)

Consideration of future steps - theme:
Another participant reflected on the ‘liberation for ventilation which could be considered in the future’ for scenario 1 (P14). She also mentioned that it would be difficult to prognosticate so early for a severe TBI patient in scenario 2 but could put a triple volt of PbtO2 monitor and was unsure if there is a requirement for the EVD and ICP monitor. She noticed that the patient is localizing which could mean it is OK to just have an ICP monitor.

Further, one participant reflected after scenario 1 that vasodilatory medication could be given around day 7 “as we expect vasoconstriction on such patients” (P15) which was assessed as incorrect by the expert physician validating the responses. This participant reflected that he would not make changes to what he mentioned during scenario 2 but further medication could be given, such as starting an anti-edema treatment when ICP goes up and ABP varies.

Consideration of overseen aspects - theme:
(P13) highlighted that he doesn’t know the patient’s response at this point, but it would have been better to consider ICP and the drainage.
The last participant in this group (P18) noticed that she didn’t look at the third page on the flowsheet and added that she would usually look for neurological signs and CSF drainage.

Consideration of additional measurements - theme:
(P14) reflected that adding the transcranial doppler and the PbtO2 monitor would be beneficial, but it is also ok to have only one monitor.

After scenario 2, (P18) said she wouldn’t change the strategy for now but is “unsure about the loss of autoregulation but it doesn’t look so good.” She also mentioned that she would be interested in seeing a plot of ICP and CPP to check for autoregulation and do a physical exam, potentially adding the Licox monitor. “Sometimes we raise the BP and see how the ICP behaves.”

Experimental group

The themes identified in the experimental group’s reflections are as follows:
- Consideration of more aggressive treatment options
- Consideration of further risks and providing medication
- Lack of information
- Consideration of overseen aspects or ruling out risks
- Realization of ICP focus
- Consideration of the flowsheet
- Realization of overseen variables
- Consideration of additional measurement

Consideration of more aggressive treatment options - theme:
Reflections about the strategies and limitations stated by participant 1 after scenario 1 were “if my approach was wrong, I would have to catch up on not being more aggressive approach early on. But this is a very common scenario so this is what I would do in this situation.”
Participant 3 reflected after scenario 1 that “I could've been more aggressive titrating down BP, but I actually wasn't thinking about it much until the end.”
Reflecting on scenario 2 (P8) said the “big thing about this case is to maximize medical therapy. Why did she have those spikes, was it because she was turning or due to sedation. Maybe we could get more hypertonic saline or mannitol, probably saline and get her back on the autoregulation curve, if that doesn't work, then go for the decompressive hemicrania. Probably do those things more aggressively if I was doing it again. If it doesn't improve after giving mannitol or saline, it depends how long the ICP is up for. If it transiently goes up to 25 for a minute it’s ok, but if its 10 min that’s a problem. When ICP was in the 30s

it was transient. If they stay up, I would give 23.4% saline and call neurosurgery, if it doesn't go down the patient needs to go to surgery." (P8)

(P9) also added that she would have reacted earlier and gone through all TIERS of therapy; trying sedation, analgesia and if that doesn't work, do osmotherapy. If that doesn't work then paralyze the patient and if that won't work either, she would call neurosurgery.

A few more reflections were stated by (P11): "I would aim for deep sedation (after seeing the flowsheet). Ideally put an EEG monitor on to help guide the dose and better control the ICP, infusion of opioids. Maybe there are any fractures missing? If she has uncontrolled increase in ICP, not improving with sedation and osmotherapy and she did have no craniectomy, then for the decompressive craniectomy we should talk to neurosurgery. Temperature is covered, CSF is covered, hemodynamic management is covered, there are no other strategies further, it would be surgery. Now it looks like ICP is controlled as it is around 15 after the 3% saline were given. At least I would be asking for the opinion for neurosurgeons, what would trigger a surgical intervention in their mind, then we can set same targets." (P11)

A similar statement was provided by (P16) saying that she would have acted earlier since there was no decompressive craniectomy or any surgery, this being a more complex case than scenario 1.

Also (P11) mentioned that in case ICP remains uncontrolled and is not improving with sedation and osmotherapy, then a decompressive craniectomy would be planned, while checking temperature, CSF, hemodynamic management.

Consideration of further risks and providing medication - theme:

Participant 7 reflected after scenario 1: "Usually, I would be worried about vasospasm and BP augmentation. I would have told this during the handover if vasospasms might be an assumption. The BP target should be higher. Did they clip the aneurysm? If it's not secured you have the risk of rebleeding. If it is clipped, then there is no risk of rebleeding." After scenario 2 she added, "this one is more acute".

Another participant (P8) provided an in-depth reflection after scenario 1 "There's an algorithm, TIER 0,1,2,3. That's what I was going through for my initial situation. You can do hypertonics before they hit 20, usually it's not right at the beginning when the patient comes in, more when they continue to have the issue. If it's given too early, hypertonic can lose the gradient for osmotic therapy. If they hit 30 and you did mannitol before, then you're stuck with nothing else, by then you are paralyzing them. I wouldn't necessarily do that very early on."

For scenario 2, (P1) mentioned that "other approaches would be to treat now or wait for brain to cool off. We see a trend, there's an ICP trend going up and down. I would put this patient in burst suppression and increase the propofol from 30 to 300. I would further examine this patient, the ICP was not 40, so my approach is to be patient and wait for the next event while getting the scan. Some centers would put a continuous EEG for a patient like this. I don't see a use for this at the moment, but it would be easy to argue for it. A limitation is the risk that the patient's ICP goes up again. But also, there is a risk to treating something in a preventative manner. That's why I would stick to my approach."

Participant 9 reflected after scenario 2 that she "usually you get information from the nurse. I would check vitals, pupils, images and compare them, if treating hyper ICP, I would give hypertonic or sedate the patient and see if it works, usually sedation works instantly. Probably I would have done a CT or talked to neurosurgery before or deeply sedate her before. I would have tried to deeply sedate her when she had hyper ICP, if it's not working, I would check if hyper ICP only in certain moments or maintained over time. If it sustained, I would be very worried also when the patient was not responding to medication. If its temporarily, maybe it's because of coughing or suctioning or touching the patient, I would be less concerned but if it's over 20 for 1 hour and after all our efforts it is still 20, she should go for decompressive craniectomy. Reacting earlier would have been better and talking to neurosurgery. Everything else is controlled (etCO2), I don't know the sodium, but she received hypertonic saline I would repeat that. I think we went through all TIERs, sedating, analgesia if it's not working then move to osmotherapy if it's not working paralyze her, if not working call neurosurgery.

Lack of information - theme:
(P8) further commented that "although the ICP waveform is constant throughout and there are 2 EVDs, I could have asked which one is open to drain. Checking that foremost is important. You get the waveform when it's closed. We wouldn't have a constant ICP unless the patient got a bolt, it sounds like they were bad enough to need 2 EVDs to drain all of the blood, they should both be open. Checking this is important. I assumed the patient was flat for a long time because of getting the EVDs and came in at 2 am."
After scenario 2, (P3) added that one limitation is "not knowing the imaging and how amenable she is for neurosurgical decompression, there are some nuances of which lesions benefit from a craniectomy. There is a lack of nuance about the lacking imaging."

Consideration of overseen aspects or ruling out risks - theme:
More reflections were provided by (P8): "There is no mass lesion that requires them to have a decompressive hemicrania because it's subarachnoid. There might be slight differences in how early they do hypertonic, there are no other guidelines. I would still do what I explained during the scenario in terms of strategy."

Realization of ICP focus - theme:
(P9) noticed in her reflections that there was a big focus on ICP in general.

Consideration of flowsheet - theme:
Besides reflections on additional monitors, one participant shared thoughts about potential time lag due to the green coloring of the additional tabs on the ecological interface after scenario 1: "It's better to see the trends of the variables. It's just that the flowsheets are like that. I would have worried about the patient earlier if I looked at the trend on the flowsheet. When the screen is in green you don't worry but the reality is that the patient isn't good. I worried when I saw the white trend." For scenario 2, this participant reflected that she would have been more aggressive in treatment for this patient. (P12)

Realization of overseen variables - theme:
(P7) reflected that she forgot to consider CPP.

Consideration of additional measurement - theme:
(P11) reflected that he would put an EEG monitor and consider missing fractures.

Control group using the ecological interface
Also in this group, similar themes came up:
- Considerations of more aggressive treatment options
- Consideration of autoregulation
- Consideration of new targets for different variables

Considerations of more aggressive treatment options - theme:
"It might change management, it's very helpful to see what has happened in the past." (P2). This participant referred to the ICP burden graph mentioning it was "clearer, the trend in the graph in each hour. We usually estimate the ICP (the average ICP is useful) and it's better to see the trend even if the patient looks clinically stable (but the ICP might have been rising). Often, the 'normal' ICP numbers aren't seen in the ICU much." (P2)
(P10) further reflected that "This is a bad pattern, often related to cerebral vasoconstriction after vasodilation, so I would be even more concerned about this patient, it gives me a different insight. I don't see a couple of waves only, I see it throughout the time. Overall, I would treat more aggressively."

Consideration of autoregulation - theme:
(P6) reflected that "in the second case, I worry that her autoregulation is impaired, we do not often do the MAP challenge. We don't have all of this information readily available. If I saw the ICP spikes and I was able to recognize that something acutely changed, it would have certainly been on my mind seeing on this interface. On the other interface I briefly thought about it but did not explore".
She further added that "interfaces are not just something that we interpret, but it also prompts us to learn more and treat in a different way. For example, if I had this information, then I would learn more about how to do autoregulation challenges at the bedside, it would actually generate a new skill set which currently my skill set in that specific domain is not very good because we don't do it and we don't do it because we don't necessarily have that information available. I think this is a very important part of it. I wish we had something like this" (P6).
Further ideas were stated by participant 15: "Cushing or bradycardic factors impact autoregulation. Here we can see that the autoregulation is intact. I wasn't able to see that directly on the standard interface, only ABP hinted a bit towards that, I was considering dysfunctional autoregulation maybe around 50% but, on the EID, I consider it around 80%. I might already consider loss of autoregulation around 1 pm area."

Consideration of new targets for different variables - theme:
Participant 10 said he "really like(s) it, it has much more information and all of it is very relevant. Maybe I need to use it more. A zoom-in waveform would be useful, P1, P2, P3 to get insights on brain compliance and over longer periods of time, to see Lundberg patterns. At this time, we cannot have a strong opinion about autoregulation and try to manipulate it. But if the patient is not doing well and the basics are not working, I would move forward, maybe MAP changed, CPP changed based on autoregulation. Maybe I could find a new target for CPP, maybe the usual target is not adequate. So, for the patients that are not responding to the basics, I think it is a good path". Referring to the 'Pressures tab', he also added that "this is exactly what I was talking about. This is a bad pattern, often related to cerebral vasoconstriction after vasodilation, so I would be even more concerned about this patient, it gives me a different insight. I don't see a couple of waves only; I see it throughout the time. I would maybe try to be a bit more conservative with the ICP maybe bringing it a bit lower and treat more aggressively."

6.3.5.2 Study reflections
As both groups got the chance to familiarize themselves with the EID, overall reflections on the study and using the EID are summarized in Table in Appendix P. The main themes captured interface comments,

learning with the ecological interface, considering other variables, scenario comments, usefulness, user-friendliness of the ecological interface, and autoregulation.

Interface comments related to visualizations used to confirm reasoning: "The summation graph is useful to confirm the conclusion or reasoning." (P1). Further comments are outlined in Table on this subtheme. Clinical relevance was another subtheme in this category: "I don't know if the bars are useful clinically if there are frequent reassessments happening" (P3), questioning the visualizations' impact in long-term usage.

Comments on colours were revealed and the subtheme "alerting signs" related to the following exemplary quote: "The red colours help to even see things are bad from 10 meters away. Without a training video, it might be confusing, what the colours represent. It's helpful to see within the box what the colours mean (high transitional normal)." (P18)

ICP trends and values were considered a further subtheme: "The actions are based on the numbers or thresholds. Maybe in the future it would be more about the ICP burden and not only about the threshold. It would give more of the feeling about prognostication. For decision-making, you might still talk to the family, but it might not change much unless there is more data about the ICP burden correlation with the prognosis. The brain is not a straight-forward thing. This is a good start for further research in the area." (P4)

Further comments linked to the "annotations related to time"-subtheme: "The ICP burden blew my mind and the annotations. The patient at 2 am had a high ICP and had an intervention that helped a lot, now the patient has been safe for the rest of the time. The values and waveform tracing are great, maybe notifications could help: watching out for normal compliance, for people who are less experienced. The numbers are from literature but we are not sure if that's the right thing so the waveforms might be good to see. Examples of normal and abnormal tracings: P1>P2> P3, abnormal: P1 and P2 higher. Sometimes this also happens when the monitor doesn't transduce properly. In the transitional zone, people might just chill but they shouldn't." (P14)

Suggestions for further improvement or "other interactions" as a subtheme were provided: "An icon or clickable button would be useful to click next to ICP time burden, as a definition, ranges and maybe a website as a reference, to explain what it means." (P11)

The theme "learning with the EID" was split into further subthemes. The EID being an "individual learning tool" was highlighted by multiple participants: "Interfaces are not just something that we interpret, but it also prompts us to learn more and treat in a different way. E.g., if I had this information, then I would learn more about how to do autoregulation challenges at the bedside, it would actually generate a new skill set which currently my skill set in that specific domain is not very good because we don't do it and we don't do it because we don't necessarily have that information available. I think this is a very important part of it. I wish we had something like this!" (P6)

Some participants also mentioned a "quick adaptation" and formed another subtheme: "The learning curve was fast for the interface. In the second scenario, I was just thinking about the case not the interface." (P1)

Other participants also noticed that there may be more of a "use for experts": "The pressures tab might be more helpful for the experts rather than the novice, whether it is true impaired autoregulation or not. A person working not regularly in neuro-ICU maybe wouldn't look at it." (P17)

The "transfer of mental plan" was identified as another subtheme: "I know the targets, I don't need them, it might be valuable for the team to share the mental plan and urgency of the situation, but for me personally

I don't need the green/yellow/red" (P5). Another participant highlighted that the "EID is great for handover" (P16).
Also, in the context of "research", participants mentioned: "EID helps to get all information. Would also facilitate academic research, could be used for training/learning" (P15).

The theme "other variables" was split into multiple subthemes. The "incorporation of other variables in context" subtheme was mentioned by multiple participants saying, e.g.: "I'm interested in seeing the MAP, ICP, CPP, and PbtO2 altogether on a graph. I could then figure what's the best for the patient on MAP or CPP. Maybe a line graph" (P10). Other measurements, variables, or techniques were further mentioned by other participants and are detailed in Table 24 in the Appendix P.
Another subtheme was named the "use of a messaging system for TIERS of therapy", as one participant commented that "overall, one could consider the next TIER of therapy through a messaging system which might be very useful" (P3). A similar subtheme called "notification" was outlined because of another participant's insight: "notification is nice to consider but I'm always thinking about it because it affects your decision making, if they are herniating or responsive to treatment. I did really like the relationship graph that was my favorite" (P8).

In the theme "scenario comments", two subthemes were captured. One named "trends of ICP" was related to the comment "Scenarios were both showing trends. Maybe if there was an active ICP of 40, that might maybe change the management approach" (P1).
The other subtheme "scenario 2 complexity" was identified due to comments like: "the first one was a little trend and not necessarily worrisome, it's something we see very frequently, and no one really knows what to do with it perfectly but the second one was not only we weren't sure what to do but we know that we will harm the patient if nothing is done. The second one was definitely more challenging" (P1).

The theme "usefulness of EID" was linked to the subtheme "a resource-efficient tool": "the EID would also be very helpful in other parts in the world where there is no nurse for each patient available (nurses visualize the trends)" (P16).
The theme "user-friendliness of EID" had one sub-theme on "time-scale adjustment" with varying comments: "It was useful, easy to navigate, flexible with just enough information" (P11), however, (P9) mentioned "I ranked the summation graph at lowest because the time scale adjustment was hard to use."

"Autoregulation" was the last theme that was captured with the subtheme "consideration of autoregulation as a concept": "We talk about autoregulation in our exams, but I don't know exactly what to do about it. Maybe it's also because I don't know enough. The notification hints towards it, I might have to control BP. I know it in theory, but I don't know how I would act on it. I would try to bring down MAP and SPB and check ICP and CPP and see what happens. If there's a linear relationship, there might be impaired autoregulation and I would have to be more tight on hemodynamic targets" (P17).

6.3.5.3 Input on the improvement of the interfaces

The control group using the standard interface mentioned that the interface is the same or very similar to the interfaces they are familiar with. Thus, they found the interface presented in this study clear and intuitive. Overall, participants in both groups highlighted the following aspects and themes as in Table 25-

26. The column on the right-hand side shows how their concerns can be improved for the next version of the interface used in future studies.

The main themes identified by the control group relate to colours, flowsheet, timescales, and information representation. The subthemes and direct quotes can be seen in Table .

Table 25: Themes of the control group and improvement ideas on the interface

Themes	**Subthemes**	**Direct quotes**	**Ways to improve in the next version**
Control group using the standard interface			
Colours	ICP and HR representation	"It's hard to see the ICP and HR colour and waveform." P2	Make lines thicker for ICP and HR.
Flowsheet	Readability	"Some parts on the flowsheet are very small." P2, P17	Portray the information in bigger size.
		"Nurses usually highlight changes on the flowsheet with a red marker or circle it. It's hard to go through black text. Nurses would also add a comment why the EVD is not draining." P6	Big changes in numbers could be highlighted on the flowsheet to call the attention of the viewer.
		"It was difficult to check RR on both flowsheet and monitor." P14	Make RR readings more obvious.
Timescales	Variable alignment	"ECG and SPO2 look atypical and not aligned." P4	Correct the phases of variables to be aligned.
		"Usually, we see a faster HR waveform, this was confusing." P13	Correct the phases of variables to be aligned.
		"HR and ABP are not synchronized which is confusing." P14	Correct the phases of variables to be aligned.
Information representation	Combining more information on a single screen	"[…] Would like to see hemodynamics and neurodynamics all on one page." P6	Variations in visualizations representation can be done.
	Zoom-in interaction	"Zooming into the ICP waveform to better see P1 and P2 would be helpful." P10	A zoom in option can be included in future versions.
	Other variables	"[…] Interested in having the waveform representation of etCO2 when patients have respiratory issues." P10	Representing etCO2 can be incorporated in new variations of visualizations display.
		"[…] Would like to have the etCO2 next to the respiratory parameters displayed, instead of next to ABP." P14	Representing etCO2 can be incorporated in new variations of visualizations display.
		"Temperature might be useful to see." P15	Temperature can be included on the display.

The themes verbalized by the experimental group related to the interface perception, tabs settings, other variables, flowsheet, and annotations.

The subthemes and direct quotes of participants can be seen in Table below.

Table 26: Themes and subthemes of the experimental group and ideas on the improvement of the interface

Themes	**Subthemes**	**Direct quotes**	**Ways to improve in the next version**
Experimental group using the ecological interface			
Interface perception	Usability	"The interface was very user-friendly. Not an overload of information. It had most of the information I needed. I forgot about the flowsheet, but I was happy to see it was available (after telling). Nothing on the interface was confusing." P1	The training could have a stronger focus on the flowsheet inclusion.
		"This interface was clear. It was easier than the first one, I understand the interface better now. It was intuitive and had all the information needed for this case." P1	
		"The bars in each range (burden) are less important to me than the actual numbers. Otherwise quite reasonable to use." P3 "It's clear enough, it was quite easy to navigate and understand." P5	
		"Everything was pretty clear but needed a reminder on the colour meaning of the ranges." P8	Adding clearer colour definitions or offer repetition during the training for more familiarity.
		"On the vitals screen, I didn't see the initial ICP so I went to the Pressures tab but didn't know that you can see the initial value when hovering over. But everything was user-friendly." P8	The training could be extended to enhance familiarity of interactions.
		"[…] I find it a bit hard to adjust the slider on the pressures tab, not very easy. It would be helpful to see both for ICP burden and pressures tab." P9	The interface contains a lot of data. A different structure for the code can be implemented to accelerate interactions.
		"It is useful to modify and navigate at different times." P11	
		"It is difficult to distinguish between the red and dark red dotted line on the trajectory graph." P16	The differentiation can be made more clearly, e.g., by changing the colour or making the dots bold.
	Quick adaptation	"It's much easier to use now, it's easy to learn, only seeing it one time suffices." P7	
	Alerts	"I liked most the percentage [ICP summary on vitals]. It gives you a heads up if something is going on, it alerts you even if you forgot to check the trend of the past 24	

		hours. Nothing confusing about the interface." P7	
	Combining visualizations	"The last graph maybe has too much information and colour if the ICP is in two places, maybe it's not needed (more info than needed). The bottom graph is sufficient, I don't need to see the ICP over MAP and CPP because we use it all the time, I don't need the visual representation of it." P5	New combinations of visualizations can be considered.
Tabs settings	Default settings	"The 3rd screen [pressures tab] was set at 6 hours use instead of the other, which was on 24 hours, I didn't catch on the acute rising ICP but everything was reasonable." P3	Set all defaults the same way for all tabs.
		"Only on pressures I didn't notice first that I looked at the first 6 hours and not last hours/24 hours." P12	Set all defaults the same way for all tabs.
Other variables	Imaging	"It would be good to have the images (CT)." P9	More information can be included to make further information accessible.
Flowsheet	Integration of flowsheet within the interface	"I just have troubles with the flowchart because I don't know how it is built, it takes more time than if you know where to find the information." P5 "I liked that the nursing chart was connected." P7 "Very clear, only problem is that some things are too small (flowsheet)." P9, P12	More training time can be dedicated for the flowsheet structure. The size of the flowsheet information can be increased.
		"I got the scenario by better looking at the interface rather than flowsheet. This is when my view changed and could not tell from flowsheet. On the flowsheet, just one number/second is written." P7	
		"I forgot about the flowsheet it was good to have it." P8	More training time can be dedicated for the flowsheet.
Annotations	More details for annotations	"[…] I would like to see more about medication (this was missing)." P11	More details on previous treatment can be included.

6.3.6 Confidence ratings

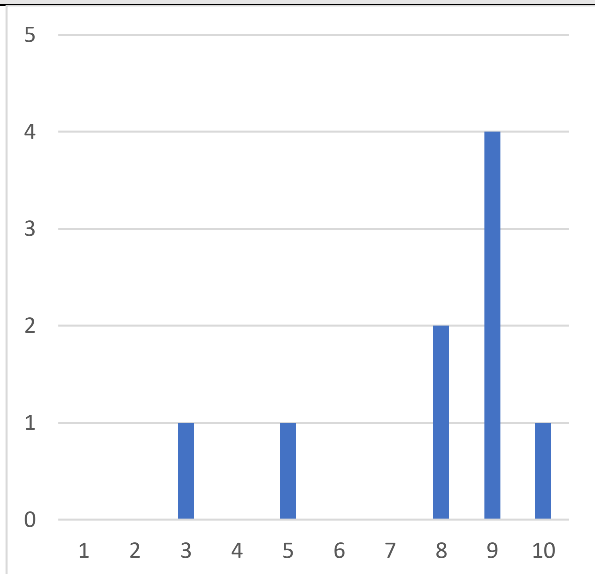
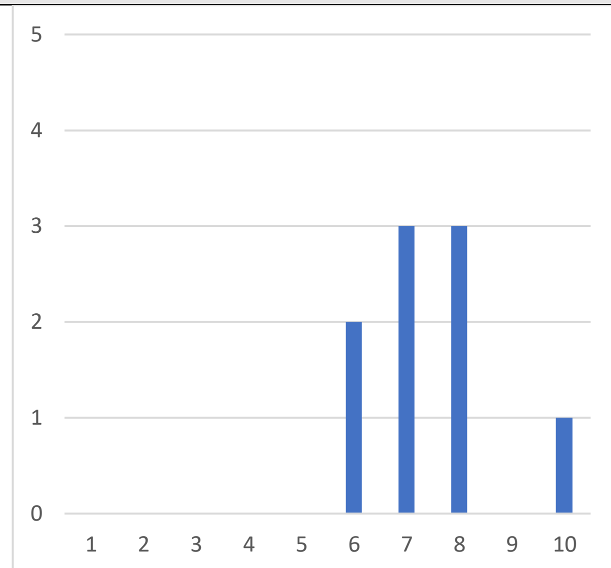
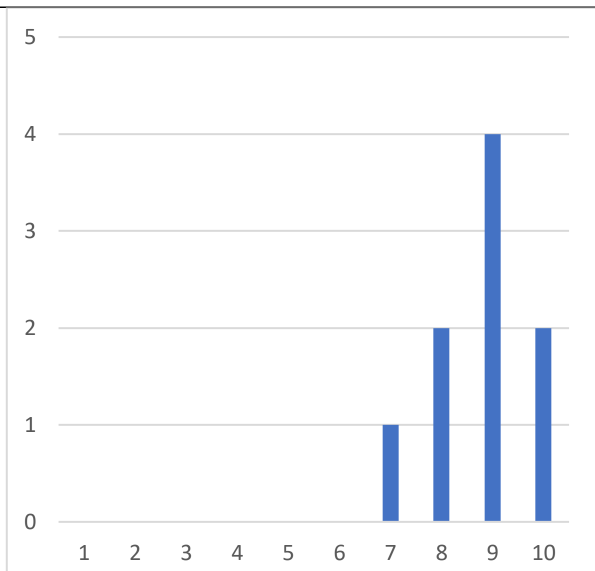
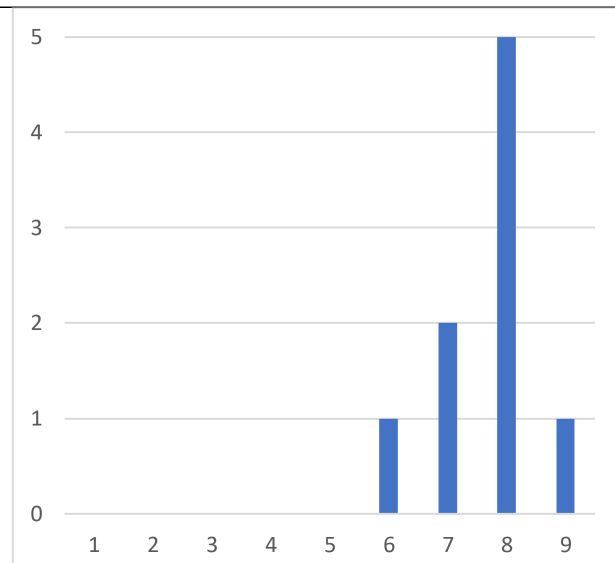
After each scenario, participants were asked to share their perceived confidence level for *handover* of the patient to the next team on a rating scale from 0 (not confident at all) to 100 (very confident).

The total average confidence for the standard interface on scenario 1 for both female and male participants is stated as 85, whereas it is 94 on average for scenario 2.
For female participants, scenario 1 was rated an average of 80 and was slightly higher for scenario 2 with 84. Male participants rated the first scenario with an average confidence level of 70 and higher for scenario 2 with 84.
Overall, it is visible that participants stated a higher confidence level on average on scenario 2 although it is a clinically more severe case. A possible explanation for this is the learning effect that participants had when they used the same interface for a second time and felt more confident.

For the ecological interface, within scenario 1, all 9 participants rated their confidence level as an average of 72 and was slightly higher with an average of 74 for scenario 2.
When looking at the average confidence ratings of female participants in both scenarios, the averages are 78 for scenario 1 and 77 with a small decrease for scenario 2.
For male participants, the average confidence ratings were slightly lower overall, i.e., for scenario 1 an average of 65 and 69 with an increase for scenario 2.
Also, the average confidence ratings for both female and male participants was overall slightly higher in scenario 2 than scenario 1.

The frequency distributions of the participants in each group are shown in Table . The 0-100 rating scale has been split into intervals of 0-10, 10-20, 20-30 etc. Thus, all ratings have been summarized on the x-axis of the graphs from 1 (being 0 to 10) to10. The y-axis shows how many participants have rated the associated scale. For example, in scenario 1 for the control group using the standard interface, 4 participants rated their confidence in the interval of 9. On the same graph, it is possible to see that participants' ratings are more scattered (between 3 to 10) when compared to the participants using the ecological interface in scenario 1 (6-10).
Among scenario 2, there is no major difference visible between both groups.

Table 27: Frequency distribution of the confidence ratings

<table>
<tr><th>Confidence</th><th>Control group using the standard interface</th><th>Experimental group using the ecological interface</th></tr>
<tr><td>Scenario 1</td><td>
</td><td>
</td></tr>
<tr><td>Scenario 2</td><td>
</td><td>
</td></tr>
</table>

Although the control group had higher confidence in their numerical ratings, it is noteworthy that they use this system daily. An explanation for the experimental group having lower confidence could be linked to unfamiliarity and using a new system. Also, during the think-aloud, participants in the experimental group showed more directed thoughts by identifying the correlations between the annotations and effects. The control group, however, was sharing all possible options they would relate to the presented case.

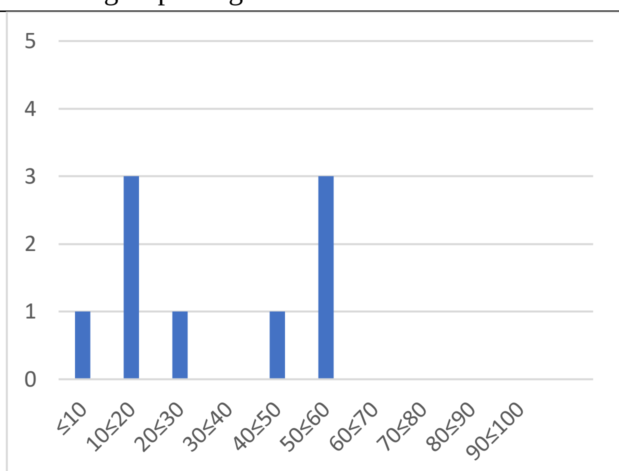
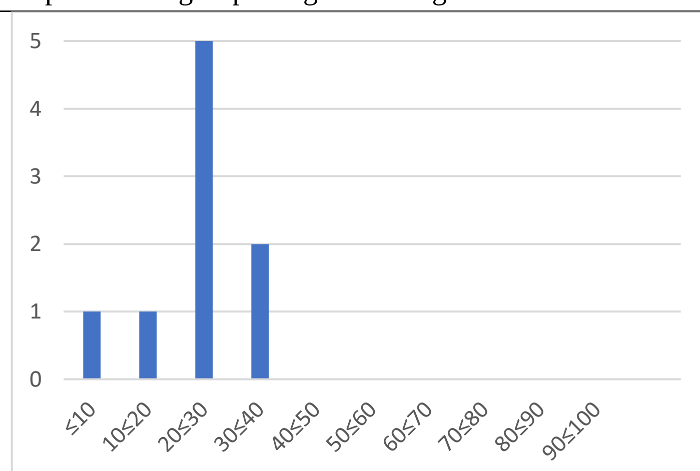
6.3.7 Subjective performance ratings

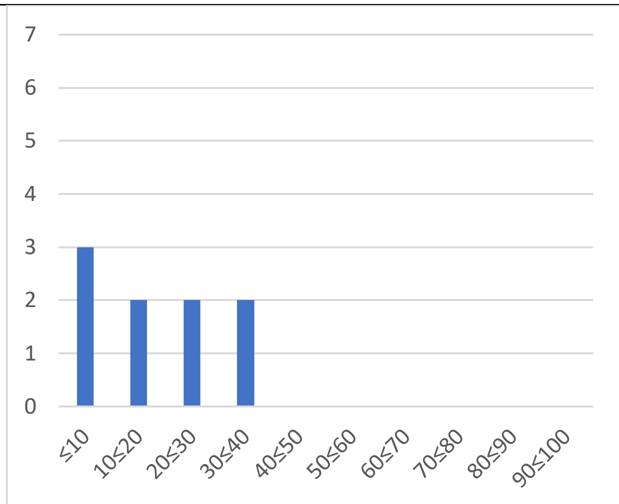
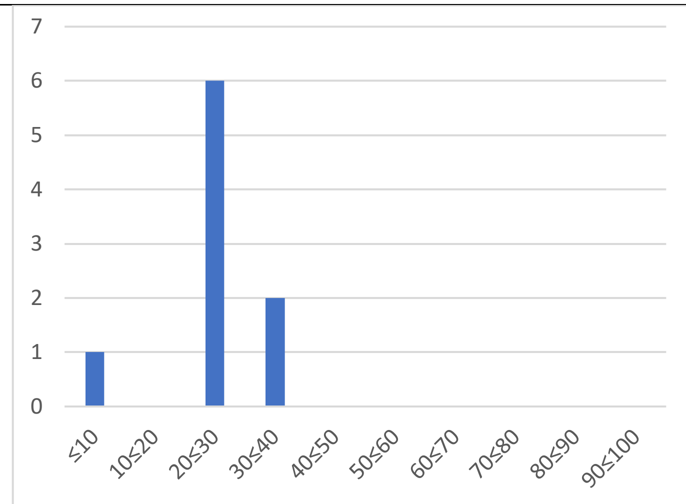
Participants were further asked to share their subjective performance ratings on both scenarios. For this, a NASA TLX questionnaire was provided to the participants. Overall, NASA TLX results are part of an extended research project and thus, only the ratings for the performance as part of the expertise development theme are outlined in this dissertation. A short definition was provided as part of the NASA TLX questionnaire: **“**How successful do you think you were in accomplishing the goals of the task set by the experimenter? How satisfied were you with your performance in accomplishing what you were asked to do?”

The scale was provided from 0 to 100, while 0 is related to great performance (or perfect) and 100 to bad performance (or failure). As other subjective workload questions were asked as part of the NASA TLX questionnaire (e.g., physical demand), performance showed the opposite rating scale. Although this has been pointed out during the study conduction to mitigate confusion or wrong ratings, four cases have been noticed that show some outliers. These outliers were reversed for the following data analysis as participants’ think-aloud and confidence levels would correlate and be more logical within this context.
Table 28 shows the frequency distribution in increments or intervals of 10. While the distribution in the control group within scenario 1 is a bit more laid out representing some ratings around 50 to 60, the experimental group perceived their performance in the higher ranges overall.

In scenario 2, the control group using the standard interface rated their performance generally high, while the experimental group perceived their performance stronger in the 20-30 ranges.
In both cases, the experimental group rated their performance on average slightly higher than the control group. However, a statistical confirmation is not possible due to the small sample size but can be investigated in future studies.

Table 28: Performance rating frequency distribution

<table>
<tr><td></td><td>Control group using the standard interface</td><td>Experimental group using the ecological interface</td></tr>
<tr><td>Scenario 1</td><td>
</td><td>
</td></tr>
</table>

Scenario 2		

6.3.8 Usability ratings

A digital usability questionnaire was shared with participants during the usability study for gathering general feedback on:

1. The usefulness of the interface and information represented on it,
2. The ease of use of the interface and related information,
3. Neurocritical care concepts and clinical information displays.

For all short statements, participants were asked to rate on a scale from 1 to 7, i.e., 1 meaning strong disagreement and 7 indicating strong agreement.

This questionnaire was given to both control and experimental group, whereas the control group using the standard interface had the opportunity to provide ratings a second time after seeing the training video of the ecological interface at the end of the study.

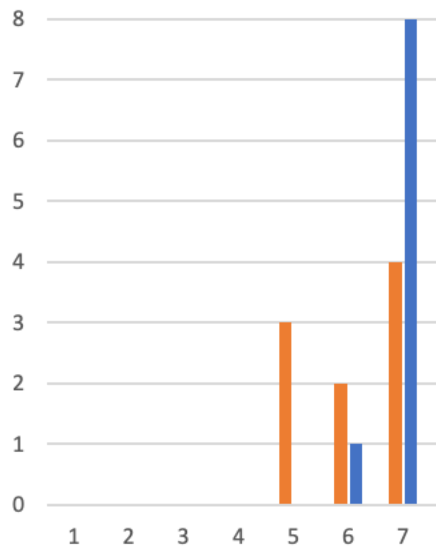
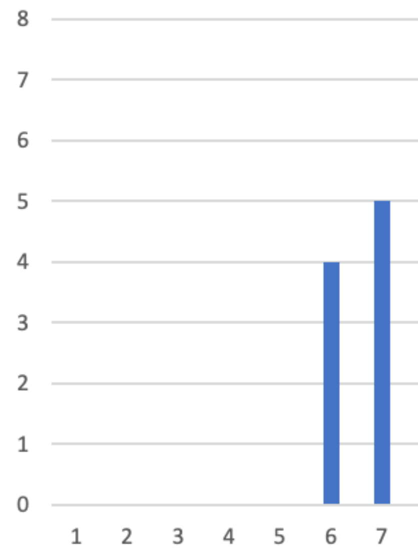
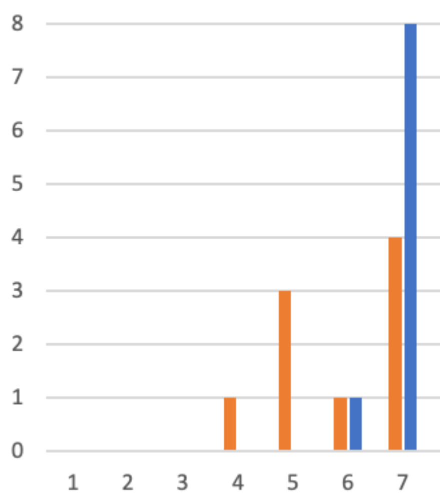
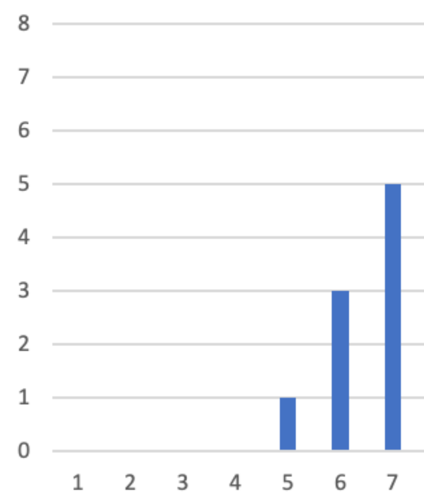
6.3.8.1 Usefulness of the interface

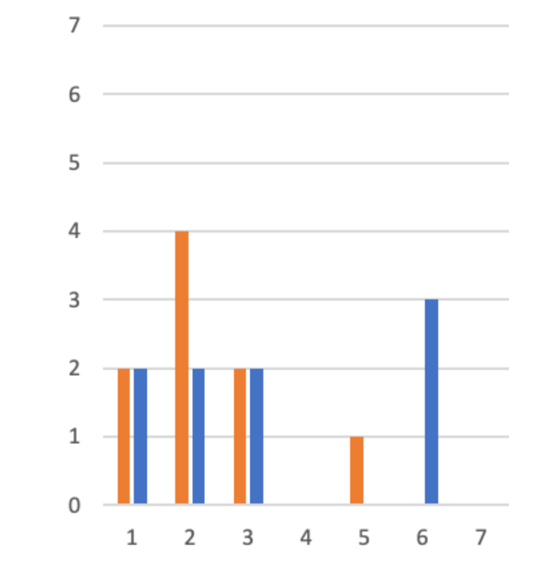
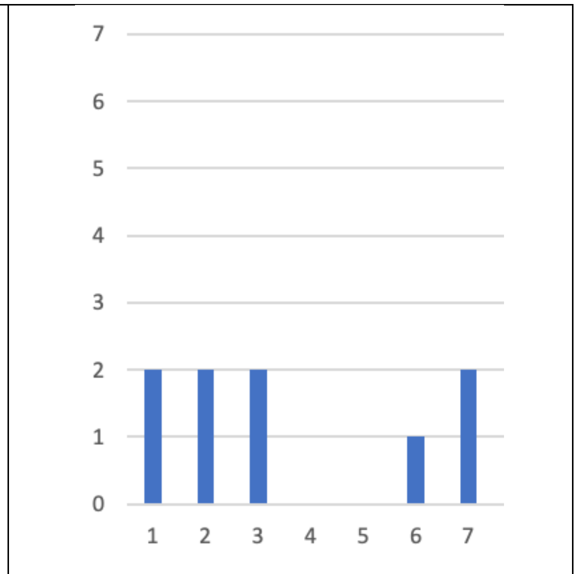
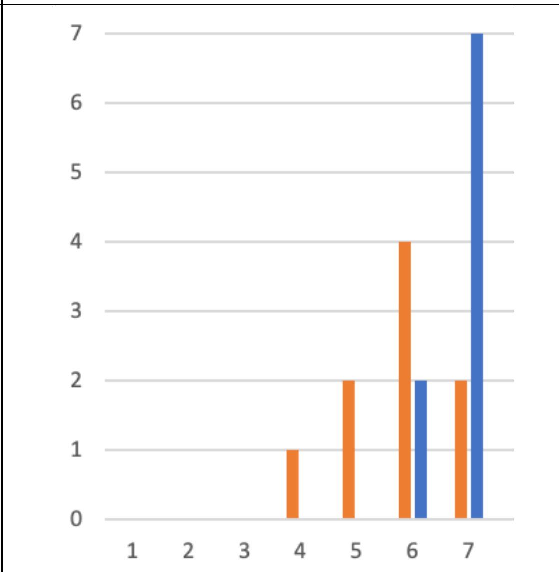
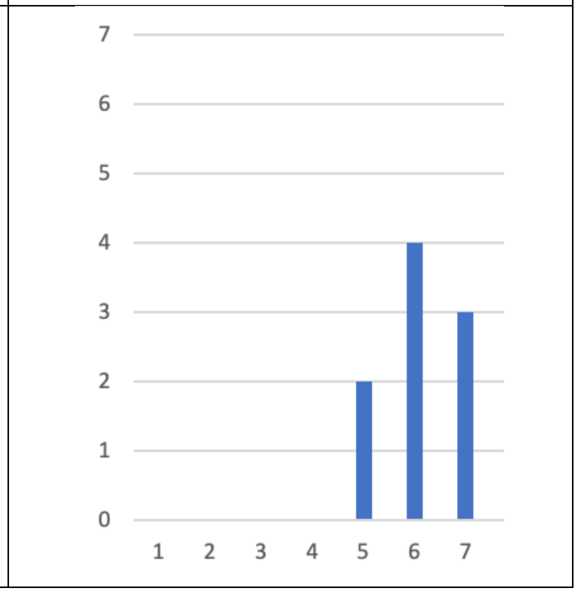
The usefulness of the interface was evaluated on statements 1 to 4 in Table 29 below (St1-St4 are used as abbreviations).

The graphs represented in Table 29 show the ratings on the perceived usefulness of the interfaces for each statement. The left column shows the frequency distribution of the control group using both standard and ecological interface, while the right column shows the frequency distribution for the experimental group. The frequency distribution in the left column (control group) shows the ratings on the standard interface in orange, while the blue bars represent how the same group rated the ecological interface later.
The control group usually rated the ecological interface higher on the four statements. This means that they find the ecological interface provides data in a more meaningful way, supports decision-making better, contains some information that is irrelevant to their decision-making process, and the information on the ecological interface is more useful to inform understanding of the patients' situation compared to the standard interface. These ratings are also in most statements higher than the ratings of the experimental group. The experimental group has not shifted back to the standard interface during the study, they would relate to their daily experience in the hospital from memory.

Also, three experimental group participants rated on statement 3 that some information on the ecological interface is irrelevant to their decision-making process. In the discussion at the end of the study, some participants provided feedback on the overall usefulness of the interface.

Table 29: Usefulness of the interface

Statements	Frequency distribution – control group	Frequency distribution – experimental group
St1: The interface presents data in a meaningful way		
St2: The interface supports my decision making		

St3: The interface contains some information that is irrelevant to my decision making process		
St4: The information on the interface is useful to inform my understanding of the patients' situation		

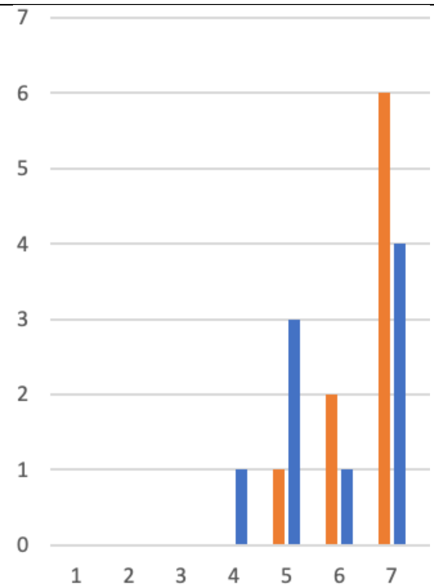
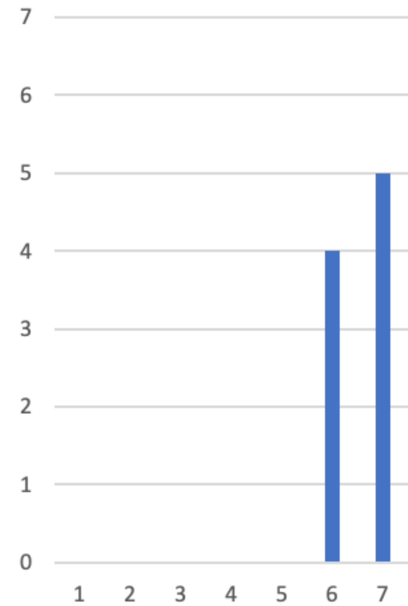
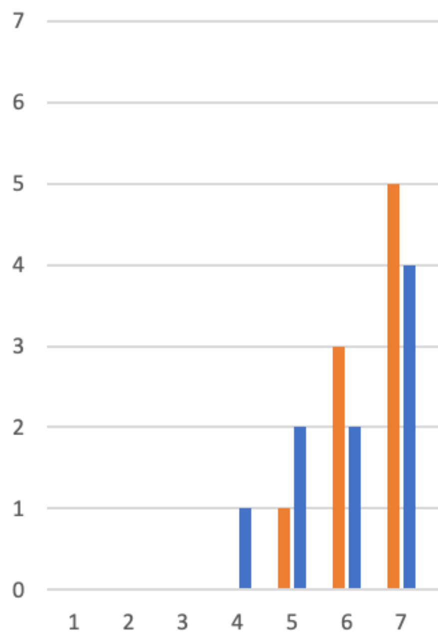
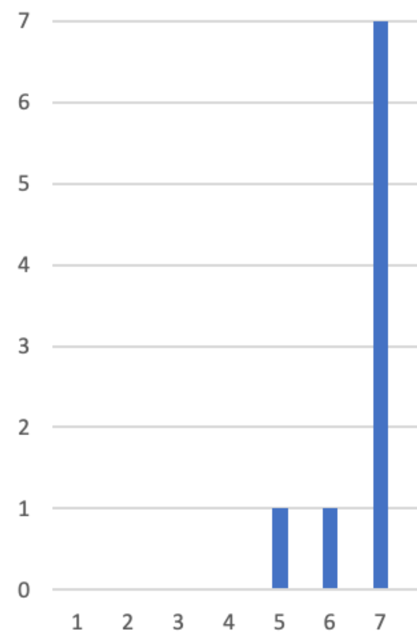
6.3.8.2 Ease of use of the interface

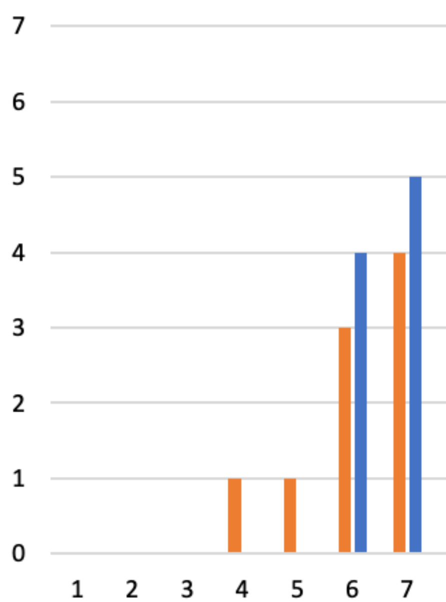
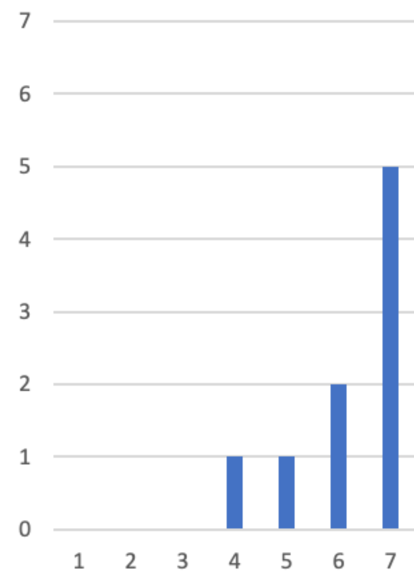
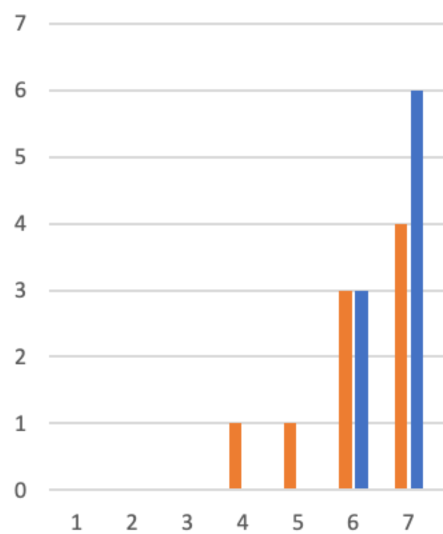
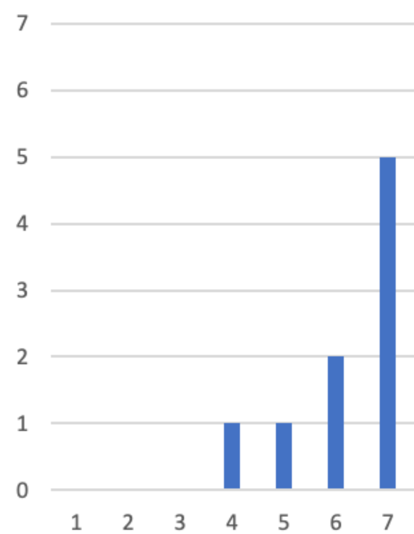
The ease of use was evaluated by showing statements 5 to 8 (St5-St8) in Table 30.

Again, ratings between the control group using the standard interface first and the experimental group looked quite similar overall. However, the experimental group rated higher on the ease of finding information needed on the ecological interface compared to the control group using the standard interface.

When the control group used the ecological interface later, the ratings generally for St5 and St6 were lower compared to their initial ratings on the standard interface showing they find the ecological interface less easy to use and less easy to find information they needed. However, there were also some ratings on St7 and St8 that indicated that information organization of the ecological interface makes more sense to them (St7). Also, the sequence and flow of different screens on the ecological interface was perceived more logical by some participants (St8) which is similar to the ratings of the experimental group using the ecological interface only.

Table 30: Ease of use of the interface

<table>
<tr><td>Statements</td><td>Frequency distribution – Standard interface</td><td>Frequency distribution – Ecological interface</td></tr>
<tr><td>St5: The overall interface is easy to use</td><td>
</td><td>
</td></tr>
<tr><td>St6: It was easy to find the information I needed on the interface</td><td>
</td><td>
</td></tr>
</table>

St7: The organization of the information on the interface makes sense to me		
St8: The sequence and flow of different screens on the interface was logical		

6.3.8.3 Neurocritical care concepts and clinical information display

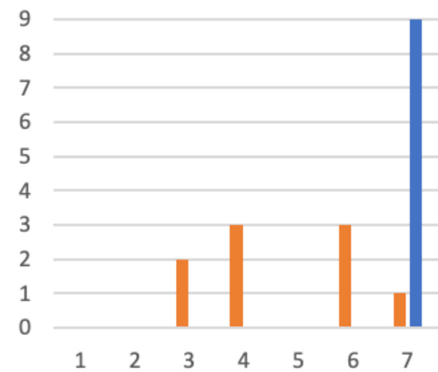
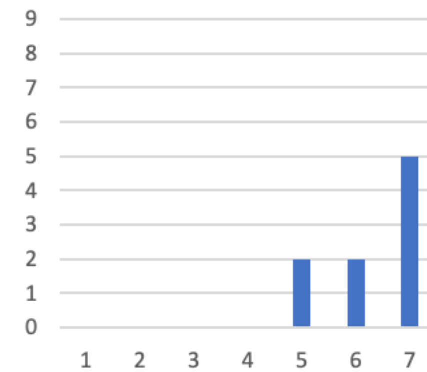
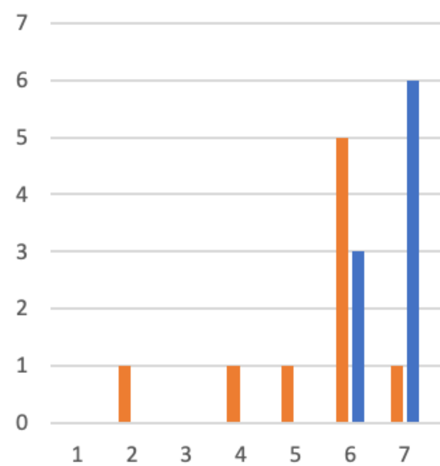
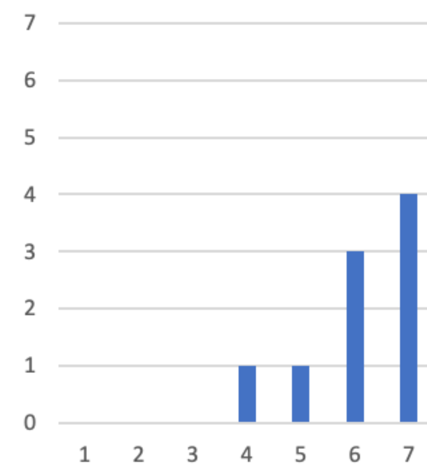
This part was evaluated based on statements 9 to 13 (St9-St13) while St13 was only presented to the experimental group (and St13 additionally to the control group using the ecological interface). The frequency distributions on each statement are represented in Table 31.

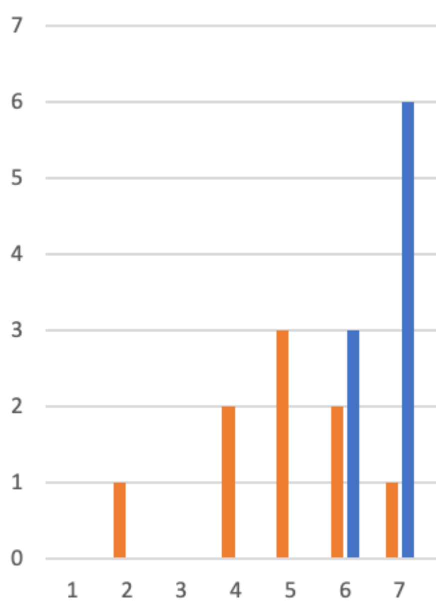
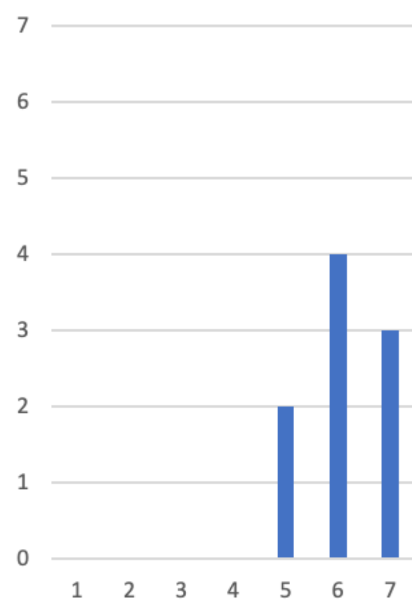
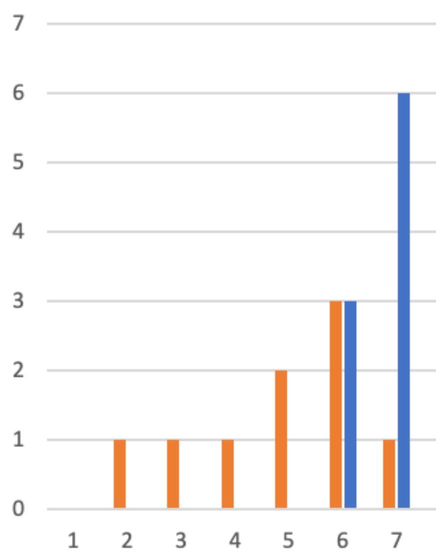
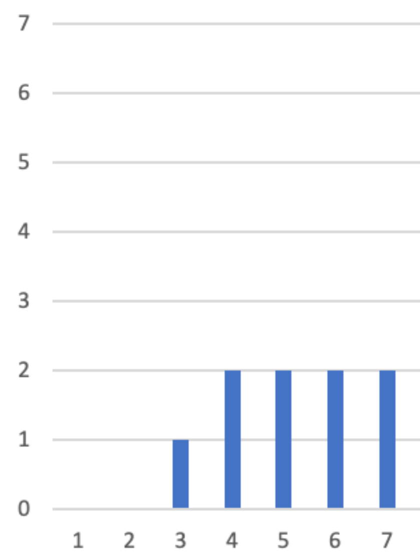
The concept of ICP (St9) was rated higher by both groups using the ecological interface compared to the standard interface.

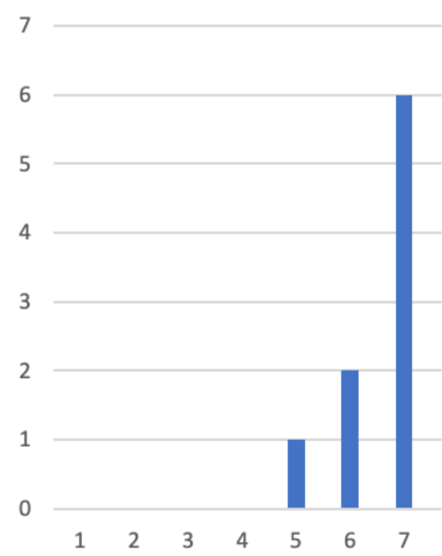
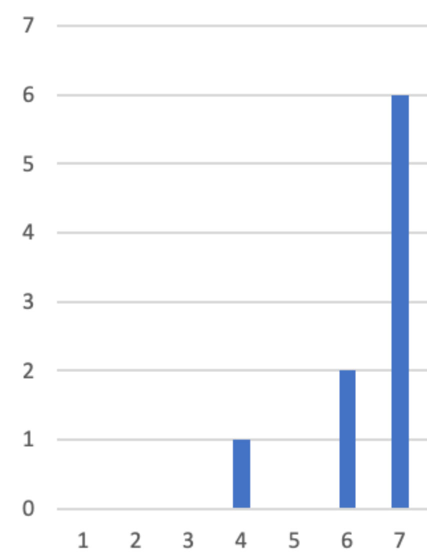
For St10 and St11, it is noteworthy that the experimental group shows slightly higher ratings related to the relationship between ICP and other vital signs, as well as better understanding of trajectory when using the ecological interface compared to the control group only using the standard interface. However, there are no big differences in the individualization of treatment among both groups; it is more distributed across 2-7 ratings for the control group using the standard interface, and 3-7 for the experimental group. The control

group using the standard interface and the experimental group both provided lower ratings for the individualized care component (St12). When the control group used the ecological interface later, the ratings increased and agreed more that individualized care is possible with the ecological interface.
Most participants have rated ‘high agreement’ to state that current ICU displays require improvement (St13) When the control group additionally rated the ecological interface, for statements 9-12 it was generally higher than their ratings for the standard interface and also higher than the ratings of the experimental group. Opinions on St13 were almost identical; both control and experimental group think that the current interfaces need improvement.

Table 31: Neurocritical care concepts and clinical information display

Statements	Frequency distribution – Standard interface	Frequency distribution – Ecological interface
St9: The interface helps me understand the concept of ICP burden		
St10: The interface helps me think about the relationship between ICP and other vital signs		

St11: The interface helps me understand the trajectory of the patient so I can better provide treatment (in regard to neurocritical care)		
St12: The interface helps me think about an individualized treatment approach for the patient (in regard to neurocritical care)		

St13: I think that the current information displays in the ICU need improvement (in regards to neurocritical care)		

6.4 Discussion

This discussion aims to highlight the challenges and limitations with various aspects.

6.4.1 Reflections on the interfaces

Firstly, during the study preparation, the decision was made to use a standard interface that's represented in many ICUs especially in Ontario, Canada. The main intention for that was to get the common and most representative experiences of study participants in the region where the research group is based, and most participants were expected to be recruited from this location. Thus, the closest design to the commonly used Philips IntelleVue model was created. In future studies, it can be further researched which bedside physiologic monitor models are commonly used across larger locational scales to lay a basis for a usability study where new (ecological interface) designs would be tested upon. The recreated IntelleVue interface was having functional limitations as we were not expecting those areas to be of interest when customizing the interface for the purpose of this usability study. However, it is important to consider carefully what features are of interest and might impact study results. We further hypothesized that additional tabs on such vital signs screen would be feasible to extend the common view which would be required to check with the company guidelines and regulatory aspects.

6.4.2 Reflections on the scenarios

The scenarios created were limited in a way, i.e., the data set was made up for a 24-hour capacity with data points for each second. The usefulness can be further investigated in future studies to identify various scenarios where longer and shorter data sets could be used to see how that impacts patient individualization, trajectory, risk perception and more. Also, the amount of neurocritical care scenarios can be expanded. In this study, we have included two different pathologies on two different severity and complexity levels which could be explored in more dimensions: providing more details, different data scales, dealing with noise (this was excluded in our scenarios), and limited overall information on further medical conditions or access to imaging, blood work etc.

6.4.3 Reflections on the flowsheets

The flowsheets provided were based on a paper flowsheet format that is used in the University Health Network hospitals in Canada, and variations of that exist which may have impacted the way of how quickly participants were able to orientate and find information they're interested in. The flowsheets were not fully filled out and limited to the potentially useful or standard information required for the scenario understanding. The way how flowsheets are filled out is another component to discuss; the flowsheets presented in this study showed data filled out in a digital format without specific notes or colours that may be used by nurses in certain hospitals to highlight outlying aspects. At the beginning of the study conduction, the researcher held back of interfering in participants' overall way of thinking out loud and clicking on the interface. It was also noticed that participants sometimes forgot to use the flowsheet in their decision-making, which could be related to the high frequency of omissions. An overall assessment of the cases was still possible even without the flowsheet but led the participants to disregard details that might have been relevant. Thus, the researcher reminded the participants at the end that relevant information can also be found on the integrated flowsheet in case the participants deem it relevant too. During the training, participants got the chance to explore the interface, interactions, and limitations to prepare for the scenarios, but this has still been a point that required reminders. In future studies, it can be highlighted additionally just before starting off with the scenario, to use all accessible data points or ask the participants to verbalize that they do not require the flowsheet information. Participants forgetting to use the flowsheet could be related to the fact that they commonly use a paper flowsheet that would be tangible, or they commonly use a separate system to open up the flowsheet on a separate screen in daily practice. If a simulation-type of study is conducted, the flowsheet can be provided in paper or online format that is used in its common way.

6.4.4 Reflections on the recruitment process

Another component that can be discussed is the recruitment process. In this study, the author of this dissertation tried to recruit participants through the collaboration with Dr. McCredie who has forwarded the recruitment email to program directors, trainee lists and experts directly especially for the interview study. For the usability study, the interview participants were a great source to reach out to again and ask for interest in participation. However, the number of participants was still lower than expected and was extended through asking organizations, reaching out to trainees directly at the critical care forum in Canada, and through the author's personal network. Getting a more international group of participants could be aimed for while investigating comparable systems or checks for their background training or cultural aspects that could play a role.

Also, during the evaluation process, it would be ideal to get a committee of experts discuss the findings and evaluate error perception etc.

In terms of demographics, we noticed that the years of experience were a factor of inclusion which made it challenging to find participants accordingly. As expertise and experience varies across participants, we have only used certain inclusion and exclusion criteria that can be extended further in future studies. For example, it can be explored if participants with a specific specialization have a different approach to dealing with the scenarios represented.

When the author of this dissertation extended the recruitment process, some trainees with a neurology residency were also considered to be a group of interest as the scenarios dealt with neurologically critical patient cases and participants had multiple years of experience working in the ICU. This is an example for recruiting compatible participants that are being trained in a different system and can be considered for future studies.

Furthermore, this study focused on trainee physicians but in future studies, other clinical personnel who use the interface can be taken into consideration too, such as nurses or neurosurgical trainees.

6.4.5 Reflections on the measures

The selected measures for the usability study can be explored further; only certain expertise development measures were taken for this usability study but can be extended to a bigger pool of measures that relate to the development of expertise within this specific field.

The conduction of the study in its full length of 90 minutes could have been a factor in having difficulty finding participants, too. As the study was conducted during the pandemic phase, it was changed to use a virtual participation opportunity rather than a simulation study that was initially considered. The virtual option on the one hand made it accessible to participants to join the call at any time and day that suits them best after or before a work shift. On the other hand, we lost all the environmental aspects that could have impacted the participants' perception and feeling of being surrounded in a close to real setting as in a simulation study.

The measures taken for assessing the usability and perception were based on a rating scale for which a bigger participant group would be considered effective to understand the detailed differences. In this study, the ratings were still included and provide a basis for a general impression that should be further explored in the future, as statistically it is not possible to state specific findings from the small data set. Follow-up studies should consider a larger participant pool and also use more quantitative approaches in case access to a larger group is enabled.

Measures used to assess confidence, for example, could be further deepened as we noticed that the participant group is very international and might have different experiences being a trainee in a different location and language. Although participants often mentioned that they felt more comfortable and confident when confronted with the interface for the second time, i.e., in scenario 2, we were expecting that they might find this scenario more challenging because a specific pattern of the waveform trend was implied (i.e., the Lundberg A-waves) and the scenario was laid out for more acuity as the physician is being called in due to some changes. This may show that their feeling of familiarity with the interface impacts their perception on confidence in handing over the patient to the next team.

Also, when rating the usability aspects of the interfaces, some participants mentioned that they would like to re-rate their previous choices for the standard interface after seeing the ecological interface, because they had a much better impression about the ecological interface, and they didn't know that better interfaces could exist.

The study showed short-term effects among the two groups; in future studies, a more long-term approach can be considered to evaluate the use of the interface and their thinking pattern over time, and its impact on mental model development. Although tendencies can already be seen in a short time, the long-term approach would enable us to deepen and explore potential variety of impressions and learning effects.

In future studies and continuation of this research work it is recommended to further link the next steps with conformity of standards, usability regulations, functionality and compatibility with other devices that the interface would be connected with in case various signals would be processed.

6.5 Conclusion

In this chapter, a usability study was outlined in which the ecological interface was compared to a standard interface of a bedside physiologic monitor commonly used in the ICUs in Ontario, Canada. The ecological interface was designed to especially support expertise development among neurocritical care novices. The measures taken for such assessment are linked to the previous chapters where expertise-relevant indicators were outlined such as the number of errors, severity of errors, reasoning and SA, reflections, confidence, and subjective performance. Usability was also taken as a measure in this study to get more feedback on the ecological interface usefulness, ease of use and extent of capturing visualizations addressing important neurocritical care concepts.

In this study, we have investigated aspects of expertise development as short-term effects or initial insights of what aspects of the interface can help to develop expertise in neurocritical care.

The study was conducted within 90 minutes with 9 participants in each group (the control group using the standard interface and the experimental group using the developed ecological interface). This study can be extended to a multiple-day or long-term approach in future study designs.

The main question for this chapter was to identify: to which extent does the ecological interface support expertise development (in the short-term) when being compared to a standard (currently used) interface? This high-level question has been split into various sub-questions.

6.5.1 Findings of the usability study summary

First, together with a neurocritical care expert, **strategies** for neurocritical care cases were analyzed. While experts have a concurrent management approach, novices handle processes in critical care rather sequentially. Further, it is possible to look at strategies between or within patients. In the usability study presented in this chapter, participants focus on a single patient (within-patient) at a time and do not have to divide their attention among various patients requiring support.

When investigating the **errors** participants made among both groups, it was noticed that the number of errors was higher in the control group as hypothesized. Also, it was noted that all participants omitted information relevant for decision-making. These were also higher in scenario 1 for the control group. When proposing a 6-number harm scale, the expert evaluator mentioned that there could be different views on classifying certain errors strictly on that scale. Thus, he suggested the inclusion of numbers in between (e.g., 2-3, 3-4) for which the distribution did not clearly show if the experimental group was making less severe errors or omissions.

When summarizing findings of **reasoning** and **SA levels**, some differences across both groups were noticeable;

Level 1: Comparing responses of scenario 1 and 2 among the two groups on level 1-SA, it is noted that the experimental group made more observations related to the ICP uptrend and the response to therapy. In scenario 2, risks they perceived were described during their initial glance too. More participants in the control group, however, noticed the waveform of ICP and shared immediate ideas for medical therapy.

Level 2: When both groups shared their concerns on the provided cases, the experimental group again pointed out the response to therapy, the correlation with other variables and the ICP pattern (in scenario 2 especially). The control group mentioned the response to therapy at a later stage and shared more thoughts on diagnoses.

Level 3: When participants shared their assessment of risks (short and long-term), the control group provided a wider range of risks overall compared to the experimental group that also related to infections in hospitals etc. The control group thus shared a more general approach on potential risks and the limitations of flowsheets not providing enough information to prognosticate. The experimental group, however, perceived risks also in combination with procedures and focused more on the control of ICP, autoregulation and individual variables such as age playing a role in brain swelling time.

Both groups shared general treatment options that were similar, however, one participant in the control group added that the transcranial doppler could be used and overall, the relationship between variables would be considered important. The experimental group focused on the maximization of therapy and letting the patient's family know about the situation.

Level 3: When inquiring about priorities and treatment plans, the control group pointed out various considerations and targets of physiologic variables and how that could be facilitated through medical therapy. The experimental group, however, proposed less variety of medication options and summarized it generally as a TIER 1 and 2 therapy requirements. Overall, the experimental group also tackled more the reasons for the ICP trend (scenario 2) and proposed to treat more aggressively in scenario 2.

Level 3: When specifically probing participants about whom they would communicate to about the patient cases shown, both groups mentioned similar clinician (groups) such as neurosurgery, radiology etc. At this point, also the control group mentioned they would talk to the patient's family.

Level 3: In scenario 1 handover, both groups summarized the ICP uptrend while the experimental group pointed out the correlation with a decreasing CPP and increasing HR level, and a poor waveform compliance of ICP. The control group had shared at the beginning of their think-aloud that there are ICP waveform compliance issues, however, they didn't point it out again during the handover summary. Both groups mentioned responses to treatment and risks they perceive. The control group listed more details to the risks they consider. The experimental group focused more on previous procedures while the suggested therapies were similar among both groups. In scenario 1, the control group considered coiling while the experimental group highlighted that they would consider it if there were no other options left anymore. In general, the type of medication suggested by both groups were not always aligned either. In the handover summary of scenario 2, the control group focused on a general overview and overall assessment (injuries, physiologic variables etc.) while the experimental group proposed to try all possible strategies first and treat more aggressively in this case.

Overall, important analyses have been outlined among all three levels, however level 3 was detailed out. This is also related to the probing questions asked when participants did not share further details during their think-aloud. Tendencies of the experimental group are more visible when trying to consider the reasons for the ICP trend or pattern more than the control group does. Trying to correlate various variables and looking more at the trend helps to shape a more directed trajectory. The control group participants shared overall relevant treatment options and risks for patients with high ICP.

The **reflections** among both groups on their strategies helped them to think about aspects they might have left out when going through the cases or suggesting other approaches or changes to their proposed plans.
Various insights were gathered through the reflections after each scenario on the chosen strategies and potential limitations. Both groups reflections were rich and only showed slight differences in the themes. The main themes related to: targets provided by experts, focus on numbers, lack of information, use of systemic approaches, consideration of alternative medication, consideration of future steps, consideration of overseen aspects, consideration of additional measurements.

This can be briefly summarized in key points for the control group:

- It was stated that they are usually being told by the staff physicians what the targets are for the patient.
- Another comment outlined that they didn't look at the ICP waveform because in the hospital, they focus more on numbers.
- A more systematic approach was considered to double check all quick assessments: the approach is called ABCDE.
- It was noticed that the CT was actually not needed although this was mentioned it a few times during the scenario that it would be needed. However, the scenario description provides the necessary information and doesn't require to check the CT again.
- Other reflections related to the absence of the patient's response, so it is hard to assess this patient.

- A PbtO2 monitor could be added. However, it was noticed that the participant is localizing, so the ICP monitor would be sufficient.
- Reflections on medication to be provided after day 7 were made.
- Another participant said he wanted to do the neuroexam himself. He also reflected that he is surprised about the aneurysm not having been treated earlier. He further mentioned that he would have focused more on ICP and CPP maybe giving more sedation in the next 1-2 days. He added some reflections on limitations; as one of the presented scenarios is a TBI case, the patient could have trauma anywhere and he focused mostly only on the brain.
- Further reflections showed that the other pages of the flowsheet may be considered. More sedation could be provided, and neurosurgery could be notified. The state of autoregulation seemed unclear, especially for the second case. A Licox monitor might be of help too.

Some themes were similar also for the experimental group while some differed: consideration of more aggressive treatment options, consideration of further risks and providing medication, the lack of information, consideration of overseen aspects or ruling out risks, realization of ICP focus, consideration of the flowsheet, realization of overseen variables, consideration of additional measurement and consideration of future steps.

Reflections made by the experimental group are summarized as follows:

- Considerations for more aggressive treatment could be made early on. Medication could be added, and EEG measurements are taken in some hospitals. There is a risk of treating a patient in a preventative manner. More details on the EVDs could be checked. The importance to look for the cause and maximize medical therapy is considered.
- Imaging is lacking due to some further reflections that would show the nuances of the case.
- The risk of vasospasms and rebleeding could have been added during the handover.
- Another participant mentioned that they usually get more information from the nurse. If after all efforts made the ICP is still high, talking to neurosurgery would be considered.

The opportunity to reflect on participants’ strategies taken helped to see if they would potentially consider other alternative treatment plans as well. Giving novices time to reflect on their actions, thinking about limitations and changes was a way of fostering and critically thinking about what could be done differently. Many of the insights they provided are very valuable and showed that there is not always a single way of analyzing or treating a patient. Overall, many of the reflections related to changes in medication or adding other tools for assessment. Some noticed that it would be useful to add more details, e.g., during the handover, sharing more risk factors to be aware of. All of these considerations are valuable.

Later, more reflections were outlined when showing the ecological interface to the control group. They were asked to reflect on how their assessments or plans would have changed if they had used the ecological interface instead of the standard interface. Three themes were raised: consideration of more aggressive treatment options, consideration of autoregulation, and consideration of new targets for different variables. Major statements are summarized below:

- The trends are clearer on the ecological interface.
- There are concerns about autoregulation being impaired. The ecological interface also prompts to learn and treat things in a different way.

- Another participant reflected that much more information is visible on the ecological interface and all of it is very relevant. He also said that waveforms could be assessed more if there was a zoom-in function. He would look for a new target of CPP in case the basics are not working well, and the patient is still not better. When seeing the Pressures tab, he identified the Lundberg pattern and mentioned treating more aggressively.

These reflections show that participants started reflecting about various aspects of the scenarios and their applied strategies and considerations on the cases. While the experimental group reflected on considering an even more aggressive treatment plan, this was also mentioned once the control group saw the ecological interface. This implies if having more insights into the data, they act differently. Considerations of autoregulation was highlighted by the control group using the EID and shows how relevant the visualizations are in supporting the understanding of major neurocritical care concepts. The fact that the control group using the EID reflected on new targets for different variables further highlights that they perceive a different level of risk or individualization considerations for various variables.

Further, reflections were shared on the overall study.
Reflections were given on the ecological interface, being useful and supportive in confirming reasoning and conclusions, making it foolproof. The ICP summary is helpful, but some mentioned different preferences for visualizations e.g., a pie chart. Aspects of the burden or peaks are better to recognize compared to a flowsheet. The colours are also cues for looking at atypical signs also from a distance. The ecological interface is perceived as helpful to prognosticate and come up with a trajectory. The annotations are useful to see everything in context.
Multiple participants stated that the EID helps to learn; learning by yourself is possible and there is a fast learning curve. Certain parts of the interface can also be used by experts daily (e.g., the summation graph) rather than novices.
Further, one participant said that the targets could help to share a mental plan and be used for urgent situations, but he wouldn't need the different colours. Another participant also pointed out that the ecological interface is great for handover and care takers, securing communication. The ecological interface can facilitate academic research and can be used for training.

Reflections also related to areas that could be improved on the ecological interface, e.g., other variables could also be integrated such as PbtO2 or EEG data.
Some comments related to autoregulation; participants in the study mentioned that it's often a concept coming up in exams, but it is unclear how to handle it. Others stated that it was more obvious to recognize it through the ecological interface.

More ideas were shared that could help improve the interfaces.
The control group mentioned a few aspects that could be improved on the interface for the next versions to be used in a study. This feedback related to a thicker representation of the HR and ICP line on the vital signs tab. The flowsheet could benefit from a bigger size as well. Some variable's timescales were not aligned, which should be improved. If possible, zooming interactions could be helpful and seeing etCO2 as a waveform as well.

The experimental group shared feedback on improving usability for example reminding more clearly on the meanings of the colour ranges, the use of the flowsheet, making the red and dark red colourings stand out more and potentially changing the combination of visualizations as too much information is presented. The default for all tabs could be the same and more clinical information could be integrated (e.g., access to the CT). Also, more details on the previous treatment could be included.

Another measure taken in the context of the usability study was **confidence**.
Although confidence ratings were higher on average for the control group compared to the experimental group, the transition from scenario 1 to scenario 2 yielded that participants in scenario 2 were more confident across both groups on average. The distributions varied slightly across the two groups; in scenario 1, the control group ratings are more scattered (ratings between the intervals 3-10) while the experimental group was around the interval 6-10.
Scenario 2 ratings overall were similar among both groups, while the control group had a greater frequency of ratings on the interval 9-10 and the experimental group more in the interval of 8.
Overall, there are confidence rating differences but both groups' ratings were not extremely different. The ecological interface may be lower on average compared to the standard interface averages at this point, which shows that participants are still more confident using the interface they have been using for a long time. However, it is noteworthy that the experimental group still showed high confidence levels that weren't too different overall and it should be tested further in the long run, when users of the interface get more familiar with the interface itself. The average rating increasing among both groups for scenario 2 is an indicator for the rising confidence by only a second time of use of the interface. More cases could be given to participants in both groups in future studies, and observed if confidence changes by the quantity of cases or time of familiarity with the interface.

Performance was pointed out as an additional measure. This has been considered as a subjective performance rating. The opposite rating scale was used here and showed higher perceived performance when ratings were closer to 0.
When participants rated their subjective performance on both scenarios, again, the control group shows a greater range for their ratings (0-60) while the experimental group lays around 0-40. More ratings were captured around 20-30 for the experimental group. This shows that more in the experimental group rated their performance higher than the control group. This could be linked to aspects such as a better understanding of the patient data and analysis.
In scenario 2, the distribution of the experimental group's ratings is again higher in in the interval of 20-30, while the ratings are more evenly distributed between 0-40 in the control group. The distributions highlight that there are no major perceived performance differences between scenario 1 and scenario 2 among both groups, however, the experimental group has a stronger rating interval of 20-30 than the control group.
The performance levels of both groups weren't perceived too differently. The distributions being wider for the standard group compared to the experimental group may relate to individual differences of interface usage in different hospitals. This aspect especially needs further investigation as the number of errors was higher in the control group that show parallels to the greater range in perceived performance.

At the end of the study, **usability** as another measure was considered. Usability has been divided into three aspects: the usefulness of the interface, the ease of use of the interface, and the deployment of neurocritical care concepts into visualizations on the interface.

Usefulness of the interface: Overall, the control group got to rate both the standard interface and at the end of the study, the ecological interface. The ratings show that the control group rated the ecological interface higher overall when compared to the standard interface. The experimental group did not show great differences either. When participants were asked to rate if the ecological interface contains irrelevant information to their decision-making, there were some participants in both control and experimental group who think there is information represented as being irrelevant. The ecological interface shows ICP in various representations on all tabs. During the development of the visualizations, they were considered to be complementary, however, participants may have other perceptions and could decide for a treatment also only using one or two of those visualizations instead of all.

Ease of use of the interface: The ease of use perceived by the experimental group was higher on some statements such as finding the information easily on the ecological interface compared to the control group using both interfaces. The control group using the standard interface first and then the ecological interface rated higher for the ecological interface when sharing their opinion on the organization of the ecological interface making sense to them as well as the sequence and flow of the interface being logical. However, they rated the ecological interface slightly lower on the ease of use and ease of finding relevant information. As all of these are initial impressions, the ease of use is expected to change over time when participants use the ecological interface more frequently.

Neurocritical care concepts and interface design: These ratings related to thinking about the neurocritical care concepts being represented or connected to the interfaces they used. These concepts focused on being able to see the relationship between ICP and other vital signs, think about trajectory, individualize treatment and the necessity of improved interfaces in neurocritical care especially.
Overall, the ratings for the ecological interface among both groups were higher. One statement related to the individualized treatment approach that would be supported when using the assigned interface for which the ratings were distributed in a range of 2-7 for the control group using the standard interface, and 3-7 among the experimental group. This shows that there are various perceptions on the support of both interfaces on individualizing care. However, when the control group used the ecological interface, they noticed a greater support of the ecological interface in this regard.

Although we have a small sample size for any quantitative assessment, there are some characteristics visible that we have been hypothesizing at the beginning of this work. The perceptions that the ecological interface in its usefulness and ease of use not being too different to the standard interface is positive as the intention was not to overcomplicate but support finding relevant information and relationships among variables. This becomes clearer especially when the control group first tries the standard interface and then provides even higher ratings when using the ecological interface. The overall representation of neurocritical care concepts seems to have fulfilled its purpose; the expectation was to let participants relate more to core neurocritical care concepts and display the information in a logical and useful way for them to make decisions. One element that might still need improvement on representing this through the ecological interface, could be the idea of individualizing treatment. However, the scenario information did not provide many individual characteristics to take into consideration such as multisystem issues or other health conditions like diabetes etc., that would require physicians to take into consideration carefully as an addition to their usual treatment plans. In future studies, more cases with more individual patient characteristics can be included to further evaluate this aspect.

Chapter 7 General outlook and contributions

This dissertation's final chapter takes us back to the overall summary of the research conducted, the limitations of the studies, connections of the results to the initially proposed research questions and hypotheses, and outlines contributions to science, technology, education, policy, society, as well as economy.

Some final words of the impact of Covid-19 on the research will be shared as a last section.

7.1 A summary of the overall work

This work disseminated multiple research questions related to challenges faced in the chosen work domain (neurocritical care), expertise development in this dynamic environment, understanding and supporting expertise development through interface design and its evaluation through interviews and a usability study.

Overall, some of the previously outlined hypotheses related to the various challenges existing in neurocritical care that especially novices face when they develop expertise. The potential of improving technology and data representations on bedside physiologic monitor interfaces can further be investigated, while including the expert and novice critical care physicians' perspectives can pose great value in the design process. Thus, it is important to explore and organize the challenges faced in neurocritical care through various perspectives (e.g., by observing with a more objective lens and by asking key stakeholders for their subjective thoughts) and discover ways to mitigate these.

There are various ways of designing interfaces; in this work, a more systematic approach is rooted in the CWA framework, analyzing multiple aspects of the work domain, tasks, and strategies to develop ideas for designing a so-called ecological interface. This approach is especially relevant in complex sociotechnical fields which relate directly to the neurocritical care environment.

Although experts try to support novices to overcome challenges during their training, a more in-depth understanding of expertise can be beneficial. This part of the dissertation related to general ways of developing expertise in critical care specifically but also guided through aspects that can be linked to interface design supporting this process. Through discussions with experts in the field and the models created through CWA, important neurocritical care concepts were outlined that novices have to acquire to develop expertise. These were implemented on static visualizations first.

Semi-structured interviews were then conducted with novices and expert critical care physicians to gain deeper insights on what can further be improved on the visualizations to support some aspects of expertise development.
The design process was iterative, and a user-centered design approach was involved to get novices' and experts' perspectives on the visualizations directly. The final design of the visualizations was implemented on an interface. In a usability study, various expertise development indicators were measured in a clinical context.

7.2 Relations of results to the overall research questions

The first part of this research examined neurocritical care and explored the understanding of challenges perceived by an observer to the unit (i.e., the author of the dissertation) and critical care physicians (novice and expert critical care physicians). The challenges related to various aspects such as dealing with a great amount of data and correlations of clinically relevant variables that require deep understanding for appropriate treatment, organizational matters, and educational aspects linked to how expertise is developed and supported. The intensity of neurocritical care can pose a high workload, especially for novices as they have to learn and deal with multiple things simultaneously. The involvement of experts in the expertise development process is crucial as they guide the novices to learn and improve skills, acquire more knowledge and exposure to a variety of cases, and stay up to date with new research findings. However, there is also a component of learning key concepts through applications and understanding data of the patient's individual measurements. Expertise is multi-faceted and various components such as years of experience, skills (e.g., pattern recognition) etc. come into play. Experts support novices in many ways, e.g., to analyze and understand such data to develop trajectories and treatment plans. Often, novices have different approaches compared to experts; while novices follow the guidelines and protocols carefully, they tend to lack the incorporation of the trends over time and focus too much on ranges and have a numerical approach. Experts understand the individual differences among patient cases and watch out for the overall trend of the patients' stay over time. Novices receive feedback regularly and get the chance to present their thought process to their clinical team. By training such skills daily, novice physicians develop expertise over time.

Summarizing the findings of the first section with the CWA models provide the overall extent of the neurocritical care work domain regarding monitoring, treatment, and prevalent challenges in the unit from the observer's perspective. The handling of tasks represented on the DL (as part of CWA) shows differences between the novice and expert critical care physician identified by the observer, whereas more details and subjective perspective of the physicians themselves can help to add more in-depth insights and thus present a fuller picture.

Regarding the initial section, some limitations can be outlined. The amount of time spent during the observations in a single neurocritical care unit was limited; more time, more teams, different processes, and more neurocritical care units could be observed in future studies. This dissertation includes only one example of perceptions of a neurocritical care unit at a specific time. This approach has been combined with discussions and interviews of a specific clinical group: the trainee and staff critical care physicians representing the key decision-makers to mitigate bias. There are limitations linked to demographic information; participants were mainly recruited in North America and have varying specialty backgrounds that may have an impact on the type of skills they already bring to neurocritical care. For example, anesthesiology physicians focus more on equipment understanding and handling during their previous program than emergency medicine physicians who develop skills that support a fast and initial treatment. Inclusion criteria were selected rather broadly to explore a general perspective of trainees and experts in this field. For very specific research questions, more narrow or specific inclusion criteria can be further investigated in future studies. The observed ways of expertise development are also limited due to previously described time and location constraints. Thus, interviews with physicians were tackled in the next section. The models represented show impressions gained through the observations and can be further iterated and developed across more dimensions on the AH or different cases selected for the DL representation. The models thus show only one example of an AH and DL for neurocritical care.

The second part of the research involves the perspectives of novice and expert critical care physicians on the expertise development theme. Although some parts were already explored in the prior part, this section focused on how expertise is described in neurocritical care, the characteristics associated with experts, differences they have compared to novices, progress indicators, and mental model conveyance to develop expertise.

Themes identified on expert characteristics related to knowledge (e.g., literature, research), training (e.g., certifications), experience (e.g., years of practice, volume of patients), skills (e.g., accurate diagnosis, understanding nuances, pattern recognition), communication (e.g., team, patient families, conferences) performance and personal traits (endurance, heuristics, confidence). Differences between experts and novices were similarly outlined and often stated that novices are learning and acquiring the same themes.

Progress indicator themes related to the knowledge in the trained area, mental model development, increasing performance, showing autonomy, reflections, and communication. As novices' mental models differ from those of experts, themes were identified showing that knowledge and literature, teaching, training, reflection and corrections, autonomy, structuring and breaking down complexity, and tools that support the mental model development, are all considered in the process of developing expert mental models. Expertise is multi-faceted, and one expert may show different characteristics than another expert.

For this part, similar limitations exist as previously described (e.g., on demographics). Although the inclusion criteria for experts were generated before gaining more specific insights into what actually describes an expert in this field, it is noteworthy that participating experts work in academic hospitals and often demonstrated a research profile. Thus, there might also be a bias in how they perceive and describe experts in their field. Another way that can be considered in future studies is to recruit participants also from non-academic settings and compare, if there may be differences in their answers on the questions.

The time to interview the participants was limited; in some cases, the time dedicated to each question was not distributed equally. A greater variety of clinicians could be integrated as perceptions of e.g., nurses could also represent differences on expertise perception, even on how they think about physicians' expertise. Also, conducting the interviews virtually may be perceived not as natural as talking to the interviewer in person, and technical issues (e.g., WIFI connection problems) may lead to misunderstandings. Asking participants in a broader context about terms like 'expertise' or 'mental models' often nudges them to think abstractly (especially the experts) but also leads them to reflect and provide recent or common examples.

While the third part is an extension of the interviews, it refers to how expertise development can be supported through interface design in neurocritical care. To identify different parts of this question, participants were asked about the importance and how they use the bedside physiologic monitors, how the data on the interface is used to develop a trajectory for patient care, which key concepts and variables are related to ICP monitoring, and what concepts in neurocritical care are important in developing expertise but difficult to learn and apply. The results of these questions were then merged with the insights gained from the CWA models to shape design ideas for an improved interface.

Using the bedside physiologic monitor requires knowledge about its functionality. It is often used to check the immediate feedback to treatment, however, often lacks the incorporation of all relevant information for decision-making. Current interfaces only represent limited data. Physicians often have to track data from various sources and (paper) flowsheets. Its usability has also been pointed out to be limited. Trainees specifically highlighted that they deal with a high cognitive workload which is a contributing factor in going with a numerical approach.

Experts' comments revealed that data is important for trajectory: checking data to track trends over time, doing tests in real-time, including ICP waveforms and patterns, assessing the time spent in certain conditions, and considering conflicting data are some examples.
Key concepts of ICP monitoring have been outlined slightly differently among trainees and staff physicians. For trainees, key concepts related to ICP monitoring were familiarity with the equipment, understanding waveforms and trends, fully understanding the pathology and correlating variables, and the pressure reactivity concept. Staff physicians further highlighted the risks of inserting the ICP monitor, learning about expected responses to treatment and recognizing pitfalls, limitations of the equipment and evidence behind it. The staff physicians added the Monroe Kellie doctrine, the ICP waveform morphology, pressure-volume curves, autoregulation, EVD drainage alterations, and the impact of certain variables.
Many variables were listed by trainees and staff physicians that correlate with ICP; examples include vital signs (BP, HR, etCO2, RR, MAP, CPP), medication, CT and more. However, staff physicians highlighted that they have a stronger focus on waveforms and trends, and numbers may give alerts.
When the difficulty about neurocritical care concepts was discussed, it was commented that it is not only about the vast content but also the mentality, skills, data and equipment related to them. Staff physicians added that the component of individualization is another challenge. Although some neurophysiology concepts were outlined such as tackling individualizing thresholds, caring about the trajectory and learning the concept of secondary insult dose, further themes referred to a careful analysis of waveform patterns, considering ICP in its context, evaluating autoregulation, creating understanding for causes of ICP elevation and having different starting points. The Bayesian approach and consideration of risks within the trajectory were further outlined by staff physicians as important neurocritical care concepts.

CWA insights linked similar aspects to the physicians' comments; through the AH, individual aspects about the patient were outlined; the patient condition and disease type and state, the damage that was caused and relationship between multiple variables were highlighted. The measured signals also provide ideas for the design of visualizations focusing on the frequency or waveforms of signals. Some challenges depicted on the AH also correlate with the interview results; technology, data retrieval (past, current, future), and capacity (storage) are examples. The DL highlights at which states an ecological interface may support novices. For example, facilitating the set of observations, identification, interpretation, or definition of tasks may be considered beneficial.
More specific design elements could be related to insights from the observations and discussions with clinicians. Examples highlight that having more data available over time, the ability to customize views on them, and identifying thresholds for individual variables, trends and correlations with other variables can benefit the decision-maker.
The initial design ideas resulting from these inputs included a visualization capturing the ICP waveform trend, another one that shows the ICP waveform peak connections to detect abnormal ICP, a summation trend of variables and an ICP dose visualization. Discussions about the initial visualization ideas (e.g., on feasibility, usefulness etc.) led us to iterate further and focus on the secondary insult dose representation of ICP, another one focusing on the individualized thresholds, another visualization representing the trajectory and one on the summation representation of the variables ICP, CPP and MAP.
This section portrays insights into possible ways of applying some methods to develop visualizations with the intention to support novices develop expertise, however, there are also limitations associated with the process and development of visualizations.

Although one can ask physicians directly about what they perceive as challenging or how they use the data to come up with a trajectory, other methods can be considered, such as contextual inquiry. This is often done by diving into the actual context (often a site visit is required) in which participants are actively watched while they perform tasks related to the specific research question. This can be limiting as the time allocated for such a visit is constrained (as in the observations section) and may lack the breadth of the linked actions and cases observed. The interview format has provided the opportunity to reflect on general, maybe recent, challenges that come to participants' minds at that moment. The interviews were also selected over the contextual inquiry due to the constraints of the pandemic. To mitigate bias and provide both objective and subjective approaches, the CWA insights were modeled through the observations and the interviews related to a more subjective perception of novices and experts in the field.

The representation of neurocritical care concepts to experts could have further limited the variety of concepts discussed; for example, asking each expert about which concepts they deem important could have represented a better approach, however, could be more time-consuming, and certain tendencies of difficulties were already visible from previous answers they provided. Many variables related to ICP monitoring were listed, but not all were incorporated into the visualizations. There is further potential to develop more constellations and contexts of ICP in correlation with these variables. More details on the initial visualizations could be investigated; some were excluded in this project due to technical feasibility or usefulness, however, this can be analyzed further in future research. Multiple variations or iterations could be developed for each visualization to assess the effectiveness of the visualizations. This could include different styles or colours of graphics for instance. A more nuanced examination could also be done for first or second (and higher)-year trainees who may have varying levels of familiarity with the equipment.

The following part examined the perceptions of novices and experts of the visualizations proposed to be included on a bedside physiologic monitor interface for neurocritical care. The novices' perspective helped identify usability aspects (e.g., understanding and usefulness), while the experts' perception was considered to evaluate whether major neurocritical care concepts are appropriately represented on the visualizations and if they think anything about the visualizations can be improved. The overall intention in this part was to further improve the design of the visualizations and add or delete certain aspects to develop a prototype as the next stage.

Four visualizations (i.e., the ICP dose visualization, individualized thresholds visualization, trajectory visualization, and summation trend chart visualization) were thus presented to the participants. Many themes among each visualization's feedback were the same for novices and experts; for example, both participant groups often outlined comments around the visualization interaction and application. However, themes of user confusion or the impact on management did not come up in both group's feedback. Some examples as takeaways from the interviews were that visualizations should be tested against various colour deficiencies, simple and complex visualizations should be included, user customization should be available (e.g., zooming in and out on the timeline), setting up default values for ranges, enabling hover over interactions to see exact values of the variables on the visualizations, providing explanations on the visualizations in a training video, enabling turning on and off of variables, and considering what can be displayed together or solely. More features were added with some iterations, and the design was changed accordingly. Overall, main concepts were represented through the designed visualizations. Some limitations of this procedure were that each visualization was shown statically (e.g., without any interactions at that point) and without a possible representation in the vital signs interface context; however, this was described verbally, so the participants should imagine the visualization in addition to the vital signs' representation.

The data in the visualizations was hypothetically developed and did not include any signal noise or non-continuous data that might be common in a real setting. Further, the time for discussing each visualization was limited, and thus, not all visualizations could be shown to all participants, but some participants gave more in-depth feedback on the ones shown. The order of showing the visualizations may have caused potential bias in terms of complexity or relevance. The first visualization shown was the ICP dose visualization which was perceived as more complicated but also showed the greatest relevance. The colour-related comments were often linked to what the participants saw first, thus favoring similar colour-coding for the other visualizations. Although some iterations were integrated as another version and discussed with participants who conducted the study later, the effectiveness of each visualization was not tested. In future studies, all variations of each visualization could be discussed. The designs of the visualizations should be assessed with regulatory compliance and standards in this context if deployed in an interface to be used in a real setting.

The final part showed insights from the usability study which examined the extent of the ecological interface's support on expertise development when compared to a standard interface. Thus, two groups were shaped (a control group using the standard interface and an experimental group using the ecological interface), and several measures were considered part of expertise development indicators in this context (based on previous findings).

Firstly, the strategies of how an expert would act when a (new) patient is admitted or assessed were represented on an information flow map. The experts handle cases with a concurrent management approach, while novices show a sequential management pattern. The overall assessment includes initial observations and a primary survey which needs to be completed first before moving to the more in-depth inspections within the secondary survey or re-assessment phase. There are also different approaches when physicians apply their strategies within or between a patient. Between patients, physicians evaluate various factors, assess the level of acuity, and assess the triage scale. Within a patient, they assess the TIERs of therapy that depend on the severity of brain damage for instance. Limitations associated with this part are the inclusion of one expert physician who was sharing the general approach in ICUs within Canada. This could potentially be compared with strategies across different hospitals and countries. The between-and within-patient information flow maps are also linked to my observations and discussions with physicians in general. This can further be examined in future studies as strategies are dynamic and may differ depending on various factors (e.g., available resources).

Parts of these strategies are outlined during the think-aloud procedure, too.

Further, errors were counted in both groups. As hypothesized, more errors were made on the standard interface compared to the ecological interface. It was noted that many expected steps were omitted in both groups. The severity of errors was rather distributed, and it is difficult to state which group was making more severe errors. However, if only the highest severity value is taken into consideration, for scenario 1, the control group showed two errors on level 4 and only one on the same level by the experimental group. In scenario 2, the highest severity level of an error was 4 again and only came up once in both groups.

For omissions, the severity level 5 was visible among both scenarios and showed up twice in both groups. A potential limitation of this part is the review process; one expert was evaluating the insights of participants and could be extended to an international expert committee, evaluating the same statements. The harm scale was widened to have sub-categories (e.g., 2-3 instead of 2 and 3 only) which also provides room for comparison of multiple expert evaluations. This is recommended for future studies.

Both groups' reasoning and related SA (level 1-3) was assessed.
Level 1: As expected, participants in the experimental group shared a more focused first glance on the ICP uptrend and response to therapy indications. In scenario 2, they also commented on risks they perceived through the representations. It remains questionable why the experimental group has not highlighted or verbalized the waveform compliance issues at this point. Potentially, the new visualizations may have taken their attention to focus on trends and ranges of ICP, rather than outlining the real-time waveform pattern, too. In future studies, the familiarity with the EID can be trained longer, so that participants get accustomed to the vocalizations and can focus on their routine procedure including the identification of the waveform patterns. Different scenarios can be integrated to let participants highlight their insights about the waveforms.
The control group had a different approach on going through all vital signs and shared immediate considerations for medical therapy at this very early stage. However, many participants in this group noticed the waveform pattern abnormality sooner.
Level 2: When initial concerns were shared, it was noted that the experimental group related these risks more to the response to therapy and correlation of variables and patterns they recognize. The control group jumped directly to diagnoses.
Level 3 (risks): The wide variety of risks was noticed among the control group. It also shows though, that they might just state any possible risk related to such cases, while the experimental group was more focused on procedures, control of ICP, autoregulation and individual characteristics of patients. An additional tool considered to get more insights on the patient state (i.e., transcranial doppler) was pointed out by one participant in the control group, and the relationship of variables was deemed important. The experimental group highlighted the need for providing more therapeutical interventions and contacting the patient's family.
Level 3 (priorities and treatment plan): Although many treatment plans and priorities looked similar overall, the control group stated a great variety of medical therapy options, and targets of physiologic variables. The experimental group shared a smaller variety of such options and proposed TIER 1 and 2 therapeutical plans. The experimental group also looked more to find the reasons for the ICP elevations and considered more aggressive treatment especially in scenario 2.
Level 3 (communication): Although clinicians from other units or departments were considered as communication points, no major differences among both groups were noticed overall.
Level 3 (handover summary): The experimental group pointed towards the correlation of ICP with decreasing CPP and increasing HR and observations on poor waveform compliance. The control group noticed the waveform non-compliance at early stages but did not point it out again during the handover summary part. Some relevant information was omitted by both groups. While the control group outlined more details of risks, the experimental group talked more about previous procedures in their summary. Both groups included comments on the response to therapy (e.g., sedation). Therapies across both groups were similar, however coiling was a consideration for the control group already in scenario 1 while the experimental group suggested maximizing therapy first. Medication suggestions among both groups varied. Another difference noted during scenario 2 summaries was the overall assessment on the case provided by the control group while the experimental group shared, they would consider all possible strategies and treat more aggressively overall.
All in all, the experimental group showed tendencies of a more focused assessment than the control group in which they outlined their reasoning related to the annotations and context of procedures and analysis of the visualizations' meanings (e.g., relationship of variables, response to therapy). Although more aspects

were shared on SA level 3 than level 2 or 1, it may be difficult to claim that the experimental group was stronger on all level 3 categories (e.g., communication was similar across both groups) than the control group. The probing questions may have also created bias or impacted the flow of their assessments. The EID included various visualizations with a focus on ICP which may have influenced the participants of the experimental group to comment more on the ICP variable. Both groups omitted relevant aspects during the handover summaries, even if they had mentioned important aspects before in their assessments. This can be emphasized in future studies, where participants can be asked to precisely repeat the main aspects of their assessment. Some limitations of the study may have contributed to this, as participants may have forgotten to cross-check all details from the flowsheet with the visualizations and scenario descriptions. Although a training was part of the study, in future experiments, there can be a different format of providing the flowsheet (e.g., in an actual paper format) that would be a more natural setting for participants. Also, the descriptions inside the flowsheet may benefit of a better data entry or highlight of nurses' notes on such flowsheets. The single numbers denoted per hour on the flowsheet may have contributed to the perception of risk or severity among the control group. Some responses show individual differences, and it is challenging to put all in one group as all trainees may relate to varying hospital standards or may be used to follow certain procedures that are different from country to country. Reasonings may relate also to the experience with the provided two cases or different procedures prioritized in different countries.

Reflections were considered as another measure but more in a qualitative context; The reflections on participants' own strategies helped to identify and fill more gaps that happened during their think-aloud process. Reflections helped to critically think about one's own actions and consider alternative solutions on analyzing or handling situations. Thus, when participants reflected on their strategies taken, they are often stuck to their previously mentioned approach, however, added some details or aspects they forgot to mention earlier. This ranged from details around the scenario description already providing all necessary information up to maximizing medical therapy. In this part, it became more visible which themes they reflected about. There were slight differences between the reflections of the control and experimental group themes, however, many reflections were also similar and overall rich in content. While participants in the control group mentioned that they have numerical approaches and follow targets that are set by experts, the experimental group highlighted that the flowsheet should have been taken more into consideration and there was a stronger ICP focus. Both groups reflected on alternative medications, future steps, overseen aspects, lacking information and adding more measurements. The experimental group shared that they could have considered more aggressive treatment options too. This aspect was highlighted once the control group used the EID, too. The control group using the EID further outlined that identifying autoregulatory aspects are clearer displayed on the EID and they would consider new targets for different variables. These insights especially show that the control group raised awareness for more details of the data and neurocritical care concepts when looking at the additional visualizations. The reflections of the experimental group were overall richer in context and detailed out significant reflections. Overall, in future studies, one can investigate if these or different themes arise when having more scenarios. One limitation in this context is the small number of scenarios and participants may have further reflections after using the EID for a longer time and among various cases. Although some themes related to missing information, often, participants were not aware that they had the information available (e.g., in the scenario description or the flowsheet). This can also be highlighted further in future studies, so that participants are aware of the information available to them, to properly analyze the case.

Reflections on the study and use of EID showed that the EID supports confirming reasoning and conclusions, and various aspects of the visualizations are helping to understand the context (e.g., annotations, trends). It also supports participants in learning and facilitates research activities, benefitting different users (e.g., experts may find the summation graph more useful in daily practice than novices). This shows that users may quickly learn and get used to the visualizations, but some graphs may be more useful for advanced clinicians. The integration of further variables was suggested which is also a limitation of this study as we have focused on ICP visualizations rather than the incorporation of many more correlating variables such as EEG, PbtO2, or others. This can be further explored in future studies. Autoregulation was a common theme that displayed novices' uncertainties of applying this concept to the standard vital signs monitor; the EID gave clues about this concept and supported the participants to become more aware of this concept in context. As various visualizations were related to this concept, it remains unclear which visualization or which combination of visualizations or notification alerted the participant to think about autoregulation. Future studies can focus on this aspect as the concept of autoregulation is a core concept that experts try to convey to the trainees.

Reflections on the interface further helped to identify parts that could benefit from improvement and overall details they deemed important or didn't value as much (e.g., certain lines and colour differences of variables should stand out more, default settings across all three tabs would be beneficial, more clinical information could be represented on the ecological interface.) A list of elements that could be considered for the next ecological interface version is highlighted in Chapter 6. The interfaces had limitations in their functionality (e.g., some areas were non-clickable because of irrelevance for the study), however, an improved prototype could help with higher fidelity and impressions that are closer to what participants usually use in the hospital. All of these aspects should be evaluated for regulatory standard if continued in future development or studies.

Confidence ratings were provided to participants in both groups; Confidence was rated slightly higher by the control group using the standard interface. However, the confidence level of the experimental group was very close to the control group's ratings showing that it has potential for further supporting confidence levels when used for a longer time. The increased confidence ratings in scenario 2 were also noticed that could relate participants' increasing familiarity with the interface. The small number of scenarios provided in this study is a limitation; it can be further explored in future studies, whether seeing more scenarios further increases the confidence levels for handover tasks. With a bigger participant pool, considerations on e.g., gender differences can be investigated.

Participants' subjective performance was also asked to be rated; The performance ratings were similar overall, however the control group using the standard interface had a wider range for their ratings compared to the experimental group. The experimental group had slightly higher ratings and a smaller range which may imply that participants perceived their performance higher when using the ecological interface where data representation was richer and provided more depth. The performance ratings among scenario 1-2 stayed in the same range on the experimental group which means that participants found their performance ratings similar across varying cases. Some limitations were associated with this part; as the rating scale was the opposite to the other measured ratings, there were some ratings provided, where we thought that participants confused the scale. As this was noticed later during the analysis in some cases, it was impossible to ask the participants again if the ratings were meant to be opposite or not. However, an assumption was made with regard to their overall think-aloud and impressions of the study and ecological interface. In future studies, we recommend that a follow-up should happen immediately or within a very short timeframe to ensure participants have not forgotten their ratings (in case this is still possible). Another alternative would be to leave out such ratings. To eliminate confusion, one can also reverse the standard performance scale.

Overall, the usability aspect was still considered relevant when designing an interface that shall especially support novices in decision-making in a complex domain. Three aspects were thus considered: the usefulness of the interface, the ease of use and the representation of neurocritical care concepts through the visualizations.

Usefulness: In terms of usefulness of the interface among both groups, the ratings looked similar. However, there were some ratings showing that information may be irrelevant for their decision making on the ecological interface. In future studies, it could be further explored which elements exactly are perceived as irrelevant and thus improve the ecological interface.

Ease of use: The control group using both interfaces rated the ease of use and finding needed information lower than the ecological interface; however, the experimental group rated these higher. The organization of the information on the interface made more sense when the control group used the ecological interface compared to the standard interface. The experimental group also provided high or very high ratings; however, the distribution is laid out between ratings 4 to 7. The sequence and flow of different screens on the interface were perceived as more logical in the experimental group when the control group used it after the standard interface. The ratings by the experimental group were again distributed among ratings 4 to 7, but 5 participants (i.e., more than half of the overall participants in this group) fully agreed with this statement.

These ratings show that participants may need some time to familiarize themselves with the ecological interface to find and use information easily. The current version of the ecological interface provides a meaningful and logical information representation and flow of screens. One previously outlined limitation of the interface development was to evaluate in more depth the effectiveness of various organizations of the screens and logical order of visualizations. The training could be extended to give participants more time to find all information easily.

Neurocritical care concepts: Various neurocritical care concepts have been linked to the generation of visualizations which shows that the ICP burden concept, the relationship between ICP and other variables, trajectory, and individualizing the approach overall are incorporated. The ratings among the control group using first the standard, then the ecological interface further highlighted that the ecological interface was perceived as stronger on the statements. However, the individualized treatment concept requires more evaluations and changes in scenarios where the individual aspects are more determinant. All participants stated that they think the current standard interface requires improvement and some rated high agreement for the ecological interface being an improvement to the current interfaces.

Overall, the control group got the chance to see the ecological interface for a short time at the end of the study and participants often realized at that point, how they would have provided different ratings, especially on the usability questionnaire. For the experimental group, it was assumed that their comparison would reflect to the standard interface they use daily in the ICU. Future studies should consider more scenarios and longer data sets that can be deployed on the interfaces. The integration of a five-minute loop on the vital signs screen was a temporary solution, however, for long studies, a longer loop can be incorporated as there are participants with varying think-aloud durations. The interface interactions could benefit from a saving function that would save the settings on each tab and doesn't fall back to the default setting. The design of the variables (e.g., the thickness of lines) and having access to more data (e.g., blood work) can be incorporated and improved for the next version of the interface.

The colour deficiency aspect was not tested for all particular types of deficiencies. However, the researcher asked to share if participants had a colour deficiency to be able to help during the study conduction if this was the case.

A bigger group of participants with various backgrounds (trainees in different levels) and different levels of familiarity with ICP can be considered in the future. Also, evaluating the interfaces with experts in the field can be an addition and imply how the ecological interface may impact teaching. This study was conducted with individuals; however, the use of the ecological interface could further be examined in a focus group process, e.g., during ward round discussions with the full clinician team. Other potential users of the ecological interface could be examined too, e.g., nurses spend a great amount of time looking at the data throughout the day and may contribute with other perspectives. The virtual format of the usability study could be reconsidered too; conducting this study in a simulation lab within a more realistic setting may highlight different focuses on the integrational use of the ecological interface in decision-making or handing the patient over to the next team. This context could include additional variables that are not only displayed on the vital signs bedside physiologic monitor.

A 90-minute session may be short to explore all distinct aspects of the ecological interface; a longitudinal approach can be considered in future studies and further measures that relate to expertise development can be investigated.

7.3 Contributions

This work impacts several aspects of contributions as theoretical and practical relevance in the scientific, technological, educational, policy and societal, as well as economic area.

Scientific impact:

This dissertation made a theoretical contribution to the healthcare CWA work linked to expertise development by investigating especially the WDA, ConTa and StrA level where novice and expert decision-makers were considered in critical care. A scoping review (Jiancaro & Jamieson, 2014) shows studies on healthcare CWA research which are often related to areas such as diabetes management (Thompson et al., 2003), anesthesiology (Jungk et al., 1999), emergency medicine (Wears, 2009) and telehealth (e.g., Burns et al., 2008), while the ICU was investigated through studies by Effken et al. (2001, 2006) and barely in neurology (such as by Lopez et al., 2010). The area of such combination as neurocritical care has thus been introduced in this dissertation and shows potential for further studies to be conducted in this area. Also, the EID approach to support novices in their decision-making within a complex sociotechnical area should be explored more in healthcare.

Although expertise and expertise development has been investigated in various chapters of Ericsson's et al. (2018) handbook on expertise or shown among the expertise stages by Dreyfus and Dreyfus (1980), the ecological interface design has been investigated in the air traffic control domain by exploring short-term effects of novice operators using the ecological interface (Borst et al., 2019). More such studies can take steps towards including both novices' and experts' perceptions in the healthcare context too.

The application area in this dissertation has been examined with critical care and neurocritical care professionals, i.e., trainee physicians and staff physicians as being the main decision-makers.

Other studies investigated expertise in the critical care context, for example with a different professional group as nurses (Welch and Carter, 2020) which shows parallels to our findings of experts and expertise (e.g., knowledge, experience, performance-related characteristics etc.,). Understanding the cognitive processes of different stakeholders is crucial. When designing tools or interfaces, the mental models and cognitive processes shall be considered if the expertise development component plays a role.

Technological impact:

The contribution in the area of technology in this project relates to the use of high amounts of data and its representation on bedside physiologic monitor interfaces used in daily practice. During this research project, we have investigated the use of Philips bedside physiologic monitor interface that are mostly used in the Ontario region to present it to the control group. However, this may not be the interface used as much in every hospital around the world. It offers standard functions and representation of vital signs retrieved from certain devices' measurements directly and hasn't changed its design much over the past years. Other advanced interfaces or software solutions have been developed such as the T3 ® data aggregation and visualization software by Etiometry (link: https://www.etiometry.com/t3-data-visualization/) or the clinician assistant in critical care software DOCBOX ® (link: https://docboxmed.com/).

Both software solutions show a real-time representation of clinical data as well and include analytics to support clinicians' view on patient data. While older bedside physiologic monitors are limited with their data aggregation capacity and customizability options for example, these software solutions provide an extended data access and storage capacity for longer time intervals with a greater time navigation feature. Centralizing or merging data representation of multimodal input, can facilitate users' workflow and workload which still needs to be evaluated in future research. Some investigations have undergone a heuristic and usability evaluation of the T3 web-based platform Lin et al., 2017) in the pediatric cardiac area. Several usability tasks such as e.g., orientation tracking, relationships between parameters and others were taken into consideration with a clinician group. In terms of analyses, the research group looked at scoring task completion and use errors as well as severity levels. Their overall recommendations for future data integration and visualization software relate to the reduction of "[…] redundant data streams. Provide user awareness. Reduce clinician cognitive demand in interacting with the visual displays. Mandate integration of data integration and visualization software with existing medical record systems. […] Provide easy time navigation. Ensure interface is flexible to different types of users and levels of expertise. Ensure software responsiveness" (Lin et al, 2017).

It is thus noteworthy that design of such data representation interfaces plays a key role in making the interaction with the system convenient and meaningful for various system users. The graphics implemented on the interface we created as the ecological interface, are examples of supporting expertise development for novices in a distinct environment. Thus, the technological representation has different conditions to match and support stakeholders on different expertise levels and professional background. MedTech companies especially can benefit of these findings and provide different ways of representing data in software solutions. Also, other industries dealing with great amounts of data can benefit from the methodological approach, core concepts and visualizations.

Educational impact:

The implemented visualizations on the interface are based on core concepts that experts especially deem relevant when they foster trainees' expertise development. The visualizations thus portray a learning component to a certain extent and shall support novices to understand complex relationships by themselves.

"The ecological interface makes it fool-proof. For someone not working regularly in neurocritical care this is extremely helpful: the trend and annotations. The ecological interface summarizes very well the flowsheet information in a more digestible way for the user and why this happened" (P17).

This statement shows that the ecological critical care trainees in general can benefit from learning with the EID. It also provides a basis for ward round discussions or during training, as well as observations and simulation examples when clinicians gather and analyze the details together or individually. Although the usability study was conducted with individual participants, further research can be conducted in focus groups or teamwork activities, and from a teaching perspective, too. Overall, the ecological interface can support the learning curve of trainees to grow more and can facilitate the communication points between trainer and trainee when creating more in-depth and objective discussions about patient cases.

Policy and societal impact:

The drafted visualizations are early-stage designs and have the intention of impacting the way clinicians think about a percent-dose, threshold or individual approach in context. When coming up with design ideas, only general aspects of the interface were checked with medical device standards. The interface thus has not been taken under strict consideration for guidelines or ISO standards as it would be needed in case of real use in the hospital. As an example, the ISO13485 Medical devices certification, ISO 62366 for Usability Requirements for Medical Devices, ISO 14971 Risk Management in Medical Devices, and the

IEC60601 for Medical Electric Equipment relating to safety and essential performance could be relevant regulations to check and comply with.
It is not only the healthcare, engineering, and systems design community benefiting from advances in design and technology but also the society in general that might be in need of better healthcare support. This work portrays new considerations for improving professional work to become more efficient and reach a higher quality of care.

Economic impact:
The Canadian Institute for Health Information provides insights into Canadian ICUs (August 2016); for example, the average ICU length of stay is around 3 days being similar to European countries. Although numbers are not represented for the year of 2023 in this report, it was noticed that the general ICU usage in Canada is rising. Teaching hospital ICUs show a higher average occupancy rate than large facilities in urban areas. In 2013 to 2014, ICU admissions are stated around 230,800 adults (excluding Quebec). The report also considers how patients are admitted: either due to urgent or emergent, or planned or elective situations. However, 8 out of 10 ICU patients are admitted because of urgent or emergent situations.
Further, ICU stays are associated with large costs for the hospitals ("3 times the average cost of a day's stay on a general ward, as ICU stays are more resource-intensive" due to the Canadian Institute for Health Information report on Canadian ICUs (2016).
Regarding the high number of admissions in the teaching ICUs, it is playing a key role in how expertise is growing so that clinicians provide the best care for their patients.
In the long run, it could be expected that clinicians (especially novices) have better understanding of the patient's condition which could guide them to achieve higher outcomes. If clinicians have more time to dedicate to the patient and don't miss out on relevant information, it will also impact the economic structure of the hospital in the long run. More patients could then be cared for or the revisit rate of previously patients cared for in the unit may be reduced. Thus, the resources will be allocated differently and provide more room for individual care.
In addition, the expected expenses for developing an ecological interface are assumed to not turn out high in its cost since the used graphs are kept basic and easy to integrate on current bedside monitoring systems. In return, with a higher success rate in the ICU, there could be more opportunities for MedTech companies to explore and integrate new ideas and advanced designs.

7.4 Impact of Covid-19

This work was conducted during the Covid-19 pandemic in which we faced several challenges in the conduction of the studies and collaboration with ICU professionals. The specialty of neurocritical care is limited internationally and together with the pandemic, we had less resources available to conduct this research. Major challenges related to the participant recruitment in the studies, as well as the time delays due to the recruitment challenges and collaboration with neurocritical care experts who were supporting us although they were working in the ICUs more than usual due to the resource limitations in hospitals. The initial intention of this work was to be designed for a simulation study with an in-person approach to create a more realistic environment, however, due to regulations making in-person study conduction challenging and could increase risks of exposure to Covid-19, we offered a virtual study format and accommodated around the participants' availability by offering study participation hours from 6:00 am in the morning until midnight (18 hour time slot) during the week as well as weekend.

Bibliography

Alkhachroum, A., Kromm, J. & De Georgia, M.A. (2022). Big data and predictive analytics in neurocritical care. Curr Neurol Neurosci Rep 22, 19–32. https://doi.org/10.1007/s11910-022-01167-w

Andrade, E., Quinlan, L., Harte, R., Byrne, D., Fallon, E., Kelly, M., ... & ÓLaighin, G. (2020). Novel interface designs for patient monitoring applications in critical care medicine: human factors review. JMIR human factors, 7(3), e15052.

Andrews, R. W., Lilly, J. M., Srivastava, D., & Feigh, K. M. (2023). The role of shared mental models in human-AI teams: a theoretical review. Theoretical Issues in Ergonomics Science, 24(2), 129-175.

Ashoori, M., Burns, C., Momtahan, K., & d'Entremont, B. (2011, September). Control task analysis in action: Collaboration in the operating room. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 55, No. 1, pp. 272-276). Sage CA: Los Angeles, CA: SAGE Publications.

Austin, E., Blakely, B., Salmon, P., Braithwaite, J., & Clay-Williams, R. (2022). Identifying constraints on everyday clinical practice: applying work domain analysis to emergency department care. Human Factors, 64(1), 74-98.

Bakken, L. L., Sheridan, J., & Carnes, M. (2003). Gender differences among physician–scientists in self-assessed abilities to perform clinical research. Academic Medicine, 78(12), 1281-1286.

Bhardwaj, A., Mirski, M. A., & Ulatowski, J. A. (2004). Handbook of neurocritical care. Springer Science & Business Media.

Borst, C., Visser, R. M., Van Paassen, M. M., & Mulder, M. (2019). Exploring short-term training effects of ecological interfaces: A case study in air traffic control. IEEE Transactions on Human-Machine Systems, 49(6), 623-632.

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), APA handbook of research methods in psychology, Vol.
2. Research designs: Quantitative, qualitative, neuropsychological, and biological (pp. 57–71).
3. American Psychological Association. https://doi.org/10.1037/13620-004

Brems, C., Johnson, M. E., Warner, T. D., & Roberts, L. W. (2006). Barriers to healthcare as reported by rural and urban interprofessional providers. Journal of interprofessional care, 20(2), 105-118.

Burns, C.M., & Hajdukiewicz, J. (2004). Ecological Interface Design (1st ed.). CRC Press

Burns, C., Jamieson, G., Skraaning, G., Lau, N., & Kwok, J. (2007, October). Supporting situation awareness through ecological interface design. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 51, No. 4, pp. 205-209). Sage CA: Los Angeles, CA: SAGE Publications.

Burns, C. M., Enomoto, Y., & Momtahan, K. L. (2008). A cognitive work analysis of cardiac care nurses performing teletriage. In A. Bisantz & C. Burns (Eds.), Applications of cognitive work analysis (pp. 149–174). Boca Raton, FL: CRC Press.

Burns, C.M. (2019). Cognitive Work Analysis: Models of Expertise, in (eds. Waard, Schraagen, Gore, Roth): The Oxford Handbook of Expertise, 451-467.

Canadian Institute for Health Information. Care in Canadian ICUs. Ottawa, ON: CIHI; 2016.

Cioffi, J. M. (2012). Loss of clinical nursing expertise: A discussion paper. International Journal of Nursing Practice, 18(5), 423-428.

Cranfield, K., & Petrucci, K. (1993, September). Interface design for clinical information systems: An ecological interface design approach. In Vienna Conference on Human Computer Interaction (pp. 391-402). Springer, Berlin, Heidelberg.

Crawford, L. M. (2019). Conceptual and theoretical frameworks in research. Research design and methods: An applied guide for the scholar-practitioner, 35-48.

Creswell, J. W. (2014): Research Design, SAGE Publications, Inc., international, 4th Edition, Los Angeles, London New Delhi, Singapur, Washington.

Christoffersen, K. (1998, October). Assessing the state of the art in integrated displays. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 42, No. 3, pp. 400-403). Sage CA: Los Angeles, CA: SAGE Publications.

Dai, H., Jia, X., Pahren, L., Lee, J., & Foreman, B. (2020). Intracranial pressure monitoring signals after traumatic brain injury: a narrative overview and conceptual data science framework. Frontiers in neurology, 11, 959.

De Georgia, M., Kaffashi, F., Jacono, F., & Loparo, K. (2015). Information technology in critical care: review of monitoring and data acquisition systems for patient care and research. The Scientific World Journal.

de-Lima-Oliveira, M., Salinet, A. S., Nogueira, R. C., de Azevedo, D. S., Paiva, W. S., Teixeira, M. J., & Bor-Seng-Shu, E. (2018). Intracranial hypertension and cerebral autoregulation: a systematic review and meta-analysis. World neurosurgery, 113, 110-124.

Donchin Y, Gopher D, Olin M, Badihi Y, Biesky M, Sprung CL, Pizov R, Cotev S. (1995 Feb). A look into the nature and causes of human errors in the intensive care unit. Crit Care Med.;23(2):294-300. doi: 10.1097/00003246-199502000-00015. PMID: 7867355.

Donchin, Y., & Seagull, F. J. (2002). The hostile environment of the intensive care unit. Current opinion in critical care, 8(4), 316-320.

Dreyfus, S. E.; Dreyfus, H. L. (1980). A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition, Operations Research Center, University of California, Berkeley.

Dziedzic PH, Suarez JI. (2022). Managing Clinical Data in Neurocritical Care. In: Healthcare Information Management Systems. Springer International Publishing; :235-245. doi:10.1007/978-3-031-07912-2_16

Effken, J. A., Kim, N. G., & Shaw, R. E. (1997). Making the constraints visible: testing the ecological approach to interface design. Ergonomics, 40(1), 1-27.

Effken, J., Loeb, R., Johnson, K., Johnson, S., & Reyna, V. (2001). Using cognitive work analysis to design clinical displays. In MEDINFO 2001 (pp. 127-131). IOS Press.

Effken, J. A. (2006). Improving clinical decision making through ecological interfaces. Ecological Psychology, 18(4), 283-318.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In Proceedings of the Human Factors Society 32nd Annual Meeting, Human Factors Society (pp. 97–101). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R. (2018). Expertise and situation awareness. The Cambridge handbook of expertise and expert performance, Cambridge University Press, 714-741.

Endsley, Mica R. (2016). "Toward a Theory of Situation Awareness in Dynamic Systems." In Human Factors. Los Angeles, CA: Sage CA. doi:10.1518/001872095779049543.

Ericsson, K. A., Prietula, M. J., Cokely, E. T. (2007). The Making of an Expert, in: Managing People. https://hbr.org/2007/07/the-making-of-an-expert (last retrieved: October 5, 2020).

Ericsson, K. A. (2011). The Surgeon's Expertise. In H. Fry & R. Kneebone (eds.), Surgical education: Theorising an emerging domain (pp. 107-121). Springer Netherlands.

Ericsson, K. A. (2015). Acquisition and maintenance of medical expertise: a perspective from the expert-performance approach with deliberate practice. Academic Medicine, 90(11), 1471-1486.

Ericsson, K. A. (2018). An Introduction to the Second Edition of The Cambridge Handbook of Expertise and Expert Performance: Its Development, Organization and Content. The Cambridge handbook of expertise and expert performance. Cambridge University Press, 3-20.

Fackler, J. C., Watts, C., Grome, A., Miller, T., Crandall, B., & Pronovost, P. (2009). Critical care physician cognitive task analysis: an exploratory study. Critical Care, 13, 1-8.

Fan, J. Y., Kirkness, C., Vicini, P., Burr, R., & Mitchell, P. (2008). Intracranial pressure waveform morphology and intracranial adaptive capacity. American Journal of critical care, 17(6), 545-554.

Feng, M., Zhang, Z., Zhang, F., Ge, Y., Loy, L. Y., Vellaisamy, K., ... & Guan, C. (2011, August). iSyNCC: An intelligent system for patient monitoring & clinical decision support in neuro-critical-care. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 6426-6429). IEEE.

Fidel, R., & Pejtersen, A. M. (2004). From information behaviour research to the design of information systems: The Cognitive Work Analysis framework. Information Research: an international electronic journal, 10(1), n1.

Flick, Uwe. (2011). Introducing Research Methodology: A Beginner's Guide to Doing a Research Project. Los Angeles: Sage.

Gibson, J. J. (1979/1986). The Ecological Approach to Visual Perception, Hillsdale, NJ: Lawrence Erlbaum Associates.

Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.

Görges, M., Westenskow, D. R., & Markewitz, B. A. (2012). Evaluation of an integrated intensive care unit monitoring display by critical care fellow physicians. Journal of clinical monitoring and computing, 26, 429-436.

Güiza, F., Depreitere, B., Piper, I., Citerio, G., Chambers, I., Jones, P. A., ... & Meyfroidt, G. (2015). Visualizing the pressure and time burden of intracranial hypertension in adult and paediatric traumatic brain injury. Intensive care medicine, 41, 1067-1076.

Hadi, M.A., José Closs, S. (2016). Ensuring rigour and trustworthiness of qualitative research in clinical pharmacy. Int J Clin Pharm 38, 641–646. https://doi.org/10.1007/s11096-015-0237-6

Hajdukiewicz, J. R., & Vicente, K. J. (2004). A theoretical note on the relationship between work domain analysis and task analysis. Theoretical Issues in Ergonomics Science, 5(6), 527-538.

Hemphill III, J. C. (2022). Pro: Neurocritical Care Big Data and AI: It's About Expertise. Neurocritical Care, 37(Suppl 2), 160-162.

Howie, D. E. (1996). Shaping expertise through ecological interface design: strategies, metacognition, and individual differences. University of Toronto.

Jamieson, G. A., & Vicente, K. J. (2001). Ecological interface design for petrochemical applications: supporting operator adaptation, continuous learning, and distributed, collaborative work. Computers & Chemical Engineering, 25(7-8), 1055-1074.

Jiancaro, T., Jamieson, G. A., & Mihailidis, A. (2014). Twenty years of cognitive work analysis in health care: a scoping review. Journal of cognitive engineering and decision making, 8(1), 3-22.

Johnson, N. J., Maher, P. J., Badulak, J., & Luks, A. M. (2017). The transition from emergency medicine resident to critical care fellow: a road map. AEM Education and Training, 1(2), 116-123.

Jungk, A., Thull, B., Hoeft, A., & Rau, G. (1999). Ergonomic evaluation of an ecological interface and a profilogram display for hemodynamic monitoring. Journal of clinical monitoring and computing, 15, 469-479.

Jungk, A., Thull, B., Hoeft, A., & Rau, G. (2000). Evaluation of Two New Ecological Interface Approaches for the Anesthesia Workplace. Journal of Clinical Monitoring and Computing, 243-258.

Kamel, H., & Hemphill III, J. C. (2016). Multimodal brain monitoring and neuroinformatics. Oxford Textbook of Neurocritical Care, 152.

Kawoos, U., McCarron, R. M., Auker, C. R., & Chavko, M. (2015). Advances in Intracranial Pressure Monitoring and Its Significance in Managing Traumatic Brain Injury. International journal of molecular sciences, 16(12), 28979–28997. https://doi.org/10.3390/ijms161226146

Kim N., Krasner A., Kosinski C., Winninger M., Qadri M., Kappus Z., Danish S., Craelius W., (2016). Trending autoregulatory indices during treatment for tramatic brain injury. J Clin Monit Comput, 30:821. Doi: 10.1007/s10877-015-9779-3. (p.242)

Koch, S. H., Staggers, N., Weir, C., Agutter, J., Liu, D., & Westenskow, D. R. (2010, September). Integrated information displays for ICU nurses: field observations, display design, and display evaluation. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 54, No. 12, pp. 932-936). Sage CA: Los Angeles, CA: SAGE Publications.

Kochanek, P.M., Clark, R.S., Ruppel, R.A., Adelson, P.D., Bell, M.J., Whalen, M.J., Robertson, C.L., Satchell, M.A., Seidberg, N.A., Marion, D.W., Jenkins, L.W., (2000 Jul). Biochemical, cellular, and molecular mechanisms in the evolution of secondary damage after severe traumatic brain injury in infants and children: lessons learned from the bedside. Pediatr. Crit. Care Med. 1 (1), 4e19.

Kushniruk, A., Patel, V., Cimino, J. J., & Barrows, R. A. (1996). Cognitive evaluation of the user interface and vocabulary of an outpatient information system. In Proceedings of the AMIA Annual Fall Symposium (p. 22). American Medical Informatics Association.

Krupp, A., Steege, L., Lee, J., Lopez, K. D., & King, B. (2022). Supporting decision-making about patient mobility in the intensive care unit nurse work environment: Work domain analysis. JMIR nursing, 5(1), e41051.

Lazaridis, C., Rusin, C. G., & Robertson, C. S. (2019). Secondary brain injury: predicting and preventing insults. Neuropharmacology, 145, 145-152.

Lee, M., Steyvers, M., DeYoung, M., & Miller, B. (2011). A model-based approach to measuring expertise in ranking tasks. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 33, No. 33).

Li, Y., Burns, C. M., & Kulić, D. (2014, June). Ecological interface design for knee and hip automatic physiotherapy assistant and rehabilitation system. In Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (Vol. 3, No. 1, pp. 1-7). Sage CA: Los Angeles, CA: SAGE Publications.

Lim, R. H., Anderson, J. E., & Buckle, P. W. (2016). Work domain analysis for understanding medication safety in care homes in England: an exploratory study. Ergonomics, 59(1), 15-26.

Lin, Y. L., Guerguerian, A. M., Tomasi, J., Laussen, P., & Trbovich, P. (2017). Usability of data integration and visualization software for multidisciplinary pediatric intensive care: a human factors approach to assessing technology. BMC medical informatics and decision making, 17(1), 1-19.

Lopez, K. D., Gerling, G. J., Cary, M. P., & Kanak, M. F. (2010). Cognitive work analysis to evaluate the problem of patient falls in an inpatient setting. Journal of the American Medical Informatics Association, 17(3), 313-321.

Mayer, S. A., Coplin, W. M., Chang, C., Suarez, J., Gress, D., Diringer, M. N., ... & Mellick, M. E. (2006). Program requirements for fellowship training in neurological intensive care: United Council for Neurologic Subspecialties guidelines. Neurocritical care, 5(2), 166-171.

McCredie, V. (2019). Minimizing Secondary Injury [Presenation]. International Symposium on Intensive Care and Emergency Medicine.

McNamara, L. A., Cole, K., Haass, M. J., Matzen, L. E., Daniel Morrow, J., Stevens-Adams, S. M., & McMichael, S. (2015). Ethnographic methods for experimental design: case studies in visual search. In Foundations of Augmented Cognition: 9th International Conference, AC 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2–7, 2015, Proceedings 9 (pp. 492-503). Springer International Publishing.

Menchetti, I., Eagles, D., Ghanem, D., Leppard, J., Fournier, K., & Cheung, W. J. (2022). Gender differences in emergency medicine resident assessment: a scoping review. AEM Education and Training, 6(5), e10808.

Miller, A., & Sanderson, P. (2000, July). Modeling “deranged” physiological systems for ICU information system design. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 44, No. 26, pp. 245-248). Sage CA: Los Angeles, CA: SAGE Publications.

Morineau, T., Beuzet, E., Rachinel, A., & Tobin, L. (2009). Experimental evaluation of a tide prediction display based on the ecological interface design framework. Cognition, Technology & Work, 11(2), 119-127.

Morse, J. M. (2000). Determining sample size. Qualitative health research, 10(1), 3-5.

National Academies of Sciences Engineering and Medicine (2022). Human-AI Teaming: State-of-the- Art and Research Needs. Washington, DC: The National Academies Press. doi:10.17226/26355.

Norman, D.A. (1983). Some observations on mental models. In D. Gentner and A.L. Stevens (eds) Mental Models (Erlbaum, Hillsdale, NJ), 7-14.

Norman, G. R., Grierson, L. E. M., Sherbino, J., Hamstra, S. J., Schmidt, H. D., Mamede, S. (2018). Expertise in Medicine and Surgery. In (eds. Ericsson, K. A., Hoffmann, R. R., Kozbelt, A., Williams, A. M.): The Cambridge Handbook of Expertise and Expert Performance, second edition, Cambridge University Press, USA, 331-355.

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: a synthesis of recommendations. Academic medicine, 89(9), 1245-1251.

Patel, V. L., Glaser, R., & Arocha, J. F. (2000). Cognition and expertise: acquisition of medical competence. Clinical and Investigative Medicine, 23(4), 256-260.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. IEEE Transactions on Systems, Man and Cybernetics, 15:234-243.

Rasmussen, J., Pejtersen, A. M., & Schmidt, K. (1990). Taxonomy for cognitive work analysis. Roskilde: Risø National Laboratory.

Ravitch, S. M., & Riggan, M. (2017). Reason and rigour: how conceptual framework guide research (2nd ed.). Thousand Oaks, CA: SAGE.

Rezai, L. S., & Burns, C. M. (2014, June). Using cognitive work analysis and a persuasive design approach to create effective blood pressure management systems. In Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (Vol. 3, No. 1, pp. 36-43). Sage CA: Los Angeles, CA: SAGE Publications.

Rincon, F., & Mayer, S. A. (2007). Neurocritical care: a distinct discipline? Current opinion in critical care, 13(2), 115-121.

Robson, C. (2011). Real World Research. A Resource for Users of Social Research Methods in Applied Settings". Third Edition. Wiley, United Kingdom.

Rodríguez-Boto, G., Rivero-Garvía, M., Gutiérrez-González, R., & Márquez-Rivas, J. (2015). Basic concepts about brain pathophysiology and intracranial pressure monitoring. Neurología (English Edition), 30(1), 16-22.

Rojas, S. S. O., Ordinola, A. A. M., Veiga, V. C., & Souza, J. M. D. (2021). The use of a noninvasive intracranial pressure monitoring method in the intensive care unit to improve neuroprotection in postoperative cardiac surgery patients after extracorporeal circulation. Revista Brasileira de Terapia Intensiva, 33, 469-476.

Rothschild, J. M., Landrigan, C. P., Cronin, J. W., Kaushal, R., Lockley, S. W., Burdick, E., ... & Bates, D. W. (2005). The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. Critical care medicine, 33(8), 1694-1700.

Rubiano, A. M., Figaji, A., & Hawryluk, G. W. (2022). Intracranial pressure management: moving beyond guidelines. Current Opinion in Critical Care, 28(2), 101-110.

St-Maurice, J., & Burns, C. M. (2014, June). User perception of data and medical record personalities. In Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (Vol. 3, No. 1, pp. 15-22). Sage CA: Los Angeles, CA: SAGE Publications.

Sarcevic, A., Lesk, M. E., Marsic, I., & Burd, R. S. (2010). Towards an efficient method for studying collaborative practices of emergency care teams.

SHANTEAU, J., & SENGSTACKE, L. D. N. (2009). Description of inpatient medication management using cognitive work analysis. CIN: Computers, Informatics, Nursing, 27(6), 379-392.

Sharp, T. D., & Helmicki, A. J. (1998, October). The application of the ecological interface design approach to neonatal intensive care medicine. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 42, No. 3, pp. 350-354). Sage CA: Los Angeles, CA: SAGE Publications.

Sheth, K. N., Stein, D. M., Aarabi, B., Hu, P., Kufera, J. A., Scalea, T. M., & Hanley, D. F. (2013). Intracranial pressure dose and outcome in traumatic brain injury. Neurocritical care, 18, 26-32.

Shinde, S. N., Kulkarni, J. V., Mohite, V. M., & Kshirsagar, A. G. (2017). An INTELLIGENT SYSTEM for PATIENT MONITORING & CLINICAL DECISION SUPPORT IN NEURO-CRITICAL- CARE. International Research Journal of Engineering and Technology, 4(2).

Shutter, L., & Molyneaux, B. J. (Eds.). (2018). Neurocritical care. Pittsburgh Critical Care Medic.

Simpkins, A. N., Busl, K. M., Amorim, E., Barnett-Tapia, C., Cervenka, M. C., Dhakar, M. B., ... & Maciel, C. B. (2020, December). Proceedings from the neurotherapeutics symposium on neurological emergencies: shaping the future of neurocritical care. In Neurocritical Care (Vol. 33, pp. 636-645). Springer US.

Somand, D., & Zink, B. (2005). The influence of critical care medicine on the development of the specialty of emergency medicine: a historical perspective. Academic emergency medicine, 12(9), 879-883.

Sorrentino E, Diedler J, Kasprowicz M, Budohoski KP, Haubrich C, Smielewski P, et al. (2012). Critical thresholds for cerebrovascular reactivity after traumatic brain injury. Neurocritical Care. 16:258–66. doi: 10.1007/s12028-011-9630-8 .

Stanton, N. A., Salmon, P. M., Walker, G. H., & Jenkins, D. P. (Eds.). (2017). Cognitive work analysis: applications, extensions and future directions. CRC Press.

Suarez, J. I., Martin, R. H., Bauza, C., Georgiadis, A., Venkatasubba Rao, C. P., Calvillo, E., ... & LeRoux, P. D. (2020). Worldwide organization of neurocritical care: results from the PRINCE study part 1. Neurocritical care, 32, 172-179.

Thompson, L. K., Hickson, J. C., & Burns, C. M. (2003, October). A work domain analysis for diabetes management. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 47, No. 12, pp. 1516-1520). Sage CA: Los Angeles, CA: SAGE Publications.

Tuckett, A. G. (2005). Part II. Rigour in qualitative research: complexities and solutions. Nurse researcher, 13(1).

University of Toronto Adult CCM Academic Advisor CBME Manual (July 2019), V1.0, Toronto.

Üreten, E., McCredie, V., & Burns, C. (2020, September). Cognitive Work Analysis Models of Neuro-Critical Care. In Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (Vol. 9, No. 1, pp. 191-193). Sage CA: Los Angeles, CA: SAGE Publications.

Üreten, E., Schaef, K., McCredie, V., & Burns, C. (2022, September). Assessing Intracranial Pressure Visualizations Displayed on ICU Bedside Physiologic Monitors. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 66, No. 1, pp. 2143-2147). Sage CA: Los Angeles, CA: SAGE Publications.

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC medical research methodology, 18(1), 1-18.

Vicente, K. J., & Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating 'direct perception in complex work domains. Ecological psychology, 2(3), 207-249.

Vicente, K. (1999). Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work (1st ed.). CRC Press. https://doi.org/10.1201/b12457.

Vicente, K. J. (2003). Cognitive Work Analysis Toward Safe, Productive, and Healthy Computer-Based Work [Book Review]. IEEE Transactions on Professional Communication, 46(1), 63-65.).

Walsh, K. E., Harik, P., Mazor, K. M., Perfetto, D., Anatchkova, M., Biggins, C., ... & Tjia, J. (2017). Measuring harm in healthcare: optimizing adverse event review. Medical care, 55(4), 436.).

Wears, R. L. (2009). What makes diagnosis hard? Advances in Health Sciences Education, 14(Suppl 1), 19-25.

Webster's New World Dictionary (1968), Cleveland, OH: World Publishing Company.

Welch, T. D., & Carter, M. (2020). Expertise among critical care nurses: A grounded theory study. Intensive and Critical Care Nursing, 57, 102796.

Wickens, C. D., Dehais, F. (2019). Expertise in Aviation. In: The Oxford Handbook of Expertise. Oxford

Wilson, J. A., Shutter, L. A., & Hartings, J. A. (2013). COSBID-M3: a platform for multimodal monitoring, data collection, and research in neurocritical care. Cerebral Vasospasm: Neurovascular Events After Subarachnoid Hemorrhage, 67-74.

Wu, C., Jeon, J., Cafazzo, J. A., & Burns, C. M. (2012). Work domain analysis for designing a radiotherapy system control interface. In Proceedings of the 2012 Symposium on Human Factors and Ergonomics in Health Care (Vol. 1, No. 1, pp. 224-228). Santa Monica, CA: Human Factors and Ergonomics Society.

Yang, M. T. (2020). Multimodal neurocritical monitoring. Biomedical journal, 43(3), 226-230.

Website links:

Coblinder (2000). Coblis-Colour Blindness Simulator (link: www.colourblindness.com/coblis-colour-blindness-simulator/; retrieved June. 2023)

DOCBOX ® (link: https://docboxmed.com/; retrieved September 4, 2023).

Etiometry (link: https://www.etiometry.com/t3-data-visualization/; retrieved September 4, 2023).

National Eye Institute (2023, August 7). Types of Color Vision Deficiency; https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness/types-color-vision-deficiency; retrieved August 9, 2023).

Expert. (pre 2018). In Wikipedia. http://en.wikipedia.org/wiki/Expert. (as referenced in Ericsson, 2018).

Appendix

Appendix A - Recruitment Letter - Interview Study

Email for program directors.

Subject: Invitation to participate in an interview of exploring expertise development and interface design in critical care

Hello Dr. ______,

Dr. Victoria McCredie, Dr. Catherine Burns and I are working on an expertise development and interface design research project. I am a Systems Design Engineering PhD candidate at the University of Waterloo and am investigating critical care challenges that relate to the use of bedside physiologic monitoring in neurocritical care and how critical care novices develop expertise in the ability to detect, diagnose, and treat neurophysiologic deteriorations. We would like to conduct interviews with nurses, trainees and staff physicians to better understand how expertise development can be supported via interface design.

Interview Purpose
Our goal is to explore the challenges nurses, trainees and staff physicians may be facing in developing expertise in neurocritical care, and possible ways of how this can be improved through interface design.

We understand the timing for this interview is not ideal given the ongoing COVID-19 pandemic. However, the aim of our research is to develop advanced designs for interfaces that support decision-making, facilitate learning complex relationships, and ultimately reduce the strain on our healthcare providers in the intensive care unit.

It would be greatly appreciated if you could forward the attached email invitation to your critical care nurses, trainees and staff physicians. The de-identified data collected via this interview will be held securely and will be accessible only to the designated study investigators.

This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB 42892).

Please contact me at euereten@uwaterloo.ca if you have any questions or concerns about this interview process.

Thank you in advance for your time, participation and support.

Email to physicians

Subject: Invitation to participate in an interview of exploring expertise development and interface design in critical care

Hello ______,

Dr. Victoria McCredie, Dr. Catherine Burns, Kathleen Schaef, and I are working on an expertise development and interface design research project. I am a Systems Design Engineering PhD candidate at the University of Waterloo and am investigating critical care challenges that relate to the use of bedside physiologic monitoring in neuro-critical care and how critical care nurses, trainees and staff physicians develop expertise in the ability to detect, diagnose, and treat neurophysiologic deteriorations. We would like to conduct interviews with nurses, trainees and staff physicians to better understand how expertise development can be supported via interface design.

Interview Purpose

We are conducting interviews as part of a research study to explore the perceived challenges of trainees and staff physicians face in developing expertise development in neuro-critical care, and possible ways of how this can be improved through interface design.

The interview will take around 45-60 minutes via an online meeting. Please suggest a day and time over the next 2 months that's convenient for you. You can access this link [INSERT LINK] to schedule a suitable time.

We understand the timing for this interview is not ideal given the ongoing COVID-19 pandemic. However, the aim of our research is to develop advanced designs for interfaces that support decision-making, facilitate learning complex physiological relationships, and ultimately reduce the strain on our healthcare providers in the intensive care unit.

Remuneration

All trainees and staff physicians who participate in the interview will be given **150 CAD** in remuneration for time taken to participate.

This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB 42892). If you have questions for the Board contact the Office of Research Ethics, at 1-519-888-4567 ext. 36005 or reb@uwaterloo.ca.

Appendix B - Recruitment Letter - Usability Study

Email for program directors.

Subject: Invitation to participate in a usability study of exploring expertise development, interface design and evaluation in neurocritical care

Hello Dr. ______,

Dr. Victoria McCredie, Dr. Catherine Burns, Kathleen Schaef and I are working on an expertise development and interface design research project. I am a Systems Design Engineering PhD candidate at the University of Waterloo and am investigating critical care challenges that relate to the use of bedside physiologic monitoring in neurocritical care and how critical care trainees develop expertise in the ability to detect, diagnose, and treat neurophysiologic deteriorations.

Study Purpose

The purpose of this study is to compare two bedside physiologic monitor interfaces and assess its effects on perceived usability and support in developing expertise.

Conducting this study will reveal the user experience of trainee physicians in critical care. By understanding the mental models, situation awareness and challenges faced when using a physiologic bedside monitor interface within critical care, we want to contribute to a better design of physiologic bedside monitor interfaces, support physicians' decision-making process and ultimately improve the quality of care and effectiveness of critical care.

Study tasks and duration

In this study, we will provide neurocritical care patient scenarios for which we want participants to think aloud and verbalize their actions taken on the provided interface. During this process, we would like to audio and screen record your shared interface. The study duration is 90 minutes.

To be able to fully see the interface we will share on the day of the study, participants should have access to a computer or screen with a minimum width of 1024 px (this link shows the resolution of your used screen: https://screenresolutiontest.com/), and be in a quiet room. Phones and iPads are usually not in the required screen size.

The aim of our research is to develop advanced designs for interfaces that support decision-making, facilitate learning complex relationships, and ultimately reduce the strain on our healthcare providers in the intensive care unit.

It would be greatly appreciated if you could forward the attached email invitation to your critical care trainees. The de-identified data collected via this study will be held securely.

This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB 43798).

Please contact me–at euereten@uwaterloo.ca if you have any questions or concerns about this study process.

Thank you in advance for your time, participation and support.

Email for trainee physicians.

Subject: Invitation to participate in a usability study of exploring expertise development, interface design and evaluation in neurocritical care

Hello,

Dr. Victoria McCredie, Dr. Catherine Burns, Kathleen Schaef and I are working on an expertise development and interface design research project. I am a Systems Design Engineering PhD candidate at the University of Waterloo and am investigating critical care challenges that relate to the use of bedside physiologic monitoring in neurocritical care and how critical care trainees develop expertise in the ability to detect, diagnose, and treat neurophysiologic deteriorations. We would like to conduct a usability study with trainees to better understand how expertise development can be supported via interface design.

Study Purpose

The purpose of this study is to compare two bedside physiologic monitor interfaces and assess its effects on perceived usability and support in developing expertise.

Conducting this study will reveal the user experience of trainee physicians in critical care. By understanding the mental models, situation awareness and challenges faced when using a physiologic bedside monitor interface within critical care, we want to contribute to a better design of physiologic bedside monitor interfaces, support physicians' decision-making process and ultimately improve the quality of care and effectiveness of critical care.

Study tasks and duration

In this study, we will provide neurocritical care patient scenarios for which we want participants to think aloud and verbalize their actions taken on the provided interface. During this process, we would like to audio and screen record your shared interface.

The study will take 90 minutes via an online meeting.
To be able to fully see the interface we will share on the day of the study, participants should have access to a computer or screen with a minimum width of 1024 px (this link shows the resolution of your used screen: https://screenresolutiontest.com/), and be in a quiet room. Phones and iPads are usually not in the required screen size.

During this process, we would like to audio and screen-record the participant.

Please suggest a day and time over the next month that's convenient for you.
You can access this link to schedule a suitable time and fill out demographic information for participation in the study: https://calendly.com/criticalcare-interview/90min

The aim of our research is to develop advanced designs for interfaces that support decision-making, facilitate learning complex relationships, and ultimately reduce the strain on our healthcare providers in the intensive care unit.

The de-identified data collected via this study will be held securely.

This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB 43798). If you have questions for the Board contact the Office of Research Ethics, at 1-519-888-4567 ext. 36005 or reb@uwaterloo.ca.
Please contact me–at euereten@uwaterloo.ca if you have any questions or concerns about this study process.

Thank you in advance for your time, participation and support.

Appendix C - Information Letter – Interview Study

This letter is an invitation to consider participating in a research study I am conducting as part of my PhD degree in the Department of Systems Design Engineering at the University of Waterloo under the supervision of Drs. Catherine Burns and Victoria McCredie. I would like to provide you with more information about this project and what your involvement would entail if you decide to take part.

Developing expertise is a complex process that takes years if not decades. Especially in dynamic and complex socio-technical systems, it is of utmost importance to highly perform. In critical care, expertise is required to provide the patient with correct and timely treatment. In case of lacking expertise, the effects of treatment can contain mistakes, incompleteness and even lead to life-threatening situations.

Conducting interviews will reveal in-depth insights of critical care nurses, trainees and staff physicians. By understanding the mental models, challenges faced in critical care and ways how expertise is being developed over time, we want to contribute to a better design of physiologic bedside monitor interfaces, support the decision-making process and ultimately improve the quality of care and effectiveness of critical care.

Participation in this study is voluntary and a remuneration of 150 CAD will be provided even if participants choose to withdraw from the study. The amount received is taxable. It is your responsibility to report this amount for income tax purposes.

We currently want to recruit participants (nurses, trainees and staff physicians) for an interview of 45 minutes, not more than 60 minutes in length to take place virtually.

In this interview, we want to hear about your experiences with the process of expertise development in critical care as well as your impression on some visualization concepts that we developed. These visualizations will be shown to you during the interview. Hearing your impressions will help us integrate some visualizations on a physiologic bedside monitor interface that will later be evaluated in a usability study.

You may decline to answer any of the interview questions if you wish so. You can request your data be removed from the study up until March 2022 as it is not possible to withdraw your data once papers and publications have been submitted to publishers. Further, you may decide to withdraw from this study at any time without any negative consequences by advising the researcher. With your permission, the interview will be audio and video recorded to facilitate collection of information, and later transcribed for analysis.

Shortly after the interview has been completed, I will send you a copy of the transcript to give you an opportunity to add or clarify any points that you wish. If you wish to add or change anything mentioned in the transcript, please contact me at euereten@uwaterloo.ca until March, 2022.

Your identity will be confidential. Any identifying information will be removed from transcripts and stored separately, and the video recordings will be deleted after the transcripts are finalized and stored. Your name will not appear in any thesis or report resulting from this study, however, with your permission anonymous quotations may be used. Data collected during this study will be securely stored for at least 7 years on a password-protected One Drive account and secure computers.

The interview will be conducted over an online platform, Microsoft Teams. Microsoft Teams has implemented technical, administrative, and physical safeguards to protect the information provided via the Services from loss, misuse, and unauthorized access, disclosure, alteration, or destruction. However, no Internet transmission is ever fully secure or error free.

The findings of this interview study will help us to come up with design ideas for a physiologic bedside interface used in neuro-critical care units. A second study is planned to investigate how interface design can support expertise development.
There are no known or anticipated risks to you as a participant in this study.
This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB 42892). If you have questions for the Board contact the Office of Research Ethics, at 1-519-888-4567 ext. 36005 or reb@uwaterloo.ca.

For all other questions or if you would like additional information to assist you in reaching a decision about participation, please contact me by email at euereten@uwaterloo.ca. You can also contact my supervisor, Professor Catherine Burns at 519-888-4567 ext. 33903 or email catherine.burns@uwaterloo.ca.
I hope that the results of my study will be of benefit to those organizations directly involved in the study, other voluntary recreation organizations not directly involved in the study, as well as to the broader research community.
I very much look forward to speaking with you and thank you in advance for your assistance in this project.

Appendix D - Information Letter – Usability Study

This letter is an invitation to consider participating in a research study I am conducting as part of my PhD degree in the Department of Systems Design Engineering at the University of Waterloo under the supervision of Drs. Catherine Burns and Victoria McCredie. I would like to provide you with more information about this project and what your involvement would entail if you decide to take part.

Developing expertise is a complex process that takes years if not decades. Especially in dynamic and complex socio-technical systems, it is of utmost importance to highly perform. In critical care, expertise is required to provide the patient with correct and timely treatment. In case of lacking expertise, the effects of treatment can contain mistakes, incompleteness and even lead to life-threatening situations. Two main parties are involved in expertise development: the person who develops expertise which is the trainee in this case and the staff physician who supports developing expertise.

The purpose of this study is to compare two bedside physiologic monitor interfaces and assess its effects on perceived usability and support in developing expertise. One of the interfaces will be a standard bedside physiologic monitor, the other will be including additional visualizations developed by the researchers.

Conducting this study will reveal the user experience of trainee physicians in critical care. By understanding the mental models, situation awareness and challenges faced on a physiologic bedside monitor interface within critical care, we want to contribute to a better design of physiologic bedside monitor interfaces, support physicians' decision-making process and ultimately improve the quality of care and effectiveness of critical care. We will assess participants responses and compare the performance on both interfaces. Participants will be randomly assigned to use one interface only.

Participation in this study is voluntary.

We currently want to recruit participants (trainees) for a usability study <u>of 90 minutes i</u>n length to take place <u>virtually</u>. In this study, we will ask for demographic data (e.g. age, gender, location) and provide neurocritical care patient scenarios for which we want participants to think aloud and verbalize their actions taken on the provided interface. During this process, we would like to audio-record and ask participants to share their screens with us. After the scenarios, we will ask participants to reflect on their taken steps.

You may decline to answer any of the questions if you wish so. You can request your data be removed from the study up until December 2022 as it is not possible to withdraw your data once papers and publications have been submitted to publishers. Further, you may decide to withdraw from this study at any time without any negative consequences by advising the researcher. With your permission, the study will be audio and screen-recorded to facilitate collection of information, and later transcribed for analysis.

Your identity will be confidential. Any identifying information will be removed from transcripts and stored separately. Audio-recordings will be deleted after transcription. De-identified transcripts and screen-recordings will be securely shared with an expert physician to evaluate the participants' actions while using the interface. Professor Victoria McCredie will not know who has participated as all data will be de-identified.

Your name will not appear in any thesis or report resulting from this study, however, with your permission anonymous quotations may be used. The dataset without identifiers may be shared publicly (such as in publications). Your identity will be confidential.

Data collected during this study will be securely stored for at least 7 years on a password-protected One Drive account and secure computers.

The study will be conducted over an online platform, Microsoft Teams. Microsoft Teams has implemented technical, administrative, and physical safeguards to protect the information provided via the Services from loss, misuse, and unauthorized access, disclosure, alteration, or destruction. However, no Internet transmission is ever fully secure or error free.

The findings of this study will help us to come up with recommendations for physiologic bedside interface design used in neurocritical care units.

There are no known or anticipated risks to you as a participant in this study.

This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB 43798). If you have questions for the Board contact the Office of Research Ethics, at 1-519-888-4567 ext. 36005 or reb@uwaterloo.ca.

For all other questions or if you would like additional information to assist you in reaching a decision about participation, please contact me by email at euereten@uwaterloo.ca. You can also contact my supervisor, Professor Catherine Burns at 519-888-4567 ext. 33903 or email Catherine.burns@uwaterloo.ca.

I hope that the results of my study will be of benefit to those organizations directly involved in the study, other voluntary recreation organizations not directly involved in the study, as well as to the broader research community.

I very much look forward to speaking with you and thank you in advance for your assistance in this project.

Appendix E - Oral Consent Form – Interview Study

Hello. I'm Ece Uereten I am conducting research about exploring expertise development, interface design and evaluation in neuro-critical care. This interview is part of my PhD studies at the University of Waterloo, in the Systems Design Engineering department in Waterloo, Ontario. I'm working under the supervision of Professor Catherine Burns of UWaterloo's department of Systems Design Engineering.

Thank you for your interest in participating in my research.

Have you had time to read the Letter of Information I sent you?

YES:

Great, then I would like to take a moment to review some main points from the Letter of Information before we continue.

NO:

(Proceed with Letter of Information summary)

Confirm the following to the participant:

- Your participation in this study is voluntary.
- If you do not want to answer some of the questions you do not have to, but you can still be in the study.
- You can decide to stop at any time, even part-way through the interview for whatever reason.
- If you decide to stop during the interview, we will ask you how you would like us to handle the data collected up to that point, whether returning it to you, destroying it or using the data collected up to that point.
- You can ask to remove your data from the study up until approximately **March 2022.**
- Collected data will be de-identified and stored securely on a password-protected OneDrive account/secure computers.
- This study has been reviewed and cleared by the University of Waterloo Research Ethics Board.

Do you have any questions or want me to go over any study details again?

Consent questions:

Do you agree to participate in this study?

If yes,

- Do you agree to your interview being audio recorded to ensure accurate transcription and analysis."
- Do you agree to your study session being video recorded with the understanding that the video recordings will be deleted after the transcripts are finalized?
- Do you agree to the use of anonymous quotations in any thesis or publication that comes from this research?
- (Trainees only) Do you agree to be contacted for a a second research study as described in the information letter (YES/NO)? How do you prefer to be contacted? (e.g., by email)
- Would you like a copy of the study results? If yes, where should we send them (email, mailing address)?
- Do you allow de-identified study data to be stored and used for future research? (YES/NO)

If no, "Thank you for your time."

Appendix F - Oral Consent Form – Usability Study

Hello. I'm Ece Uereten, and this is my colleague Kathleen Schaef, we are conducting research about exploring expertise development, interface design and evaluation in neuro-critical care. This usability study is part of my PhD studies and Kathleen's master studies at the University of Waterloo, in the Systems Design Engineering department in Waterloo, Ontario. We are working under the supervision of Professor Catherine Burns of UWaterloo's department of Systems Design Engineering.

Thank you for your interest in participating in my research.

[*If the LOI was provided in advance*]

Have you had time to read the Letter of Information I sent you?

[*If the LOI was provided in advance and the participant responds that they have read the LOI*]

Great, then I would like to take a moment to review some main points from the Letter of Information before we continue. [*Proceed to review the highlights of the LOI, be sure to include risks and what will happen with their data, and confirm the important points about voluntary participation and withdrawal listed below.*]

[*If it is not possible to give an LOI to the participant, or if the LOI was not sent in advance, or the participant responds that they did not read the LOI in advance, then proceed to go through the full LOI in detail with the participant and confirm the important points about voluntary participation and withdrawal listed below.*]

Confirm the following to the participant:

- Your participation in this study is voluntary.
- If you do not want to answer some of the questions you do not have to, but you can still be in the study.
- You can decide to stop at any time, even part-way through the study for whatever reason.
- If you decide to stop during the study, we will ask you how you would like us to handle the data collected up to that point, whether returning it to you, destroying it or using the data collected up to that point.
- You can ask to remove your data from the study up until approximately **December 2022.**
- Collected data will be de-identified and stored securely on a password-protected OneDrive account/secure computers.
- This study has been reviewed and cleared by the University of Waterloo Research Ethics Board.

Do you have any questions or want me to go over any study details again?

Consent questions:

Do you agree to participate in this study?

If yes,

- Do you agree to your study session being audio and screen- recorded?

- Do you agree to the use of anonymous quotations in any thesis or publication that comes from this research?
- Would you like a copy of the study results? If yes, where should we send them (email, mailing address)?

If no, “Thank you for your time.”

Appendix G - Demographic Questionnaire – Interview Study

Are you currently a trainee in critical care medicine?/ Are you currently a staff physician in a critical care (in Canada or the USA)?

Yes o
No o

Which year of residency are you in?/ How many years have you been in independent practice?

What specialty certifications do you hold?/ What specialty and subspecialty certifications do you hold?

What types of patients are cared for in your ICU? (only staff)

Do you work in a university or community hospital? (only staff)

Do you have residents or fellows training in your ICU? (only staff)

Which gender do you identify with, do you prefer not to say?

What is your age?

Appendix H - Demographic Questionnaire – Usability Study

Are you an ICU trainee physician (fellow) or early career staff physician in Canada or U.S.?
Yes in Canada o
Yes in the U.S. o

When did you graduate from med school and in what country did you graduate from med school?

Which year of fellowship program are you in currently? If you are not in a fellowship program or have graduated, please let us know how many years you have worked in the ICU.

Do you have any red/green or vision deficiency?
Yes o
No o
Prefer not to say o

What specialty certification(s) do you hold?

Which subspecialty certification(s) do you hold?

Which gender do you identify with?
Female o
Male o
Other o
Prefer not to say o

What is your age?

How familiar would you rate your familiarity with ICP monitoring on a scale from 1 to 10? (1 never used before -10 use every day)
To conduct the study, you will need a desktop or computer available in a quiet room. Please check yes, if you can meet this requirement.
Yes o
No o

Appendix I - Interview Study Questions – Staff, expertise

Can you share with us, what tasks are involved in your current role?

If you were to describe your CLINICAL area of expertise within critical care, what would it be and do you have any academic focus- research/education/quality improvement?

When you consider people who are classed experts in critical care medicine, what characteristics are you considering?

Probes: When did you know that you were an expert?

When you think of people who are developing their expertise but not an expert yet, what comes to your mind as the difference between them and an expert like yourself?
Probes: If you were to think back to your training in critical care, what were the main skills you think you developed over time, that made you an expert in this area?

Imagine you have a trainee standing next to you, can you tell us how you are trying to convey your strategies (mental models) to the trainees? How do you explain/teach certain strategies to approaching the patient case?

How do you support trainees to develop their own mental models in critical care?

How can you tell your trainees are progressing? What are some of the indicators you look for?

Which information sources do you use to retrieve past and current data on the patient's health status?
Probes: Which of these data sources do you use to trend past data to enable the prediction of future states?
Probes: How is the flowsheet covering the gaps represented on the bedside monitor, i.e. are you able to effectively understand the past and current states (as a combination of flowsheet and monitor data), to predict the future states

What are difficult neurocritical care concepts where trainees struggle with the most?

What do you find hard to teach them?

Can you share with us, what are the key concepts that you recall about ICP monitoring?

How did you learn about ICP monitoring during your training? (Was the emphasis on ICP numeric thresholds, trends, or specific waveform interpretations?)

If you have a patient who has high ICP, what do you pay attention to when using the monitoring system especially? Probes: How often do you check the patient's status?

What would a critical ICP case look like to you? Why?

How do you teach ICP cases to your trainees? Are there any concepts that are challenging to convey?

Probes: Are there any neurocritical care cases that you think are difficult for trainees to understand? If so, why? Can you share some examples?

When monitoring intracranial pressure, what concepts do you find difficult to visualize on the current bedside monitoring system?

What do you think could help as a visualization feature on a bedside monitor to support explaining normal or difficult ICP cases?

Probes: Or do you find ICP monitoring straight-forward to monitor, assess response to treatment and predict patients' states?

Which other physiologic variables come to your mind that you consider when evaluating ICP values or waveform pattern changes?

Is there a level that makes you act quicker? what do you see on the monitor that makes you say, I'm going to give them all the TIERS of therapy I can think of?

So when you zoom in, how high of an energy state is this? Are they going to die? Is this brain damage or is the brain at risk? How do you see this on a monitor and respond to this?

Neurophysiologic Concepts

Staff physicians were presented three key neurophysiologic concepts.

o Secondary insult dose

o Individualized thresholds

o Trajectory

Staff physicians were asked if there was anything that was missing from the concepts presented.

Appendix J -Interview Study Questions – Trainees, expertise

Can you share with us what tasks are involved in your current role?

What area of expertise and critical care are you looking to develop more during your training?

When you consider people who are classed experts in critical care medicine, what characteristics are you considering?

When would you know you're an expert or what do you need to do to become one?

When you think of people who are developing their expertise, but they're not an expert yet. What comes to your mind as the difference between them and an expert like you?

If you think about your training in critical care, what are the main skills you think you develop over time that also make you an expert in this area?

Imagine you have a staff intensivist standing next to you. Can you tell us how they are trying to convey their strategies to the trainees?

How do you develop your own mental models in critical care?

How can you tell you are progressing? What are some of the key indicators you are looking for?

From your experience, can you share with us what you think trainees struggle with the most when they develop expertise over time?

How are staff physicians supporting to overcome these challenges?

Can you think back to an example with a recent patient; How did you identify the past and current states of the patient's health by using the bedside physiologic monitoring?

What are some difficult neurocritical care concepts that you recall?

How did you learn about ICP monitoring? During your training, was the emphasis more on ICP numeric thresholds, trends, or maybe more in the direction of waveform interpretations?

What are some key ICP concepts you recall?

Which other physiologic variables come to your mind that you consider when evaluating ICP values or waveform pattern changes?

What do you see on the monitor that makes you say I really need to apply all the TIERS of therapy I can think of? Is there a certain level that makes you act quicker?

What would you consider as a critical ICP case?

Visualizations – staff and trainees

We would like to show you some static sketches of ICP visualizations now. These visualizations are designed to be displayed as an addition to the standard bedside physiologic monitor. For each visualization, we will first explain the goal of the visualization and we would love to hear your feedback on how you believe it could be used on a bedside physiologic monitor.

ICP Dose:

For each visualization, a general discussion took place. Some probes include:

- Do you believe there is value in having both styles of dose visualization?
- Is there anything you find confusing in the visualization?
- What is the most appropriate title for the left y axis on the bar chart
- Is average ICP a common way to communicate ICP in the ICU?
- What time frame are you most interested in when evaluating a patient?

General discussion of concerns trainees had with the visualization.

Individualized Thresholds

For each visualization, a general discussion took place. Some probes include:

- In your current workflow, how do you understand a patient's trend and historic values?
- How do you track ICP targets for patients?
- What time frame do you tend to look at for patients with ICP monitoring? General discussion of concerns trainees had with the visualization.

Trajectory

For each visualization, a general discussion took place. Some probes include:

- How is a patient's ICP typically communicated?

- What time frame are you most interested in for evaluating a patient's change in ICP?

General discussion of concerns trainees had with the visualization.

Summation Trend Chart

For each visualization, a general discussion took place. Some probes include:

- What relationship of variables is most important in the neuro-ICU?

- Is autoregulation evident in this visualization?

- Do you have a preference for the different visual representations of the visualization?

General discussion of concerns trainees had with the visualization.

Final Questions

- What visualization or combination of visualizations do you believe would provide the most value on a bedside physiologic monitor. Please feel free to say zero or multiple.

- Based on what we showed you today, is there anything we are missing that would help support decision making, and expertise development in neuro-critical care?

Appendix K - Usability Study Questions

Reflections:

Was everything clear or anything confusing to you on the interface? What exactly?

Would you change anything about your approach to detect, understand and predict the patient's state?

What other approaches would you consider?
Why did you choose to go with the approach used when going through the scenario instead of others?
What were the limitations of the chosen strategy?

NASA TLX:

Mental Demand: How mentally demanding was the scenario?0- Very Low 100-Very High

Physical Demand: How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating etc)? Was the task easy or demanding, slack or strenuous, restful or laborious? 0- Very Low 100-Very High

Temporal Demand: How much time pressure did you feel due to the rate or pace at which the scenario occurred? Was the pace slow and leisurely or rapid and frantic? 0- Very Low 100-Very High

Performance: How successful do you think you were in accomplishing the goals of the task set by the experimenter? How satisfied were you with your performance in accomplishing what you were asked to do? 0-Perfect 100-Failure

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance? 0-Very Low 100-Very High

Frustration: How insecure, discouraged, irritated, stressed or annoyed (high rating) versus secure, gratified, content, relaxed and complacent (low rating) did you feel during the scenario ? 0-Very Low 100-Very High

Confidence rating: How confident do you feel about your approach and handover? 0-Not confident at all 100-Very confident

Final remarks:
Is there anything you would like to share with us (e.g. on the study, scenario, interface)?

Appendix L - Usability Study Scenarios

Scenario 1 (easy case: Subarachnoid Hemorrhage):
"Norma Jean is a 52-year-old female patient who presented 30 hrs ago to a peripheral hospital with sudden decreased level of consciousness (initial GCS E2V2M5). She was intubated in the ED and an emergency CT/CTA demonstrated diffuse subarachnoid hemorrhage with severe intraventricular extension (mFISHER 4) secondary to a ruptured anterior-communicating artery aneurysm. She was admitted to your ICU 24 hrs ago, and her repeated CT on arrival demonstrated worsening hydrocephalus. She therefore underwent emergency insertion of bilateral frontal EVDs. She is currently mechanically ventilated and sedated with a low-dose propofol."

Scenario 2 (difficult case: Traumatic Brain Injury):
"Jessica is a 26-year-old female patient who has been admitted to the ICU after being involved in a high-speed motor vehicle accident. She sustained a severe TBI (initial GCS post-resuscitation E2VNTM4), with evidence on the initial CT head of multicompartmental intracranial haemorrhages that have not required operative management until now. An intraparenchymal ICP monitor has been inserted to continuously monitor the ICP and an external ventricular drain (EVD) has been inserted to facilitate CSF drainage. She just came back from a follow-up CT head and the nurse calls you at bedside because something has changed…"

Appendix M - Training of Interface - Usability Study

Hello and thank you so much for your interest in participating in our study on interface design in neurocritical care. In this training video, we will walk you through the different parts of the interface that will be tested and evaluated in this study.

Vitals Tab

The first screen you will see is the vitals screen. This screen is similar to commonly used bedside physiologic monitoring in the ICU currently. On the top header, you can see that there are two other tabs that I will explain further later in this video. On the header, you can find the patient's name and sex, Rebecca Wong, Female. In addition, the current date and time is also displayed in the top right- hand corner. Note that this date and time may be different than the current date and time as we are simulating a scenario.

The “i” icon on the top left-hand corner has information about the scenario that will be evaluated, and the patient’s flowsheet filled in by the nurses. For this patient, the flowsheet is currently left blank. In the scenarios presented to you later, you might find different documentation times which might be different to the common starting time at 9:00 am.

The main section of the screen includes the patient’s real time vitals including HR, ABP, ICP, SPO2, Respiratory rate, Cerebral perfusion pressure and end tidal CO2. Each vital sign has an upper and lower limit shown by the top and the bottom number to the left of the real time vital sign.

Looking at the bottom of the vitals screen, the only clickable interactive element is patient info for this study. The patient info pop-out gives basic information on the patient such as the patient ID, name, sex and age.

On the right-hand side of the screen, there is an ICP summary for the patient. Below this summary at the bottom of the screen is where you can select the time frame of interest and the two options are the past hour and past 6 hours for that patient. This ICP summary recalculates every 30min. At the top of

the ICP summary, you will find the ICP time burden bar. This bar represents the percentage of time that the patient has spent in various ICP ranges which we refer to as “burden”.

The green range represents normal ICP and ranges from 0-15mmHg (millimeter mercury). The yellow range represents transitional ICP and ICP values between 15-20mmHg. The light red range represents high ICP and ICP values between 20-30mmHg and the dark red represents very high ICP and values of ICP greater than 30mmHg.

The percentages on the bar graph indicate the percentage of time spent in each ICP range for that selected period. For this patient on the 1-hour summary, they spent 25% of the past hour in a normal ICP, 25% in a transitional ICP range, 40% in a high ICP ranges, and 10% in a very high ICP range. At the bottom of the ICP summary, the trend for ICP over the selected time frame is shown as well as the percent change of ICP.

A positive number indicates an increase in ICP over the selected time frame. Below the ICP trend, the average, minimum and maximum ICP are shown for the selected time frame.

One last element on this screen is a notification icon on the top right-hand corner to the left of the date and time. A notification for autoregulation reading “impaired autoregulation?” will appear if the % change in 1 hour is greater than 30% (and the ICP value is more than 15 mmHg) or if ICP is greater than 20 mmHg. This does not mean the patient is not autoregulating, however, it may be something to consider based on how ICP is changing.

ICP Burden

The second tab is the ICP burden that can be accessed by the top header. This screen contains information on the patients’ ICP trend over the selected period and how much time the patient has spent in each ICP range. Looking at the bar graph, the left y axis represents ICP Time Burden as a percentage, the right y axis represents the average ICP, and the x axis represents time. The white line represents the average ICP and is read on the right y axis, whereas the bars are read on the left y axis.

Looking at each of the bars, in this view, the bars represent a 1-hour period with each color on the bar showing the amount of time that the patient spent in a particular ICP range. For example, from 2:00 to 3:00, this patient spent 88% of their time at a high ICP range, 12% of their time at a transitional ICP range and their average ICP over the hour was 22mmhg. From 3:00 to 4:00, this patient spent 68% of their time at a transitional ICP range, 32% of their time at a normal ICP range and their average ICP over the hour was 16mmHg. If you hover over the bar, the max ICP value for the hour will appear as well as the percentages for each range including the average.

If you forget the ICP ranges, the ICP ranges are described by the top right box adjacent to the graph which shows the high and low value for each respective range. These ranges are also shown on the top of the bar graph in the legend and by clicking on the box, you are able to hide them from the graph. This is also a feature for the average line if you click on the white circle.

To adjust the time scale of the graph there are two options. The first is to use the slider at the bottom of the graph. This slider shows the bars for high ICP. You are able to increase the number of bars shown or navigate through the timeline for that day. The second way to adjust the time scale of the bar graph is to use the duration selector to the right of the date. The date selector is not interactive on the interface as we will only be showing 1 day of data.

The last section of this screen is the annotations. For the time-period selected, annotations for medication, patient position etc. will appear and can be filtered by annotation type using the filter icon to the right of type. The time stamps of the annotations can also be found at the bottom of the bar graph as blue points along the x axis and the timeline adjuster. For this view, you can see there were annotations between 2:00-3:00 and 4:00-5:00 which corresponds to the annotations panel on the right- hand side.

Pressures

The third screen is the pressures screen and can be accessed by the top header. Another way to access this screen is by clicking the hour of interest on the ICP burden screen and then the detailed view will open on the pressures tab. For example, on this patient you may be interested in looking at 3:00-4:00 in more detail because there is an annotation indicated by the blue dot and in addition, the ICP was high for this hour. (You can click on the hour in the x axis to bring you to this time frame on the pressures tab.)

The top graph is a summation graph for CPP and ICP which sum to MAP based on the key hemodynamic equation that CPP is the difference between MAP and ICP.

In this graph, CPP represented by the purple area and ICP represented by the light grey area are stacked on top of one another to give the sum which represents MAP represented by the red line. For example, at 00:20 the patients' CPP is 60mmHg, and their ICP can be read by the height of the white area which is about 30mmHg at 00:20. The patients MAP is the sum of both of these variables (ICP and CPP) represented by the red line and is approximately 90mmHg at 00:20.

The exact values are available when you hover over the time of interest. On the top graph, the variables can be viewed one at a time by deselecting the variables by clicking on the legend for CPP, MAP and ICP.

The bottom graph shows the patients ICP as a grey area and threshold targets for transitional, high and very high ICP represented by dashed lines.

These threshold targets are the same as those represented on the ICP burden screen. These targets may be adjusted by a senior critical care clinician or by orders from neurosurgery. On the current interface, you will not be able to adjust these.

For example, if a patient's ICP was stable at 18mmHg, the visualizations may be adjusted so that 18mmHg represents a normal ICP for that patient rather than transitional. When targets are not customized the default targets for each range are less than 15mmHg for normal, 15-20mmHg for transitional, 20-30mmHg for high, and greater than 30mmHg for very high.

For this patient, from 2:00-2:30, the patient's ICP was in high and it dropped down to a transitional range and then normal range starting at about 4:00. If the ICP grey area is below the yellow dashed line, this indicates that the patient is in a normal ICP range. If the ICP grey area is above the yellow dashed line and below the light red line (or second dashed line) this indicates that the patient is in a transitional ICP range. If the ICP grey area is above the light red line and below the dark red line, this indicates that the patient is in the high ICP range. Lastly, If the ICP grey area is above the dark red dashed line, this indicates the patients ICP is in the very high ICP range.

The time scale can be adjusted the same way as the ICP burden tab and the annotations are also viewed on the graph.

Conclusion

At this point, you have completed the training video for the interface. The interviewers will be happy to answer any clarifying questions you may have about any of the graphics or sections of the interface prior to beginning the scenarios. Thank you.

Appendix N - Preparing the interface for the usability study

The interface has been prepared as a team with Rayyan Quraishi who implemented the interface, while Kathleen Schaef and I have been providing feedback and design directions on the visualizations shown. Data from the interviews and research background were used for the overall design of the interface. As examples, personas and storyboards were developed. To implement the interface different technologies and frameworks were tested (e.g., GatsbyJS, Contentful, ChartJS, Recharts, D3.js, NextJS, MongoDB). After testing various options, we initially used GatsbyJS as the primary framework on top of ChartJS for charting and graphing. Contentful was used to source the data. At later stages, we required more flexibility and thus used NextJS, Recharts & D3, and MongoDB as frameworks, charting tools and database. The website was hosted on Vercel's hosting platform. To find data that could be used for vital signs' numeric values and waveforms and examples of TBI cases, physionet.org was found to be a reliable source (it is managed by members of the MIT Laboratory for computational physiology). Here, we were able to find data provided by other researchers who published on TBI and thus had relevant data (e.g., ABP, ICP, HR, SPO2, RR) in this context too (a publication of Kim et al., 2016, and Goldberger et al., 2000 for physionet). Snippets of the data sets were taken to create the visualizations. Some variables were calculated from the physionet data (e.g., MAP was calculated through ABP, or CPP=MAP-ICP). Data was sourced from CSV files. The big data sets helped to create the waveforms of the vital signs. Further, a 5-minute loop was created for the vital signs' waveform representation. We required a larger data set for the trends shown in the other tabs. Thus, our clinical collaborator Dr. McCredie created an Excel file including all variables represented on the two additional tabs in a 1-second interval and correlation of variables. Two scenarios were outlined first followed by the data set that was created accordingly.

Appendix O: SA tables

Table 12: Scenario 1 - SA level 1

	Standard interface	EID
Initial impressions (level 1)		
Themes	Codes	
Vital signs check	• Numeric values check within range or not • Waveform check (e.g. brain compliance p2>p1; autoregulation/waveform looks fine, waveform non-compliant) • Initial concerns raised (e.g. herniation, dysfunctional autoregulation) • Initial questions raised (e.g. unsure if on vasopressors)	• Numeric values check within range or not • Observations on trend of ICP • Initial concerns raised (ICP hypertension) • Initial questions raised (e.g. HOB elevated, ICP monitor location of insertion) • Correlation of response to therapy

Table 13: Scenario 1 - SA level 2

	Standard interface	EID
Concerns raised (level 2)		
Themes	Codes	
Physiologic Variables	• Adjustment of CPP, CPP could decrease • Control of BP (SBP) • Control of ICP • HR looks ok • etCO2 is low • Watch out for sodium abnormality • Keep MAP below 90 • ICP might increase • Waveform: non-compliance Motor response is improving but not opening eyes	• ICP trend increasing • CPP decreasing trend • Pupils look ok Waveform is concerning
Effect of variables on health state	• Rebleed • No concerns yet about vasospasm/DCI • Hypotension	• Rebleed • No concerns yet about vasospasm/DCI • Herniation • DCI

	• Bradycardia with Cushing reflex • (prolonged) hydrocephalus • Herniation	
Patient state generally	• Patient is ok • Values not that high to take action now • Unsecured aneurysm	• Not concerned about anything acute • Worried about cerebral perfusion • Worried about ICP trend (average increasing) • Unsure about medication taken to decrease BP
Procedures	• Could hyperventilate the patient to bridge to aneurysm intervention • Check EVDs • Might need a shunt	• Might require a CT • Wouldn't do any other intervention for now • Unsure about drainage target • Check blood work (glucose, blood gas analysis correlation with etCO2)
Correlation of response to therapy		• ICP might be high due to sedation, coughing, pain (low on propofol), waking up, unsure about EVD position • No reception of chloride/mannitol yet, no treatment for ICP hypertension • Response to therapy is well (after EVD)

Table 14: Scenario 1 – level 3 SA

	Standard interface	EID
Short and long-term risks (level 3)		
Themes	Codes	
Physiologic variables	• Adjust CPP • Control CO2 • Decrease ICP • BP can relax once aneurysm is secured	• If ICP increases and CPP decreases: worried about cerebral perfusion • Maintain CPP
Proposed procedures	• Sedation • Drain CSF • Clip/coil, secure aneurysm • Check EVDs • Shunt	• Sedation • Check drainage volume and function, Drain CSF • Labs: blood glucose, patient position • Correlate blood gas analysis with etCO2 • BP medication • Give hypertonics, hyperosmolar therapy

		• If hypertonics don't work, call neurosurgery to assess EVDs • Start saline if ICP 20 • Could try to wake her up (since propofol was decreased and is low dose) • Check pain and medication • Repeat CT if nothing works • Secure aneurysm • Consult INR
Risks	• Rebleed • (prolonged) Hydrocephalus • Vasospasm (longterm) • DCI (longterm) • Hyponatremia • SIADH 9 (syndrome of inappropriate secretion of antidiuretic hormone) • Cushing reflex • Secondary SAH • Ischemia • Vasoconstriction • Infections • EVD clotting • Level of consciousness might worsen • Bradycardia • Disabilities (longterm) • Hypertension • No hypertension or hypoxia for now	• Rebleed • Edema • (delayed) cerebral ischemia • Vasospasm (e.g. 48hrs-7 days) • Not concerned about acute neurological deterioration as imminent risk • Blocked EVD but it's draining • Less concerned as long as pupils are normal

Table 15: Scenario 1 – level 3 SA

	Standard interface	EID
Priorities and treatment plan (level 3)		
Themes	Codes	
Physiologic variables	• Adjust CPP (to 60-70) • ICP high but wouldn't adjust for now • etCO2 is within range	• Treat BP • Observe BP, SBP • CPP control • Keep etCO2 low • Raise RR

	• watch out for SBP (<140 or <160) • DBP ok • HR oscillates but not much • keep normoxic • keep sodium balance appropriate • BP control • MAP ok (90) • CPP looks like on a downtrend but wouldn't do anything now • Pupils check: no anisocoria currently (i.e. no pressure on cranial nerve) • Not hyperthermic • Lytes are normal • FiO2 currently good • Could increase RR	
Medication	• Minimize sedation interruptions to prevent ICP spike risk • Keep sedation • DCI prevention medication • Provide nimodipine • Provide mannitol or lassix • Decrease propofol to avoid risk of hypotension • Increase propofol • Give antiepileptic prophylactic medicine • Increase sodium chloride • Consider anti-seizure medication	• Increase sedation (e.g. mannitol), analgesia to reduce pain • Hypertonic saline • Unsure if patient is on pressors
Procedures	• Check EVD • Plan for coiling/clipping • Check in with surgeons and neuro-interventional team • Monitor ICP closely • HOB 30% • Check consciousness, neuro control (e.g. pupils)	• Secure aneurysm • TIER 1 therapy (intracranial hypertension therapy) • TIER 2 therapy • Consult about aneurysm • Don't wake the patient up • Wake up patient if vitals stable • Neuro check

	• CT check	• Check EVD drainage • Elevate HOB • Call neurosurgery • INR • CT if GCS or neuro changes • Check electrolytes and blood work (glucose)
Risks	• Hypertensive • Prevent risk of hypotension • Transtentorial herniation • If GCS lowers, that's worrisome • Brain could swell up • Control bleeding • Avoid fever • hyperglycemia	• Rebleed • Vasodilatory effects

Table 16: Scenario 1 – level 3 SA

	Standard interface	EID
Communication (level 3)		
Themes	Codes	
Clinical personnel	• Neurosurgery • Neurovascular • Nurses • Anesthesiology • Endovascular • Radiology	• Neurosurgery • Neuro-interventionalist • Neurovascular • Nurse • Attending fellow, resident
Patient family	• Patient family	

Table 17: Scenario 1 – level 3 SA

	Standard interface	EID
Summary and handover (level 3)		
Themes	Codes	
ICP uptrend	ICP increases	• ICP increases • ICP increased while CPP dropped • Poor waveform compliance • HR also increased
Response to treatment	• Sedation • drainage	• good response to EVD
Communication with others	• neurosurgery • team	• neurosurgery • radiology

	• neurology • endovascular team	
Suggested procedures	• scan/CT • drainage • osmotherapy • sedation • consider coiling • further monitoring (e.g. of variables: GCS, neurostatus, pupils, fluid balance, urine output, ICP, CPP, HR, BP, oxygenation, ventilation, pain control, etCO2) • mechanical ventilation • HOB 30 degrees • No pressure on the head	• Scan/ CT • Check that EVD is appropriately draining • Maintain good ventilation and sedation • Wait for plan to secure the aneurysm • Close monitoring (ICP, neuroexam, pupils, control MAP and CPP) • BP target should be higher • More BP control once aneurysm secured • Surgery only if ICP crises are acute, consistently over 20 and no ICP control possible
Previous procedures	• EVD placed	• Hypertonics • Increased RR • Increased CO2 • Already on propofol which might become an issue • Patient came in after the OR • Unsecured aneurysm
Risks	• Has hydrocephalus and could rebleed • Vasospasm • DCI • Transtenotorial herniation • Hypertension • Bradycardy • Agitation	• Rebleed (if not clipped) • Vasospasm • BP augmentation • Already on propofol which might become an issue • Hydrocephalus might get worse
Suggested therapy/medication	• Increase sedation • Prophylactic antiepileptic treatment • Let the patient sleep another day • Nimodipine • Increase ventilation to keep etCO2 in the low range	• Sedation (increase) • Analgesia • Neroperiphine • Low threshold for more hypertonic saline • Consider coiling if patient wasn’t coiled • Nimodipine to prevent vasospasm • Titrate down propofol and the goal is to wake her up

Table 18: Scenario 2 – SA level 1

	Standard interface	EID
Initial impressions (level 1)		
Themes	Codes	
Vital signs check	• ICP uptrend • ICP looks appropriate • ICP was around 22 for a few hours: would have reacted to that but now it's lower • Waveform looks appropriate • CPP over 70 ok if ICP normal • CPP ok • CPP elevated • CPP a bit too low • HR increased • HR fine • BP is above target due to TBI (this is fine), check ABP • Would augment BP a bit • etCO2 reasonable • SBP, MAP are ok for TBI • Saturation ok • RR ok • RR a bit on the higher side • Make sure CO2 on ABG is related to hemoglobin • MAP stable • MAP looks unstable with values around 100 sometimes • SBP stable • Vitals look fine	• ICP uptrend • Spikes of ICP • ICP trend high and very high • ICP peaks correlated to lower CPP • ICP ok in transitional • ICP reasonable • ICP was in dangerously high ranges over the past 24 hours • Hemodynamically stable • CPP ok • BP is preserved • Saturation OK • etCO2 ok • ABP ok

Suggested procedures	• Interest in seeing results of scan/CT to check rebleed or blockage • double check CT • Drain CSF • Increase vent settings • Check vent settings • 'high ICP management' • Osmotherapy while investigating the cause • Double check tubes • check if a FAST has been done • check abdomen softness • repeat GCS • check ECG • could increase RR to keep etCO2 around 35-40 • Pupils ok • Strength in limps ok • Check ICP waveform • Double check positioning	• Unsure if CT was made and CSF was taken • Check CT • See EVD placement to check if there are signs of edema • No surgical treatment yet but has EVDs • Consider ICP values from both monitors and correlate • Check positioning
Ruled out therapies/diagnoses	• No Cushing effect • No focal deficit (eyes) • Is slightly sinus tachycardiac • No brain compliance • No respiratory distress • Not too much sedation that would have a repressive effect • Motor response improving on coma scale	• No emergency at the moment • Low brain compliance (p2>p1) • Lundberg A-waves (unsure of the cause) • Maybe patient woke up or was moved when going for CT
Communication	• Nurse (why did she call) • Usually alerted by nurse if ABP increased	• Nurse (ask about concerns)
Previous procedures	• EVD shows progressive increase • CSF hasn't been drained (maybe the EVD was clamped thus ICP spiking) • Drainage not too much • Unsure where EVD monitor is placed	• unsure whether vasopressors have been given (doesn't seem so) • response to therapy is well (ICP currently is 15) • 3% saline was given only once • 1 dose of hypertonic saline worked but ICP increased again

	• Propofol fine	• Previous procedures did not help ICP to go back to normal • Would have done anything earlier
Risks	• Blunt cardiac trauma	• brain trauma patients can have other issues too e.g. pulmonary contrusion, bleeding, pneumothorax, fractures • signs of edema • worsening edema • mass lesion
Suggested medical therapy	• check sedation • could increase propofol to reduce pain • Treat with 3% (saline) • fentanyl	

Table 19: Scenario 2 – SA level 2

	Standard interface	EID
Concerns raised (level 2)		
Themes	Codes	
ICP uptrend	• High ICP	• High ICP • ICP went down without treatment at some point (maybe during scan) • ICP does the same thing as MAP (it shouldn't be doing that) • CPP has drops • CPP is going below 60s • Unsure about cause of spikes • Patient currently seems to improve
Suggested therapy	• Check TIERs of therapy • Check HOB • Check temperature • Consciousness control	• Might require escalation of TIERs of therapy • Check if motor response is improving • Check oxygenation and position changes • Check vasorestrictions/vasodilation explaining high ICP • Check termperature/febrile • Check neurological states (pupils) • Would like to check calibration of EVD at the bedside and prescription of neurosurgeons and colleagues

Suggested medical treatment	• Check sedation • Check pain control • Might give hypertonic fluid	• Didn't receive vasopressors to compensate BP during Lundberg waves • Exams of sedation • Would try 23.3% saline
Risks	• Rebleed • Herniation • Risks could yield to brain death	• Currently not herniating • Decompressive craniectomy might be needed • Concerned about ischemia in the second wave
Diagnosis	• hypertension	• decreased brain compliance • Lundberg waves
Previous therapy/procedures		• EVD working fine • Unsure why EVD is not draining more • Unsure if patient was rebleeding • Unsure when the last scan was done • Unsure if rebleed was due to monitor insertion • Wants to understand thinking of physician who prescribed osmotherapy • Saline was given 3 hours ago and would have expected an immediate response • Questions: was the patient getting turned, suction, sedation off
Communication		• Discuss with neurosurgery if it worsens further

Table 20: Scenario 2 – SA level 3

	Standard interface	EID
Short and long-term risks (level 3)		
Themes	Codes	
Risks depend on ICP trend	• Increase of ICP every hour: need to see what happened before (not easy to see on flowsheet)	• Could have another high ICP period due to further brain injury • If her ICP is under control, would be interested to see if the patient can handle a higher CPP to get her back to normal autoregulation • If ICP cannot be controlled: might have permanent deficits or die
Longterm risks	• Neurological recovery not optimal/ disability	• Secondary brain injury

	• Swelling might continue in the next 5-1o days • VAP • Sepsis • Lines infection • EVD infections • Pneumonia • General ICU infections • TBI cases can have better outcomes after 6-12 months after injury	• Likely to survive but unsure about brain, function and capacity at this point • No promising neurological recovery • Brain death • Problems with cerebral perfusion • Bleed • Ischemic complications
Short-term risks	• Brain swelling, requirement for surgical intervention if management fails • EVD clamping can lead to increased ICP • ABP and ICP may increase, HR might decrease • EVD might not function well for drainage • Respiration problems • Blood flow within the brain could be impacted, herniation (within the next couple of hours) • Edema • Secondary brain injury • Sustained bleed might lead to hypotension and no perfusion in the brain • C-spine injuries • hydrocephalus	• diffuse intracranial hypertension, could be refractory • herniation • death • no decompressive hemicraniectomy fairly soon • has dangerous spikes that are really high • keep close eyes on pupils, make sure it's not changing • impending herniation
General risks	• "I don't have enough information, it's too soon to prognosticate" (P14)	• BP varies, so make sure CPAP level is low and no signs of agitation or harass • Patient improves now • Swelling will happen faster because the patient is 26 years old. It's important to get the her through the swelling period without herniating (even if she had a craniectomy); many minute to minute and hour to hour decisions • Ventilation risk • Infection risk

		• Immobility risk • Weakness/decondition/agitation when waking up risks • Herniation • Ischemia • Edema • Rebleed • Ischemic strokes
Proposed procedures	• Transcranial doppler and serial examinations • Check HOB 30 degrees • Check C-spine collar • Pain control • Maybe paralyze for mechanical ventilation • Seizure prophylaxis • Blood sugar control • Control temperature • Check relationship between PaCO2 and etCO2 • Make sure they're not biting on the tube and endoventricular tube is still inserted • Listen to lungs (check chest injuries, pneumothorax, hemothorax) • Check cardiovascular if bleeding appears (into chest, abdomen, pelvis, lungs, bones) • Neurological check (pupils equal and reactive), GCS, localizing signs	• Act on ICP and CPP • Get more (neuro) monitors • Prepare the family • Give opiods • Medical therapy has not been maximized yet: try bolus, 23% hypertonic, alternate with mannitol if it doesn't work • Increase medical therapy and higher goal for drain/drain CSF • OR if medical therapy doesn't work • Could increase sedation, more hypertonic saline to shift her autoregulation back • Sedation and possible paralysis can be done • Call neurosurgery • Check EVD drainage • Repeat CT

Table 21: Scenario 2 – SA level 3

	Standard interface	EID
Treatment plan and priorities (level 3)		
Themes	Codes	
Physiologic variables	• Control temperature • Decrease temperature to 35-36 degrees	• Control BP based on knowing about pressors

	• Check ICP and CPP getting better to make sure there is no coagulopathy • Keep ABP 140/90 • Get SBP less than 140 or 160 • Check sodium to be on the higher side • Keep normal oxygenation • Avoid CPP above 75 or 80 • Keep CPP to remain 60/70 • ICP below 20	
Medication	• 3% mannitol • Make sure she's sedated/increase sedation, analgesia • Seizure prophylaxis • Hypertonic fluid • Consider increasing fentanyl or propofol • Would give 3% or 23% saline if ICP is too high • Consider paralyzing patient (rocuronium bolus or infusion) • Extend saline • Wouldn't consider antiepileptic medication as in SAH case • Start antiedema treatment • Consider hypertensive medication if ABP 140/90 • Manage pain and sedation	• Mannitol, titrating/redose hypertonic treatment/sodium • Benzos (prophylactically if seizures in acute phase) • osmotherapy
Procedures	• talk to neurosurgeon or TTL • ICP management: position, HOB, C-spine collar • Drain more CSF, 15cc/hr • When ICP is high, drain 1-2 cc to decompress the brain just acutely • Check EVD draining • Consider autoregulation challenge • Check if patient is in pain • Check if patient is too awake • Consult more experienced (neuro)intensivists • Check kidney tests	• Next wave: treat more aggressively • Ask nurse if spikes relate to care (provided or not-induced) • Assessment of rebleed in CT • Troubleshoot EVD if it doesn't work (should drain) • Rediscuss with neurosurgery • TIER 2, if TIER 2 hasn't been given she needs decompression • Confirm single system injury • Check pupil reactions • Would have treated patient already at 4 am/ in general earlier

	• Pupil, GCS and neuro response check • Repeat blood work with CbC coex	• Might hyperventilate to bridge interventions
Risks	• Rebleed • Edema getting worse? • Refractory intracranial hypertension • Hypothermia • Make sure there is no bleed elsewhere • Prevent secondary injury • Prevent hypoxemia and hypotension	• Lack of autoregulation and increased risk for secondary brain injury • coughing or discomfort or secretions might correlate • Expansion of bleed • Make sure patient is not seizing

Table 22: Scenario 2 – SA level 3

	Standard interface	EID
Communication (level 3)		
Themes	Codes	
Clinical personnel	• Neurosurgery • Nurses • Radiology	• Neurosurgery
Patient family	• Patient's family	

Table 23: Scenario 2 – SA level 3

	Standard interface	EID
Summary and handover (level 3)		
Themes	Codes	
ICP uptrend	• ICP uptrend, borderlined • High dose ICP and spikes • ICP is fine at the moment but keep a close eye • Keep CPP around 70-75, 60-70 • Make sure SBP and MAP are adequate at all times (SBP 160 or 140) • etCO2 should be around 35-40 • check HR • young patient • ABP increased and decreased • During admission stable signs with MAP 80-90 • ICP increased all day without increasing the settings or medication	• Periods of high ICP • Had several ICP crises • In the past hour ICP is controlled around 15 • CPP around 70

Response to treatment	• easy to oxygenate and ventilate • ICP corresponds to the EVD, drainage is 5-10 ccm, normal function	• If spikes are related to nursing care: avoid stimulation until therapy is under control
Communication with others	• (neuro)surgery • TTL	• Advise the nurse if ICP is over 20 for 1-2 minutes • Neurosurgery to drain or for decompressive craniectomy • radiology
Suggested procedures	• Check EVD • Check and drain CSF • Check CT, reconsider CT if refractory ICP increase • Consideration of decompressive craniectomy • Check acute change in GCS, motor response and pupils • Make sure all other injuries are addressed • Be careful with temperature management • Repeat blood work with CbC coex • Check ABG to assesss PaCO2, acid base and lactate check • HOB 30, C-cpine collar	• Check and drain more EVD • Check scan • Treat aggressively • Might go to surgery • Low threshold for decompressive craniectomy to call neurosurgery • Try all strategies • Young patient
Previous procedures	• Unsure if EVD is clamped or not working • Waiting for scan • Maintained sedation, ventilation • Sedated with minimum dose of propofol • Ventilated with full support, etCO2 normal range	• Currently EVD open to drainage of CSF • EVD not draining a lot • Received 3% bolus but ICP increased up to 20 • Gave her 23.4% saline • Only one dose of hypertonic saline given
Diagnoses	• Was tachycardic and hypertensive • Pupil reaction is ok	• Noncompliant waveform, autoregulation off • Normothermia initially
Risks	• Swelling • Intracranial bleed • Check HR for no development of bradycardia • Brain herniation	• Seizures • Bleeding • Complication of monitor insertion • Concerned about exam changes, pupil changes

	• Currently she has decreased brain compliance	• Could herniate very quickly • Refractory ICP • Poor neurological recovery
Suggested medical therapy	• Give 1 dose of 3% saline • ICP control • Sedation paralysis as bridge to neurosurgery • Stop sedation for neurovitals every hour as long as ICP is fine • Keep sodium on higher side • Make sure no pain and optimal sedation	• 3% saline • More hypertonic saline • Mannitol • Sedate more • In more advanced situations give neuromuscular blockade • osmotherapy

Appendix P: Reflections

Table 24: Reflections on the study using the EID

Themes	Sub-themes	Direct quotes
Interface comments	Visualization used as confirmation for reasoning	“The summation graph is useful to confirm the conclusion or reasoning.” P1
		“The EID looks pretty nice and is probably much more accurate. The greens are reassuring. Probably, I would use the same approach with the targets, but this way it would be more reassuring, every second counts.” P4
		“The interface is very sophisticated. The ICP dose is incredibly helpful.” P6
		“I know what to look for when I see the ICP summary on the vitals.” P7
		“She has no green in the ICP summary bar which is worrisome. The trend is decreasing though, when I see this, I directly want to go to the ICP burden tab. [Looks at last 6 hours view on ICP summary,] it’s very good to have this on the first interface. [clicks on and off the low/transitional colours on ICP burden tab] she's very sick, we are undertreating this patient or we are not annotating what we are doing to her. I'm much more concerned than in the standard interface.” P14
		“The ICP summary is useful when you know what happened during the day like a brief summary, useful for monitoring during the day.” P11 “I also love the ICP summary: this is very helpful.” P6
		“My brain struggles with the stacked chart, easier would be a pie chart.” P6
		“The EID makes it fool-proof, interventions might not look good enough, for someone not working regularly in neurocritical care this is extremely helpful, trend, annotations. The EID summarizes very well the flowsheet info and is more digestible for the user.” P17
		“The ICP burden shows annotations together at 8 pm with saline could help to think about why this happened. [Pressures tab:] There is no significant improvement after saline was given. I would have been much more on top especially in the SAH case and immediately called neurosurgery, check ABG, and started osmotic therapy, sedated, and paralyzed. Probably I would have hyperventilated the patient. Also, in the TBI case, I would have been more aggressive about management on ICP. Would have given one more dose of osmotic therapy and would have also hyperventilated the patient a bit more. I would have called neurosurgery early, do they need decompression, it is a young patient.” P17
		“I would only look at the Pressures tab in the extreme patient when I don't know what's going on. If I can get the information from the vitals tab and ICP burden tab, I probably wouldn't go to pressures.” P17
		“I might have not seen the spikes through the flowsheet. The spikes visually shown help. First, I focused on interventions. I would have to correct myself: it was not coughing because it was for multiple hours.” P18
		“The ICP burden is easier to detect. The EID is clearer what happened overall on the whole day.” P18

	Clinical relevance	"I don't know if the bars are useful clinically if there are frequent reassessments happening." P3
	Alerting signs	"I don't personally like the bars but they say 'look at all the red' and do something, from that perspective. I don't like the colours, they are jarring, and once I use the monitor, I know that ICP is bad. It might be helpful for a radiologist potentially." P3
		"The red colours help to even see things are bad from 10 meters away. Without a training video it might be confusing, what the colours represent. It's helpful to see within the box what the colours mean (high transitional normal)." P18
		"Trending up is great to see, and better to prognosticate. It might also trigger the nurse. I Would be interested if there are dampings in the waveform or abnormalities." P18
	ICP trends and values	"The EID is definitely better than what we currently have in the hospital, we can't do the trend of ICP, so this was extremely helpful." P1
		"The actions are based on the numbers or thresholds. Maybe in the future it would be more about the ICP burden and not only about the threshold. It would give more the feeling about prognostication. For decision-making, you might still talk to the family, but it might not change much unless there is more data about the ICP burden correlation with the prognosis. The brain is not a straight-forward thing. This is a good start for further research in the area." P4
		"Sometimes the neurosurgeon mentions ICP is high if 22 or 25, not 20. I don't know how they come up with the exact number. We work together. Otherwise, I go with the usual number from guideline." P4
		"Weekly data would be useful too. Typically, patients have the ICP for 7 days usually, after that you might have to change it. 24 hours is good to see the trend." P4
		"I like to see the evolution over time, that's why I put the summation graph at the top." P5
		"I like to see the actual number; it means more to me. Retrospectively, I would have turned off the colours. I would prefer the real-time ICP (instead of average ICP). For spikes, I'm alarmed by the nurses, and I trust them. A very customizable display might also have an effect on who uses it (they might set it up differently). It might lead to conflicts in medicine. Putting up pre-sets are good and needed." P3
		"Customizable thresholds/ranges for ICP are good." P2
		"On vitals screen it is helpful to have the ICP from the monitor and ICP from the EVD, even if there's no reading from the EVD (e.g., when it's blocked)." P5
		"On the standard interface I see data in the moment, so 8 variables maybe are shown in the moment. But what about peaks? The EID could help for trajectory." P15
	Annotations related to time	"Having annotations according to time is very useful. "P1
		"More details to the annotations would be nice." P3

		"You have to make sure that there is an effect between the variables and the annotations. We typically ask the nurses for the relevant cause of action." P4
		"It's also helpful to see what the team has already done. There are usually two reasons why the ICP increased (due to BP, T, CO2 etc.). The annotations are very important, in reality if we document our intervention, this is very helpful (e.g., vent settings, patient position)." P5
		"It is key to have the annotations (e.g., EVD clamped, drainage volume etc.). It is very helpful what exactly the interventions were, and it's so much easier to see all on one spot. [checks pressures tab]: someone got concerned around 8 pm [used the thresholds graph and looked whether the notification popped up:] yes, I would have arrived at the ICP spikes much sooner! You almost encourage me to think about a dream world where I have all the information in one place. I think the one thing is missing here is what is happening with the EVD and the drain, so if there was a way to see what was happening to see what is happening with the drain." P6
		"The ICP burden blew my mind and the annotations. The patient at 2 am had a high ICP and had an intervention that helped a lot, now the patient has been safe for the rest of the time. The values and waveform tracing are great, maybe notifications could help: watching out for normal compliance, for people who are less experienced. The numbers are from literature but we are not sure if that's the right thing so the waveforms might be good to see. Examples of normal and abnormal tracings: P1>P2> P3, abnormal: P1 and P2 higher. Sometimes this also happens when the monitor doesn't transduce properly. In the transitional zone, people might just chill but they shouldn't." P14
		"Autoregulation relates to how sick the patient would get in the next hours. The time of injury would also be good to annotate to keep track where we are on cerebral edema phase etc." P14
	Other interactions	"An icon or clickable button would be useful to click next to ICP time burden, as a definition, ranges and maybe a website as a reference, to explain what it means." P11
Learning with the EID	Individual learning tool	"There's a lot of learning by your own." P1
		"Interfaces are not just something that we interpret but it also prompts us to learn more and treat in a different way. E.g. if I had this information, then I would learn more about how to do autoregulation challenges at the bedside, it would actually generate a new skill set which currently my skill set in that specific domain is not very good because we don't do it and we don't do it because we don't necessarily have that information available. I think this is a very important part of it. I wish we had something like this!" P6
	Quick adaptation	"The learning curve was fast for the interface. In the second scenario, I was just thinking about the case not the interface." P1
		"I love how user-friendly it is. Only seen it once and could immediately understand. It really changed management rather than flowsheet." P7
		"Probably when you use it more and more, you get used to check the average ICP line stronger. Also checking the clinical picture is important." P12

	Use for experts	"The summation graph would really help the more sophisticated physicians on a daily basis, for me it's more like a nice-to-have." P6
		"The pressures tab might be more helpful for the experts rather than the novice, whether it is true impaired autoregulation or not. A person working not regularly in neuro-ICU maybe wouldn't look at it." P17
	Transfer of mental plan	"I know the targets, I don't need them, it might be valuable for the team to share the mental plan and urgency of the situation, but for me personally I don't need the green/yellow/red." P5
		"EID is great for handover." P16
		"ICP targets are not so much useful for patient caretaker, but good to secure communication between teams and know what the targets are. It can be tricky with neurosurgery, if the targets are clear its better." P11
		"Pressures tab would be great for teaching, an attending might show us the relationship between MAP/CPP/ICP and interventions." P17
		"During handover, it could be helpful visually to show when what happened. I would be more confident to say the trend using the EID. During handover it would be very clear when something happened. Training is needed for sure." P18
	Research	"EID helps to get all information. Would also facilitate academic research, could be used for training/learning." P15
		"Also, for research it would be great and for real-time monitoring, as well as for quality improvement. Interventions and responses would be visible." P18
Other variables	Incorporation of other variables in context	PbO2 might have helped to have more individualized thresholds. P1, P4 "I'm interested in seeing the MAP, ICP, CPP and PbtO2 altogether on a graph. I could then figure what's the best for the patient on MAP or CPP. Maybe a line graph." P10
		EEG data would also be helpful in a simple form. P1, P3
		Seeing temperature would have helped (e.g., to check if the patient is febrile). P3, P5, P4
		etCO2 as a graph would be helpful. P3, P5 "PbtO2 compared to the etCO2 would also be great, maybe I want to hyperventilate the patient, if ventilating too much, I want to see the etCO2 and brain tissue waveform." P10
		"Noninvasive monitoring with NIRS could be added if it has benefits." P1
		"I could try to increase MAP around 10 or 20% and see how ICP and CPP would respond. So, changing the MAP target. I could do this especially with the brain tissue oxygenation catheter. I would correlate ICP with MAP and brain tissue oxygenation by increasing (both). I could also check the microdialysis for lactate etc. We often used brain tissue oxygenation last year but not at the moment. I would prefer to use it, it gives more meaningful information, as same as hemodynamic monitoring. If you only have pressures, you don't have the full

		picture, but if you can also see the oxygenation, you understand better what's going on." P10
		"I would want to see all waves also EEG, ECG, EMG. Especially for giving a quicker assessment this would be very important. I would also reduce the physical workload. Representation on stations also outside of the ICU would be very helpful too." P13
	Use of a messaging system for TIERS of therapy	"Overall, one could consider the next TIER of therapy through a messaging system which might be very useful." P3
	Notifications	"Notification is nice to consider but I'm always thinking about it because it affects your decision making, if they are herniating or responsive to treatment. I did really like the relationship graph that was my favorite." P8
Scenario comments	Trends of ICP	"Scenarios were both showing trends. Maybe if there was an active ICP of 40, that might maybe change the management approach. P1
	Scenario 2 complexity	"The first one was a little trend and not necessarily worrisome, it's something we see very frequently, and no one really knows what to do with it perfectly but the second one was not only we weren't sure what to do but we know that we will harm the patient if nothing is done. The second one was definitely more challenging." P1
		"Scenario 2 was more demanding." P11
Usefulness of EID	A resource-efficient tool	"The EID would also be very helpful in other parts in the world where there is no nurse for each patient available (nurses visualize the trends)." P16
		"This interface is definitely more valuable than what I currently have in the unit." P3
		"The interface is easy to use and similar/consistent with other interfaces we use in consumer life." P6
		"It can happen that new devices are produced just because they can and then you have all these flashy colours, big screens and touch screens, they are beautiful, but I mean, I don't want to watch a movie on the monitor." P11
User-friendliness of EID	Time-scale adjustment	"I ranked the summation graph at lowest because the time scale adjustment was hard to use." P9
		"It was useful, easy to navigate, flexible with just enough information." P11
		"We also change the EMRs so often. Many have EPIC. It's extremely customizable it's overwhelming. It can be powerful." P6
Autoregulation	Consideration of autoregulation as a concept	"It's in the back of my head, not something we emphasize frequently but something we should consider more. I have never been asked by senior faculty member. I would look at the reactivity index. We use transcranial doppler, but we don't talk about this concept in a regular basis." P3
		"In the second case, I worry that her autoregulation is impaired, we do not often do the MAP challenge. We don't have all of this information readily available. If I saw the ICP spikes and I was able to recognize that something acutely changed, it would have certainly been on my mind seeing on this interface. On the other interface I briefly thought about it but did not explore. Also impaired

		autoregulation is not a diagnosis, it is a symptom of something. Either way, the intervention or something bad is going on with this patient, you can argue even with the spikes alone you would initiate that treatment while you are on the phone with the surgeon. But the thing that might change here is that it would have prompted me sooner to do a bedside autoregulation challenge." P6
		"As fellows we often don't think about autoregulation challenge, maybe those who are neurocritical care specialized." P6
		"I really liked the interface; it was less than a minute about things I was confused. The first thing I do is look at the trend of MAP and CPP and correlation, was easy to see autoregulation curve." P8
		"I really like it much more, all the information is very relevant. Maybe I need to use it more. A zoom-in waveform would be useful, P1, P2, P3 to get insight on brain compliance. Over a longer period of time, it's good to see Lundberg patterns. At this time, we cannot have a strong opinion about autoregulation and trying to manipulate it. "But if the patient is not doing well and the basics are not working, I would move forward, maybe MAP changed, CPP changed based on autoregulation. Maybe I could find a new target for CPP, maybe the usual target is not adequate. So, for the patients that are not responding to the basics, I think it is a good path." P10
		"This is exactly what I was talking about. This is a bad pattern, often related to cerebral vasoconstriction after vasodilation, so I would be even more concerned about this patient, it gives me a different insight. I don't see a couple of waves only; I see it throughout the time. I recently had a case where the nurse called him and noticed a lot of variation even within the normal ranges, the patient didn't have compliance." P10
		"To understand autoregulation there is a graphic (tries to imagine it and transfer the concept). The EID is easier to show images and graphs to show this visually. Sometimes you do a test, but with the monitor you can directly see it. You would have a recording of the MAP challenge, as a proof." P16
		"There is a pressure challenge we use sometimes. Even if you know that the autoregulation is intact or not, it wouldn't lead to any further action to treat the patient. The same idea is to keep the number in the range. It wouldn't mean I would act differently." P4
		"If you think about autoregulation, you want to think about the relationship of CPP and ICP. Usually, I'm not used to see CPP in graphical representation." P6
		"The Lundberg waves are visible, which is very concerning, there are repetitive spikes, maybe she's not responding to treatment. Her CPP goes low, but we need to subtract 10 mmHG for the HOB positioning. I would be interested in seeing the waveform pattern during the Lundberg area, mostly interesting for the transitional periods. The system is not perfect [refers to transducers]." P14
		"Here we can see that the autoregulation is intact. I wasn't able to see that directly on the standard interface, only ABP hinted a bit towards that, I was considering dysfunctional autoregulation maybe around 50% but on EID I consider it around 80%. I might already consider loss of autoregulation around 1 pm area." P15

		"We talk about autoregulation in our exams, but I don't know exactly what to do about it. Maybe it's also because I don't know enough. The notification hints towards it, I might have to control BP. I know it in theory, but I don't know how I would act on it. I would try to bring down MAP and SPB and check ICP and CPP and see what happens. If there's a linear relationship, there might be impaired autoregulation and I would have to be more tight on hemodynamic targets." P17
		"For impaired autoregulation I'm not sure what people would do. I would need to review the ICP/CPP/MAP graph [pressures]. I would need to plot the concept of autoregulation in my head again. It will take more time to check." P18