Taming entropy: designing, evaluating and applying decision-support systems for risk management in industrial and non-industrial settings

by

Michelle Naef

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

Department of Chemical and Materials Engineering
University of Alberta

Abstract

After being assigned a process fire investigation in my second year of professional practice, and later working in maintenance leadership following multiple fatality incidents, I wondered why process safety outcomes did not seem to be improving despite constant discussion and an incredible pace of technological innovation in both automation and information systems. I pursued research in the field of chemical engineering and risk management to answer the question: *why are process safety incident frequencies rising despite advanced process control and rapidly improving technology?* This question seemed to be a classic process control problem, a "delta" or difference between the way things are and the way business and operations leaders said things should be. Over a decade of experience working in process facilities and control rooms had provided clues to the nature of the underlying problem, and I focused my research on the latent value I believe exists in improving the quality of human decisions at the human-computer interface.

The research summarized in the following dissertation began with a small pilot study where I pursued a practical application of a novel predictive analytic method developed by Suresh et al., (2019). Using the layered approach to user interface assessment, I developed a visual representation of the causal maps generated using the hierarchical approach and asked participants to narrate their problem-solving process when presented with a simulated process fault. It became clear that developing a defensible, reliable method for evaluating risk-based decisions at the human-computer interface could be a valuable tool in determining what types of information system and automation interface features had real utility to field workers. I developed a large-scale study demonstrating an application of the situational design model to a process industry challenge. 35 participants faced an abnormal situation similar to the one

presented to the pilot study participants. The instruments and techniques applied can be further refined and adapted to offer significantly more analytic power to design feature assessment and to industrial training and assessment processes.

In a parallel research stream, I examined the structure of information systems in the process industry, observing that many of the classification structures used in analytics and in human-centric risk management activities were similar, but not standard. Miscommunication and challenges in organizing relevant data seemed to be persistent challenges that limited the effectiveness of these efforts and required frequent intervention by subject matter experts. I wondered how asset management ontology knowledge bases could be applied more effectively to improve understanding of process systems in risk management contexts, and potentially improve the efficacy of predictive analytics by reducing the number of spurious or misleading connections made by algorithms and optimization processes. I developed a case study and applied an asset management ontology as a communication and learning tool for a group of stakeholders engaged in COVID-19 response planning in a commercial building. The results demonstrated the value of industrial risk management approaches to a new application, and the use of ontology knowledge bases as a risk communication tool.

Major loss incidents in the process industry are complex, multi-variate networks that connect people, assets, environment, and information systems. Mapping these connections, uncovering the unexpected consequences of different interactions and better equipping decision-makers at the front line is a necessary step in reducing the frequency of major loss incidents. This research offers theoretical, methodological and empirical contributions supporting new directions for industry in pursuing safer, more reliable operation.

Preface

This thesis is an original work by Michelle Naef. The pilot study described in Chapter 2 was performed under the supervision of Dr. Carrie Demmans Epp as part of the requirements for the University of Alberta course titled "Intelligent User Interfaces" delivered in Winter 2021. The study was compliant with the requirements of the Research Ethics Board approval held by Dr. Demmans Epp. The study described in Chapter 3 received research ethics approval from the University of Alberta Research Ethics Board, Project Name *"Demonstrating the effectiveness of the situational design model in evaluating user interface design features"*, Pro00128584, approved on March 2, 2023. The study concept was developed by Michelle Naef with Dr. Lianne Lefsrud being the Principal Investigator, as part of a larger research project entitled "Developing a Safety Management Systems (SMS) Model", supported financially by the Natural Sciences and Engineering Research Council.

This thesis contains three manuscripts submitted as articles to peer reviewed publications. The contributions of all authors are detailed in the preface to each Chapter. Chapter 2 of this thesis has been published as M. Naef, K. Chadha, L. Lefsrud, *"Task loading in the days of big data"*, Journal of Loss Prevention in the Process Industries. Chapter 3 of this thesis has been accepted for publication as *"Smooth operator: aligning performance assessment methods with design and operating objectives"*, Journal of Loss Prevention in the Process Industries. Chapter 4 of this thesis has been accepted for publication as *"Application of engineering thinking for risk assessment in a Canadian elementary school"*, Journal of Building Engineering.

Dedication

To my wonderful husband Thomas, and my sons Jonathan and Jacob, the reasons I work safe.

Acknowledgements

My sincerest and deepest thanks to my advisors Dr. Lianne Lefsrud and Dr. Dave Shook for their patience and generosity in supporting this journey. It was a great honor to learn from and work with you both, and I am grateful to have been granted the opportunity to question and be challenged. Thank you to Dr. Vinay Prasad for your guidance and support as a member of my supervisory committee, and the suggestion to dig into the Tennessee-Eastman problem. Thank you to Dr. Don Kennedy for your perspective and the levity you injected into our interactions and to Dr. Biao Huang for your feedback as a member of my committee.

This work was motivated by all the coworkers, friends and family who have taught me so much about mechanical equipment, chemical processes, scientific inquiry and the general brilliance of humans, thank you all so much.

Table of Contents

List of Tables

List of Figures

Chapter 1

1.1 Introduction

The chemical process industry is a heavily integrated, international network of industrial sites where raw materials are processed into intermediate and consumer products, from pharmaceuticals to fuel. As the world population increases, demands and pressure on the process industry increase, along with scrutiny from regulators and the public (Theodore & Dupont, 2022). Industrial sites that store large quantities of hazardous materials, chemical processes that harness heat and pressure, high volume transportation routes for feed stock, intermediate and finished products - all pose potential risks of explosion, fire, environmental contamination or toxic release. The risk management efforts of the chemical process industry are of concern to everyone in the world, as supply chains and production become increasingly interconnected globally (Gehman et al., 2016), and “learning from incidents” to predict and prevent future loss events is the cornerstone of process safety research (Dahiya & Dahiya, 2018).

One persistent model remains the foundation of much of the research and development toward reducing process safety incidents – the safety pyramid shown in Figure 1-1 (Lefsrud, 2019a). Figure 1-1 insets the occupational safety pyramid in the larger process safety pyramid which dates to 1931

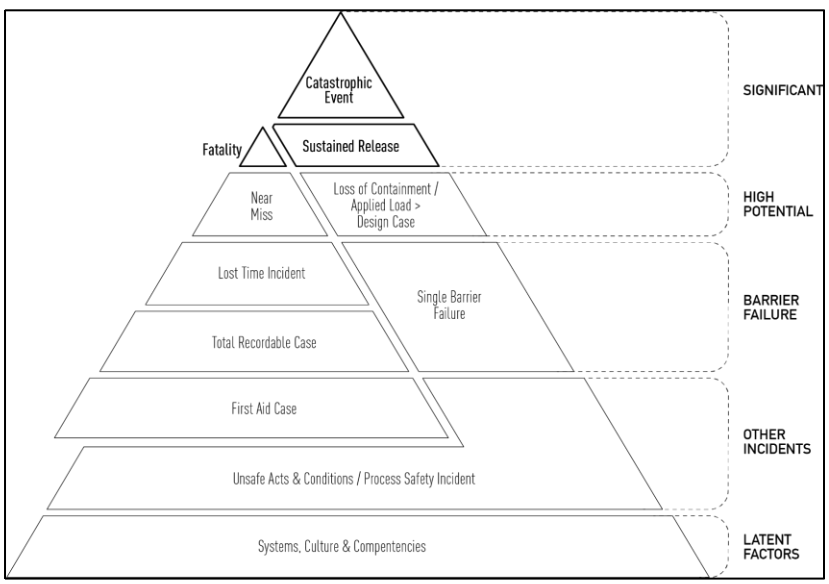


Figure 1-1 Safety pyramid depicting the relative predominance of latent factors like systems and culture, unsafe acts and conditions, barrier failures, and significant catastrophic events

and is attributed to Heinrich. The 2019 version reflects maturing ideas about how the pyramid alone may be limiting loss prevention efforts, challenging the idea that addressing any of the "latent factors" underlying any types of incidents will lead to an overall reduction in incident frequency that includes the reduction of high-consequence loss incidents (McSween & Moran, 2017). The pyramid-based mental model for process safety continues to form the conceptual basis for employee training programs, corporate and regulatory policies relating to safety and the misalignment between the information model and reality suggests an explanation for often-disappointing progress toward addressing the growing concerns around stagnant safety results (Silaipillayarputhur, 2018). Marsden (2018) compares statements from the investigations of two major Process Safety incidents, Longford (1998) and BP Texas City (2005) noting that the perception in both organizations was that declining minor incident frequency meant they were actively reducing exposure to high consequence events. The information systems that support process safety efforts in industrial applications often reflect "pyramid-based" model focusing on standardization, routinization, and task discretization – the idea that reducing any errors over many repeating tasks will contribute to a reduction in the errors that initiate major loss events. This conception neglects the fact that not all errors lead to major loss events, and that some of the causal paths from "error" to "top event" are effectively unknown or obscured. Evaluating the accuracy of the mental models used in risk-based decision making after major loss events is an assessment performed too late.

Process safety incidents are complex causal networks of assets, people and informatic interactions (Gerbec, 2013). Figure 1-2 illustrates the complexity and interconnection between data sources, measurements, assets, and people contributing to incidents in the process industry and shows how some connections are mapped on information models while others are

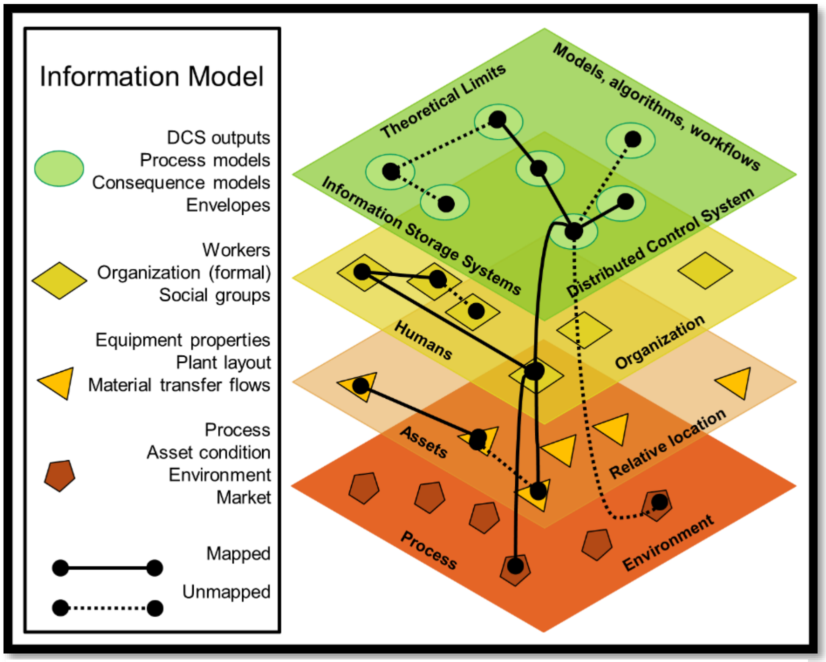


Figure 1-2 conceptual layers interacting in process facilities, information flows identified. Behavior can be conceived to coalesce around measurable information flows.

unmapped. The primary users of process safety information systems are skilled employees, like panel operators, who rely on the information to make decisions. The human interface to the data is important in shaping the perception of risk, and design objectives that do not specifically include improving user understanding of the system and its current state leave latent value in the Human-Computer system. This is of particular importance during the low-frequency, high consequence events of most interest in loss prevention as these abnormal events are poorly identified by automated systems and require human intervention. Excessive focus on "noise"; learning to work with or around ineffective informatic interfaces and focusing on high-volume but ultimately low value observation and reporting undercuts the utility of human processing power (Miller & Parasuraman, 2007). A resilient system requires recognition that well-designed assessment and scrupulous auditing is necessary both to verify establish performance, but also to explore and discover and intermediate processes (Jain et al., 2018). Hidden information flows need to be sought out and mapped, and sources of entropy or disorder, such as conflicting information or incompatible conditions must be contained. This conception illustrates mapped and unmapped flows of information, for example, an instrument sends a mapped flow of information to a transmitter (two assets) or a human collects a sample from the process and enters observations on the colour to the product quality log on the green layer, a mapped flow intersecting three levels. An unmapped flow might involve a malfunction in the digital control system linked to electrical storms, a re-occurring problem that has not yet been discovered by workers. A detailed deconstruction of ontologies and corresponding information system for the purposes of process engineering was presented in (Morbach, Yang, & Marquardt, 2007), lending some credibility to this conception and its utility in the process industry. Morbach et al. provide a comprehensive outline of the discrete information systems involved in defining even basic operational tasks. Determining the purpose and status of the protective barriers as well as the nature of the hazard in an operating facility can require knowledge and efficacy of all linked systems.

Alignment is fundamentally distinct from compliance, demonstrated in a unique study among maintainers in an operating facility (Kanse et al., 2018). When front-line work is complex and requires judgment and decision-making, compliance creates conflict and can ultimately undermine the achievement of business outcomes. Workers might be pursuing high-level compliance to imperfect procedures or pursuing performance metrics that do not assign proportional value to the outcomes of risk-related decisions. Kanse et al. effectively describe the increasing entropy experienced by many process-contacting workers, in attempting to conceptually simplify a management system, complexity and conflict are experienced by people who ultimately have to synthesize incompatible, rigid and inflexible lists of orders. The complexity experienced by workers interacting with management systems is complicated further by the addition of informatic systems, many of which are disconnected and have their own "languages" for input and interaction. The vendors who distribute these information systems intend for their products to be tools to achieving operational and business excellence and often seek to automate routine decision-making and repetitive tasks. This routinization is accomplished, in part, through the adoption of database storage systems and proprietary software packages that effectively alter the decision-making heuristics of the workforce (Agar, 2014a; D. A. Kennedy & Whittaker, 2006; Morgan, 2002). The utility and value of many informatic solutions becomes questionable when the cognitive effort expended by individual workers to simply track routine tasks begins to exceed the effort they are able to expend on interacting meaningfully with the process. There is potential risk associated with the adoption of some management systems and informatics, as incomplete or incorrect mental models can lead to sub-optimal decision-making.

Two broad questions motivated the research documented in this thesis. The first sought to explore why major loss incidents continue to occur, despite advances in automation and a growing body of historical knowledge from which to learn and apply lessons. Human error continues to persist as a finding in root cause analyses, and the study of human factors during abnormal situations demonstrates room for improvement (Perry, 2016; Strobhar, 2014f; Strobhar & Harvey, 2011; Woodcock & Au, 2013). Throughout the development of the field of process safety, there has been recognition of human factors, summarized by (Moura, Beer, Patelli, Lewis, & Knoll, 2017) with the detailing of increasingly complex models used to describe organizations and their operation around high-risk technologies. Research into both operations (Venkatasubramanian, 2019b) and maintenance (Dalzochio et al., 2020) have sought to analyze historical data and develop techniques that will better predict abnormal operating situations and loss events. The second question followed from the first, asking why novel approaches to automation and analytics were not operationalized fully, or achieving the expected potential.

The human interface can be designed to achieve the desired outcomes and build competence over time, or it can suffer the fate of many brilliant DCS modifications and conceptual models: bypassed by panel operators and never used (Shook et al., 2003).

1.2 Research Objectives

The human-informatic interface is a frontier that challenges the operationalization of many advances in research and design (Bainbridge, 1983). The following research objectives were developed and explored in the studies presented in this report:

- Explore the gaps between information models and reality.
- Interrogate the ways in which people acquire knowledge from process informatics.
- Determine what types of informatic features and designs best support risk-based decision making.

As described in section 1.1, information models associated with complex systems span multiple conceptual layers and some connections are unmapped or unknown. These unmapped connections pose a particular challenge to data scientists and analysts when attempting to identify spurious correlations between variables (Venkatasubramanian, 2011). There have been calls to develop common or even standardized information models for application in the process industry (Mathew et al., 2012; Morbach et al., 2007) but adoption of these models, thoroughly grounded in the theory and experience of the domains of interest has not been as enthusiastic as might be expected. The process industry has more readily adopted informatics designed for commercial purposes, like Enterprise Resource Planning (ERP) software, or controls-specific interfaces which do not incorporate the complete underlying architecture of designed informatic architectures. ERP packages, for example, are largely structured to support the recording of financial transactions and inventory management; a conceptual layer vital to the operation of a business, but one that bears little resemblance to the conceptual layer of industrial asset management or operations. The categorization and relationships that are implied by simple graphics, for example, can communicate spurious relationships between measurements that humans integrate into their mental models of the underlying systems (Thacker & Sinatra, 2019). This has not been broadly explored in the process industry.

The human—informatic interface represents the meeting of two information models; one intentionally designed and executed technologically, and the other developed internally by the human user through training, observation, interaction, and association (Morgan, 2002). Because complex, automated systems require human supervision and intervention - particularly during abnormal or emergency events – the ways in which people acquire knowledge from the controls interface and associated informatics is a topic that can not be ignored by process engineers. Simulators are often close copies of operating interfaces and part of formalized training programs, but few consistent practices for evaluation of knowledge and problem-solving skill have been adopted in the process industry (Colombo & Golzio, 2016; Kluge, 2008; Manca et al., 2012, 2013). With the pace of technological advancement, particularly in automation and analytics, there is a need to prioritize implementation and identify the relevant benefits of new ideas and designs (Buddaraju, 2011; Ikuma, Koffskey, et al., 2014; S. T. Lee et al., 2017). Few studies have been performed in the process industry meaningfully tying design features - particularly in informatics and user interfaces - to business and operating objectives (Doe et al., 2005).

1.3 Literature Review

As industrial process facilities become increasingly interdependent, and high performing, the human operators that interact with the process must acquire and refine the cognitive skills necessary to effectively supervise the assets (Venkatasubramanian, 2003). A gap has been established between the traditional skills and routine tasks of process operators, and the modern demands of the highly automated workplace (Golightly et al., 2018; Golightly & Dadashi, 2016; Strobhar, 2014a). Advances in artificial intelligence and machine learning have been pursued with great enthusiasm in industry and research (Dalzochio et al., 2020; Venkatasubramanian, 2019b) and the automation of repeating and routine processes has shifted the role of the process operator toward supervision and away from routine diagnostics and interventions.

Assessing the effectiveness of these advances has not been pursued as enthusiastically, and there has been little recognition in the process industry that informatic tools and control system interfaces have significant influence on how process-contacting workers understand the underlying system. That understanding – the mental models developed through interactions with process information systems – influences decision-making (Naderpour et al., 2015; Ngo & Kunkel, 2020; Whitley et al., 2018). Information models can cause the development of sub-optimal knowledge if they are designed using flawed foundational assumptions (Nakai et al., 2017; Svenson, 1989). Lefsrud, Fox, Cooper, & Taminiau, (2019) discuss how professional engineers and accountants strategically embed ambiguity in how petroleum reserves categories are defined, so that they can maintain interpretive flexibility. Kutsch and Hall (2010) described an expansion of the "taxonomy of ignorance" driven specifically by popular Project Management trends that is relevant to the process industry, classifying the ways in which human understanding diverges from reality. Environmental releases, asset damage, loss of public trust, injury and death in workers and ultimately financial performance can result from ambiguous interpretations of information (Maume-Deschamps et al., 2015). Over time, the presentation of information in a consistent – but misleading or inaccurate - format can completely replace previous education and knowledge (Sandvik, 2008; Thacker & Sinatra, 2019). Evaluating the effectiveness of an engineering artifact is a necessary element of the design process, and evaluating the results of human-computer interactions poses challenges to designers and vendors offering new solutions to industrial consumers.

1.3.1 How is operator decision-making measured?

If a process control interface is intended to augment or support operator decision-making, then the evaluation of designs and design features must effectively measure the change in decision-making performance. Competence assessment, and closely linked operator training in the chemical process industry has not progressed with the same enthusiasm as technological advancement (Mkpat et al., 2018; Ottewell, 2011; Ross, 2013; Strobhar, 2014c). Evaluating the quality of decisions in high-risk events is effectively an assessment of “problem-solving skill,” as defined in the field of Educational Psychology (C. B. Lee et al., 2019). Evaluating a complex, dynamic situation with multiple variables requires effective individual management of psychosocial influences - from perceived stress to individual confidence - and immediate task load, which can be aggravated by the interfaces and data systems intended to facilitate information retrieval and diagnosis (Bruch & Feinberg, 2017; Haji et al., 2016). These influences tend to be discussed as “stress” or “organizational factors” in the process safety industry, with broad recognition that stress or inappropriate organizational pressure negatively impact the quality of decisions (Hendershot, 2011).

Decision-making and human error evaluation are further complicated by the rough separation of two types of decisions, those that are routine or “recurring” within the scope of the role, and the non-recurrent decisions that present a unique challenge. The cognitive processes involved in recurrent versus non-recurrent decisions are distinct (Greiff et al., 2012; C. B. Lee et al., 2019; Seel, 2006) and simple measurements or single “right/wrong” outcomes do not offer substantial insight into how informatics support or hinder those decisions (Dindar et al., 2020; C. B. Lee, 2010). An error can be easily identified and counted when a repeating, recognizable scenario is presented, and the decision-maker applies the correct procedure to address the abnormal condition (Strobhar, 2014e). The utility of that error measurement relies on several factors, the most important of which may be that not all procedural responses are effective, nor are operating procedures always rigorously maintained (Kanse et al., 2018).

Advances in process automation and efforts to proceduralize operations have reduced the human decisions necessary for many routine operations, but “abnormal operation” includes some recurrent scenarios as well as new situations that require creative diagnostics or selection of interventions (Eljack & Kazi, 2016). If modern process operators are expected to supervise increasingly complex automated facilities and intervene largely in events that can’t be automated, the evaluation of their decision-making competence must extend beyond simply counting the number of procedural errors made. Decision-making at the human-process interface – particularly meaningful evaluation of non-recurrent skills - has not been fully explored in the process industry, and that presents a challenge for evaluating new technologies or design features.

1.3.2 How are DCS/UX features designed and evaluated?

Distributed Control System (DCS) panels or screens are becoming a principal source of information about the processes they represent, communicating asset relationships, process measurements and representations of the process automation architecture visually. Simulator training for process operators is becoming increasingly important (Susarev et al., 2017) as the requirement for in-field interactions decreases (D. Cameron et al., 2002; Ottewell, 2016; Patle et al., 2014), and process simulators used in training and workplace assessment typically adopt the operating interface, or a similar interface often offered by the vendor providing the DCS. DCS screens are not simply windows into the process for the panel operator, they are learning tools that build associations and relationships that humans will later use to recall information from their long-term memory (Elsawah et al., 2017). Adhitya et al., (2014) demonstrated a methodology to evaluate the effect of an “Early Warning” system for decision support, which aimed to reduce the time lag taken for operators to process information and develop a diagnosis during abnormal events. They highlighted the need for process safety engineering to focus on the human factors element of decision support tools. Aditya et al. developed a detailed experimental design to quantify the human-factor effect of the tool, and while the tool was able to reduce the time to diagnosis, researchers noted that this did not correlate to diagnosis accuracy. Evaluating the effectiveness of a decision-support tool is often a multi-variate problem, with the weight of accuracy and speed-to-decision context-dependent.

Multi-scale modelling and sensible conceptualisation of independent, dynamic systems is increasingly important in research and industry as developing and applying predictive analytics relies on a thorough understanding of the “information map” (Lapkin et al., 2011). Shanqing, (2012) summarized the results of multiple trials assessing the effectiveness of proactive monitoring display features, which offered a “look forward” view of consequences which supported operator decision-making in context and suggested directions for future work in applying predictive displays. As the pace of research and innovation accelerates, the selection of which tools, and how they are best operationalized becomes more challenging, and continuing the human factors work in decision-support tools becomes more urgent. Pruning large-scale information models and interpreting the results of multi-variate, multi-state statistical computations is not trivial. This theme extends into the design and “de-cluttering” of DCS interface screens for panel operators (Mishra et al., 2020; Mitchell et al., 2005), where the objective is often to intelligently reduce the number of attention-grabbing display elements during abnormal situations. The selection and presentation of display data can augment operator decision-making performance - or hinder it - and without a reliable way to measure baseline performance, it becomes challenging to evaluate whether a design feature or change has effectively improved performance.

Building effective information models is urgently needed to realize the full potential of recent advances in analytics (Lu et al., 2020; Macchi et al., 2018; Miskinis, 2019; Padovano et al., 2018). The evaluation of “effective” requires the adoption of systematic and reliable assessment techniques for the desired outcomes – improved decision-making by people interacting with the process.

1.4 Contributions

Table 1-1 summarizes the gaps identified in literature review, mapped to the research outcomes of the three studies that make up this thesis. As research progressed, the direction shifted away from developing an integrated management system or informatic and toward the evaluation criteria used to test designs and innovations in decision-support interfaces. Demonstrating the effectiveness of any decision support requires a systematic and defensible methodology to evaluate improvement in decision-making capability; and the decisions of interest in process safety engineering are not easily evaluated using simple psychological experiments.

Table 1-1 Summary of research contributions mapped to gaps identified in literature review

Gap identified	**Research outcome**	**Contribution**	**Type**	**Chap.**
Process safety incidents persist despite detailed investigations, response programs and regulation, "worker training" common causal factor.	Prove the efficacy of integrating plant information from discrete sources in improving risk-based decision making.	Expand utility of an existing informatic technique – ontology knowledge bases	Empirical	3,4
(P. T. Bullemer & Nimmo, 1998; Halim & Mannan, 2018; Venkatasubramanian, 2011)				
Process safety incident responses in operating companies do not reflect current research and in some cases are regressive or counterproductive.	Equip workers to make better decisions around risk by reducing barriers to critical information – access and interpretation	Utility of an integrated information access tool	Empirical	2
		Model for risk-based decision making in the process industry	Theoretical	3
(Baker et al., 2020; Baker & Lefsrud, 2019; Marsden, 2018; Silaipillayarputhur, 2018)				
Human factors analysis recognizes multiple, often dependent influences in human-factor failures not reflected in program responses and frameworks.	Improve quality of informatics interactions, including incident reporting/fault detection	Expand the model for human-factor reporting with dimensions of competence and readiness	Theoretical	3
		Reinforce value of scenario-based testing in technology adoption	Empirical	2,3
(Lefsrud, 2019b; Mazmanian et al., 2013; Moura et al., 2017; Nakai et al., 2017)				
The absence of integrated causal models in the process industry limits the effectiveness of data collection and reporting,	Test the efficacy of a hierarchical approach to causal modelling using transfer entropy.	Practical application for esoteric knowledge. Conceptually complex, not disseminated to operating facilities	Empirical	2

limiting effectiveness of data-driven techniques				
(Eizenberg et al., 2006; Hu et al., 2010; Moura et al., 2017; Suresh et al., 2019)				
Organizations with high degrees of risk maturity struggle with engagement and effectiveness in risk management.	Identify practical gaps in risk management in the process industry and demonstrate the value of ontological structures in risk assessment	Demonstrate the efficacy and utility of existing informatic structures in risk management	Theoretical	4
(Blunden et al., 2019; Chu et al., 2019; Li et al., 2010)				
Information structures and HMI designs can lead to inaccurate mental models resulting in lower quality or inconsistent decision-making.	Demonstrate the efficacy and benefit of aligning assessment techniques with design objectives	Demonstrate an industry-applicable method to evaluate risk-based decision making	Methodological	3
(Agar, 2014b; Basoglu et al., 2007; I. Cameron et al., 2017; Marshall et al., 2018; Revell & Stanton, 2016; Sandvik, 2008)				

The gaps identified and research objectives developed into three distinct research questions:

1. How can predictive analytics be integrated into interface design to augment operator decision making?

This question was explored in a pilot study and led to detailed exploration operator decision-making and design evaluation techniques for user interfaces. The review of decision-making assessment, which incorporated elements of human factors analysis as well as operator training and assessment led to educational and psychological assessment techniques targeting problem-solving. A detailed, full-scale study was developed in response to the question:

2. How can the situational design model be applied to assessing operator decision making?

In a parallel examination of literature and ongoing work in integrated management systems and industrial software design, the concept of ontology knowledge bases and their utility in structuring disparate sources of information was explored. Asset management ontologies were of particular interest, aligning well with existing hazard identification techniques, like HAZOP and Process Hazard Analysis, generally conducted by unit, block and subsystem. A case study was conducted to answer the question:

3. How can asset management ontology knowledge bases be applied to improve understanding of process systems in a risk management context?

The answers to these questions – the conclusions to the three studies described in this dissertation - offer theoretical, methodological, and empirical contributions toward the broader questions posed in Section 1.1.

1.5 Thesis Outline

To improve industrial performance in loss prevention, the human-technical interface and the role informatics play in risk-based decision making must be understood and more deliberately managed.

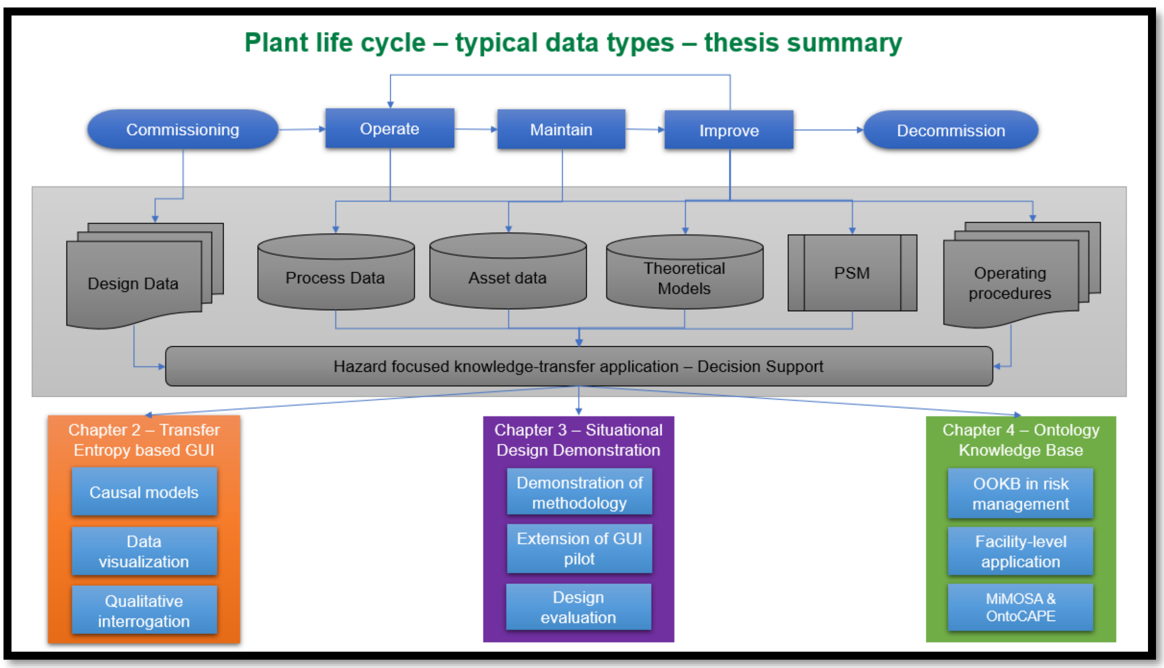


Figure 1-3 typical data types relative to plant life cycle – persistent obstacles toward integrated management systems

This thesis presents three studies exploring that interface and the role of informatics in decision-making, followed by conclusions and recommendations for future research. Figure 1-3 summarizes the relationships between the life cycle of an operating facility and the variety of data storage systems with which process-contacting workers interact in order to make risk-based decisions. Innovations in automation and process safety management seek to integrate these discrete systems and equip human operators with sufficient relevant information to make high quality decisions. The three manuscripts that form Chapters 2, 3 and 4 develop concepts critical to the design of effective decision supports for risk-based decision making.

1.5.1 Integrating predictive analytics into interface designs

Chapter 2, titled "*Decision support for process operators: Task loading in the days of big data"* details a pilot experiment testing the effectiveness of a visual display representing transfer entropy between variables on a typical DCS display. The display was built to represent the Tennessee-Eastman control problem and presented study participants with a dynamic process simulation of a heavily automated industrial process. The study employed qualitative research techniques eliciting detailed responses from participants as they attempted to diagnose and respond appropriately to a simulated abnormal situation. The NASA Task Load Index was applied to establish relative cognitive load during "normal" and "abnormal" operating conditions. This study generated further, more detailed research questions and demonstrated some of the challenges associated with evaluating user understanding of complex systems during typical performance testing scenarios.

1.5.2 Systematic assessment of operator decision-making

Chapter 3, titled *"Smooth operator: aligning performance assessment methods with design and operating objectives"* describes a full-scale expansion of the pilot study detailed in Chapter 2. Research into Applied Psychology and Educational Psychology suggested the "Situational Design Model" (Zhong & Xu, 2019) could be adapted for use in industrial applications, building on prior work from within the process industry to develop common competencies and necessary theoretical knowledge. The study documented in Chapter 3 presented a similar abnormal situation to participants and employed test/control groups to examine the effect of the transfer entropy-derived causal maps on decision-making performance. 32 participants completed the two-phase study, and the response to common assessment instruments like surveys and knowledge exams provided insight for the design of future studies. Appendix 1 contains the research ethics board approval and study summary, and Appendix 2 contains the learning readiness survey employed in the study as well as the marking instructions which include the marking rubrics, theoretical basis for the rubric, and the instructions that were presented to study participants. The situational design model holds promise as an industrial training and design evaluation tool that would more closely align the objectives of business and operations with metrics for direct comparison between designs.

1.5.3 Ontology knowledge bases and risk-based decision making

Chapter 4, titled "*Application of engineering thinking for risk assessment in a Canadian elementary school"* details a facility-level study undertaken as part of a 2-year risk management exercise in an elementary school. The facility owners and operators elected to seek risk management support grounded in industrial process safety, and Chapter 4 summarizes the ways in which ontology knowledge bases – particularly asset management knowledge bases - can support decision making during abnormal operations. The ontological breakdown used to communicate the building mechanical systems and simplify the movement and activity levels of students can be applied to any occupied facility used for any purpose and demonstrates the benefit of direct measurement over the challenges and expense of traditional ventilation models and projections.

Chapter 5 summarizes the conclusions and limitations from all three studies and offers recommendations for further study. An alphabetical list of references follows the concluding chapter.

Chapter 2

Chapter 2 of this thesis has been published as Michelle Naef, Karan Chadha, Lianne Lefsrud, "*Decision support for process operators: Task loading in the days of big data*", Journal of Loss Prevention in the Process Industries.

Contributions of the authors as follows:

Michelle Naef: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – Original Draft, Writing – Review & Editing

Karan Chadha: Investigation, Formal analysis, Writing – Original Draft

Lianne Lefsrud: Supervision, Funding Acquisition, Writing – Review & Editing

Executive summary

Modern chemical processes rely on distributed control systems to make the repetitive and routine adjustments to maintain steady operation. Operators are still required to "supervise the (system) supervisor" and intervene when variables exceed pre-programmed parameters to avert major incidents. Research in human-computer interaction and advanced process control has often focused on data-driven methods for fault detection as distinct from operator effectiveness. In this paper, we explore the application of a novel data-driven fault-detection technique to enhance operator decision support. During a simulated abnormal event, three users attempted to diagnose the root cause of a process upset using a traditional or standard interface, then with the addition of causal maps, in a A-B-A single-subject design. The causal maps were derived using a hierarchical method that could be applied to a wide range of chemical processes as an online, adaptive augmentation for abnormal situation management. Using a think-aloud technique, the three participants developed high quality insights into the process without negatively impacting the overall task load. These preliminary findings challenge prevailing wisdom in process control interface design, which often focuses on de-cluttering displays at the cost of information resolution.

2.1 Introduction

Modern chemical processes rely on automation in the form of Distributed Control Systems (DCS), where repetitive adjustments to the expected variations in the process are directed by Process Logic Controllers (PLC) which remotely control valves and powered equipment like motor drives. In industrial processing facilities, the operating panel is a Graphical User Interface (GUI) that receives the transmitted signals of instruments located on the equipment. Some variations in system parameters are outside the ranges programmed into the PLCs, which cause an alarm notification to display on the GUI. Alarms require operator attention. From the readings displayed on the GUI, the operator must diagnose the cause and determine the appropriate intervention(s) to restore the process to a steady state. The diagnosis and selection of intervention is a high-level cognitive task, based on multiple inputs with a series of possible outcomes for each possible intervention. Operators must evaluate not only the readings on the panel, but the exogenous influences like weather and business priorities as well as predicting the most likely outcomes of each possible intervention. Fault diagnosis requires the competent execution of multiple subtasks, and the application and synthesis of knowledge in thermodynamics, kinematics, chemistry, and process control.

In this paper, we explore the effectiveness of using causal maps for improving operator decision-making. The causal maps were generated according to the hierarchical method proposed by Suresh et al., (2019) communicated as a series of directed graphs to demonstrate strongly- and weakly-connected process variables. These digraphs were simplified and overlaid on typical DCS screens developed around the 2015 revision of the Tennessee Eastman Problem (Bathelt et al., 2015). Participants were asked to think-aloud as they acquired understanding of the simulated plant's operation and reasoned the likely root-cause of a fault as it propagated through the plant. An adaptation was introduced in a second section of the test, and in the final section, users were asked to return to the original interface and perform another fault diagnosis. Participants engaged with the adaptation and increased the number of correct inferences made in their narratives during and after accessing the map overlays. These early findings encourage further development of this tool to improve operators' understanding of complex causality in industrial process settings.

2.2 Literature review and context

Bainbridge, (1983) is credited with the development of the term "Automation Paradox" which predicted that as process automation became more complex, the cognitive requirements for operators would increase. Operators would have fewer routine tasks during normal operation and limiting process interactions to transient or abnormal operations. These operating states would require more challenging analysis and decision-making to restore process stability, while carrying more significant consequences for suboptimal or delayed interventions. Operators intervene or interact with the system less often, and the necessary interactions during abnormal situations require significant knowledge of both normal and abnormal operating states (I. Cameron et al., 2017; Venkatasubramanian, 2003). Further, scrutability of the controls architecture tends to be poor, and distributed control systems are largely designed and maintained by specialized vendor firms. The project described in this paper sought to design and test a causal map overlay applied to a typical DCS screen as users attempted to diagnose the root cause of abnormal operating states.

2.2.1 Automation and human reliability

Shu et al., (2016) highlighted the weak adoption of automated fault detection and diagnosis methodologies in the "big data era" despite decades of innovation and research that offered many new ways to diagnose faults. Studies into human reliability in the process industry tend to focus on operator error in specific tasks, and do not tend to explore learning or understanding developed through interaction with the interface (Cheng et al., 2019; Lithoxoidou et al., 2018; Lucke et al., 2018; Zhang & Dong, 2014). This type of evaluation often fails to reflect the complexity of the cognitive tasks examined, that suboptimal interventions are the result of multiple individual decisions.

Kluge et al., (2014) described the cognitive requirements associated with seven aspects of occupational competence among process operators. As indicated by their application of the term "non-transparency" (Kluge et al., 2014) recognize that process control interfaces necessarily abstract the complexity of the interconnected systems and process operators must fill in the detailed concepts in order to make optimal decisions. The current state must be inferred from the abstract representation monitored by the panel operator. Abnormal situations challenge the "coupling and interaction" (Kluge et al., 2014) relationships that are developed when processes are observed in steady state operation. The dynamic characteristics of the abnormal situation mean that many of the simple, cause-and-effect couplings observed when variables are within normal ranges no longer apply. The ability to understand and predict dynamic effects becomes central to good decision-making once the steady state is perturbed by a fault or series of faults (Kluge et al., 2014; Yang et al., 2014).

Accurately perceiving and responding to mitigate risks requires synthesis of multiple goals, knowledge of possible dynamic effects, understanding of couplings and interactions as well as awareness of the current state. (Dai et al., 2016) summarized a framework for smart chemical process operation identifying "safety risk intelligence" as an essential feature of effective control system design. The concepts of integrated and hybrid methods for system supervision have been researched in multiple engineering disciplines with overlapping but not identical objectives (Lei et al., 2013; Rebello et al., 2019). Opportunities for synergy have not been widely seized. (Doe et al., 2005) remains unique in the field of abnormal situation management, where direct improvements to operational outcomes in the short and long-term were identified and directly attributed to interface designs that improved operator understanding of system behavior. The automation paradox suggests that simply reducing the number of routine operator interactions may be creating a worsening skills gap in process operators, supported by industry reporting around process safety incidents. There is an opportunity in the field to develop more in-depth, quantifiable assessment methodologies for evaluating operator effectiveness with respect to DCS interfaces, which would contribute to the design evaluations and offer business justifications for continued investment in operator decision support systems.

2.2.2 Causal maps

Lei et al., (2013) directly addressed the impact of rapid technological development and lagging efforts toward increasing operator effectiveness in their introduction of a novel method for data-driven causality analysis. Hybrid methods typically combine data driven and model-based automation techniques. Expert systems are applied to reduce the dimensionality of causal networks (often large and sparse) by eliminating weakly connected variables and addressing indirect causality. Weakly connected components are often eliminated using threshold values and statistical testing if changes in one variable cannot be directly associated with statistically significant variation in the values of the other.

Lithoxoidou et al. proved the utility of transfer entropy in generating a Signed Digraph (SDG) model of the Tennessee-Eastman Control problem (TEP), widely used as a benchmark test for automation and control system optimization research (Lithoxoidou et al., 2018). The same control problem was used to further refine fault diagnosis algorithms (Duan et al., 2013; Gharahbagheri et al., 2017; Landman & Jämsä-Jounela, 2016; Peng et al., 2015) culminating in the hybrid method presented by (Suresh et al., 2019). The hybrid method generates a "causal map" of the TEP at two levels, in the form of signed digraphs at the unit-wide and sub-system levels. Decoding the SDG requires advanced knowledge in theoretical mathematics, process control and chemical engineering. While the nodes are relatively easily defined as "measurements" and "manipulated objects", the edges represent the abstract statistical relationship of "significance" determined using the concept of transfer entropy.

2.2.3 Application and design

Our central hypothesis is that a causal map can improve user understanding of relationships leading to a more accurate mental model for causality in complex industrial systems than common interface design practices currently foster. Developing a method to effectively communicate connectivity and interactions, dynamic effects and increase transparency could offer an immediately useful solution to the decades-old problem of "human reliability" in industrial process operation.

To test our hypothesis, a typical DCS screen was constructed for the Tennessee Eastman Control Problem, and a visual overlay designed to represent the transfer entropy and significance values between variables that could be toggled on and off by participants. Few studies in the field of process automation seek to evaluate user understanding of the causality between variables in the process, and cognitive load is frequently assessed only in terms of "stress" or "distraction" leading to mis-operation (Kluge et al., 2014). This pilot study was designed to specifically evaluate participant understanding, and assessed different dimensions of task loading using the NASA TLX scales. TLX is an abbreviation of "task load index" (Hart & Staveland, 1988).

If data-driven methods can be feasibly incorporated into existing DCS screen designs and improve user understanding, there is potential to augment operator performance and effectively access organizational knowledge that is otherwise difficult to integrate. This study sought to evaluate an application of data-driven causality analysis in improving typical distributed control system - if a visual interpretation of the causal maps would lead operators to improved understanding of the complex systems represented on the screen. In typical DCS screen design, control loops are omitted from the display or heavily abstracted, for example, which requires operators to mentally keep track of a key causal relationship. If causality can be communicated explicitly, the redundancy of many alarms becomes apparent, which could open new possibilities for prioritizing and sorting alarms. It may be possible to augment operator performance by including different information on screen, as opposed to attempting to reduce visual clutter and finding ways to prune the number of redundant alarms that ring in.

2.3 Methods

The layered approach to user interface (UI) evaluation detailed in (Paramythis et al., 2010) was used to deconstruct this design project and identify suitable evaluation criteria for the UI design and observational methods for data collection. This UI was evaluated at the applying adaptation layer (AA), with both qualitative and quantitative instruments measuring user perceptions of timeliness, obtrusiveness, and controllability. In this study, timeliness evaluated whether the adaptation appears when the user requires prompting or support, obtrusiveness was defined as "the degree to which the adaptation interferes with task performance" and controllability was defined as "the degree to which users can choose to interact with the adaptation."

This study combined a single-subject quasi-experimental model (Chiang et al., 2015) with the think aloud technique (Klarborg et al., 2012) and applied the NASA TLX (Hart, 2006; Hart &

Staveland, 1988) to evaluate the perceived task load increase, if any, between the typical interface layout and the GUI incorporating the adaptation. The single subject model is an experimental technique used to observe a causal relationship between an adaptation and the user's performance of a particular behavior or task (Chiang et al., 2015). The task described to participants was a root-cause identification of the source of a process disruption observed propagating through the simulator. Task performance was evaluated by counting the number of correct inferences vocalized by participants as they "thought aloud" throughout the root cause identification (Chang & Johnson, 2020).

The NASA TLX scales were developed to evaluate user perceptions of task load and was selected for this study as the most widely-published evaluation of cognitive and physical load with respect to industrial automation technology (Hart, 2006). User perceptions of cognitive load were used to develop a more detailed qualitative evaluation of the overall effectiveness of the causal map adaptation and contextualize user responses around the timeliness, obtrusiveness, and controllability of the interface. The NASA TLX scales use a semi-quantitative measurement like a Likert scale, where each extreme represents a "maximum" or "minimum" interpretation of a different type of load, as evaluated by the user. With sufficient observations and sample size, the scales can be applied in quantitative analysis with 21 discrete steps between each extreme. With small sample sizes, the utility of the scales is predominantly qualitative.

2.3.1 Think-aloud and qualitative analysis

Participants in the study were prompted to think aloud as they interacted with the GUI and observed the process simulation. As they developed understanding of what was shown on the screen, they vocalized inferences, speculations and predictions about the relationships between variables, and causal links in simple and complex combinations of system elements. Participant narratives were analyzed and synthesized using an application of Grounded Theory (Glaser, 2016; Sherer, 2019) identifying each section of the test as a case.

2.3.2 Causal map development

Suresh et al., (2019) calibrated the diagnostic accuracy of their hierarchical approach using the Ricker (Bathelt et al., 2015) test bed, which was used in the same form (base case, unmodified) for this study. Suresh et al. used an edge-weighting factor of w = 0.14 as a lower bound and demonstrated validity in all the disruption modes for the base case. The algorithm identified the root causes effectively in all 21 disruption modes (Suresh et al., 2019). The hierarchical method detailed in was developed using the directories referenced in (Bathelt et al., 2015), in the base case (Mode 1) configuration. That configuration was replicated in this study, which was calibrated to the disturbance causes originally defined by the creators of the Tennessee Eastman Problem (Downs & Vogel, 1993).

Suresh et al., (2019) used a sequential numbering to identify each variable as a node, and calculated transfer entropy between variables. For this experiment, we used colored ellipses to highlight the nodes (variables) under consideration in each causal map, and arrows representing transfer entropy in the same color linking the ellipses. Table A-2 identifies the simplification made for the user test, variables measuring composition were omitted from the base map as participants did not have access to the time-delayed molar composition measurements during the first two tests. We used curved arrows to avoid any visual confusion between the causal map overlay and the line diagram representing the process. The subsystem graphs mapped easily to the GUI and were drawn using a different color to distinguish them from the unit-level maps. As the magnitude of transfer entropy between variable pairs reflects the degree of influence, the highest-value (>0.8) connections used a heavy line weight, and the lowest value connections

(<0.1) were shown with a thin, dotted line. Connections with transfer entropy values between 0.1 and 0.8 were shown with a standard line weigh The MATLAB App Designer GUI was rebuilt in Simulink for the third pilot test, which allowed the incorporation of tunable parameters and considerably increased the simulator speed.

2.3.3 Equipment

Participants were required to have reliable internet connectivity and access to the Zoom conference application. Researchers operated the simulation and used screen sharing to display the interface to participants, who controlled the cursor. The user interface uses visual cues only, and no additional hardware or software was necessary to complete the study. Pilot group participants were asked to participate in a ten-minute, semi-structured interview following the study which took place through video conference. The entire conference call with each participant was recorded and manually transcribed.

2.3.4 Procedure

Participants opened the GUI displaying the line-diagram representation of the plant. Participants were prepared through the initial invitation to expect some process disruption and reminded by the researchers during the preamble that they would be asked to identify the cause of the disruption after a fault had propagated through the plant. All participants had the option of choosing to observe "normal operation", which launched the simulation without a pre-programmed disruption. The GUI home screen was developed using the principles and process detailed on a plant of similar complexity in (Normanyo et al., 2014). Figure 2-1 and Figure 2-2 show screen captures of the interfaces used for Participants 1 and 2, both the A (baseline) and B (intervention) test interfaces.

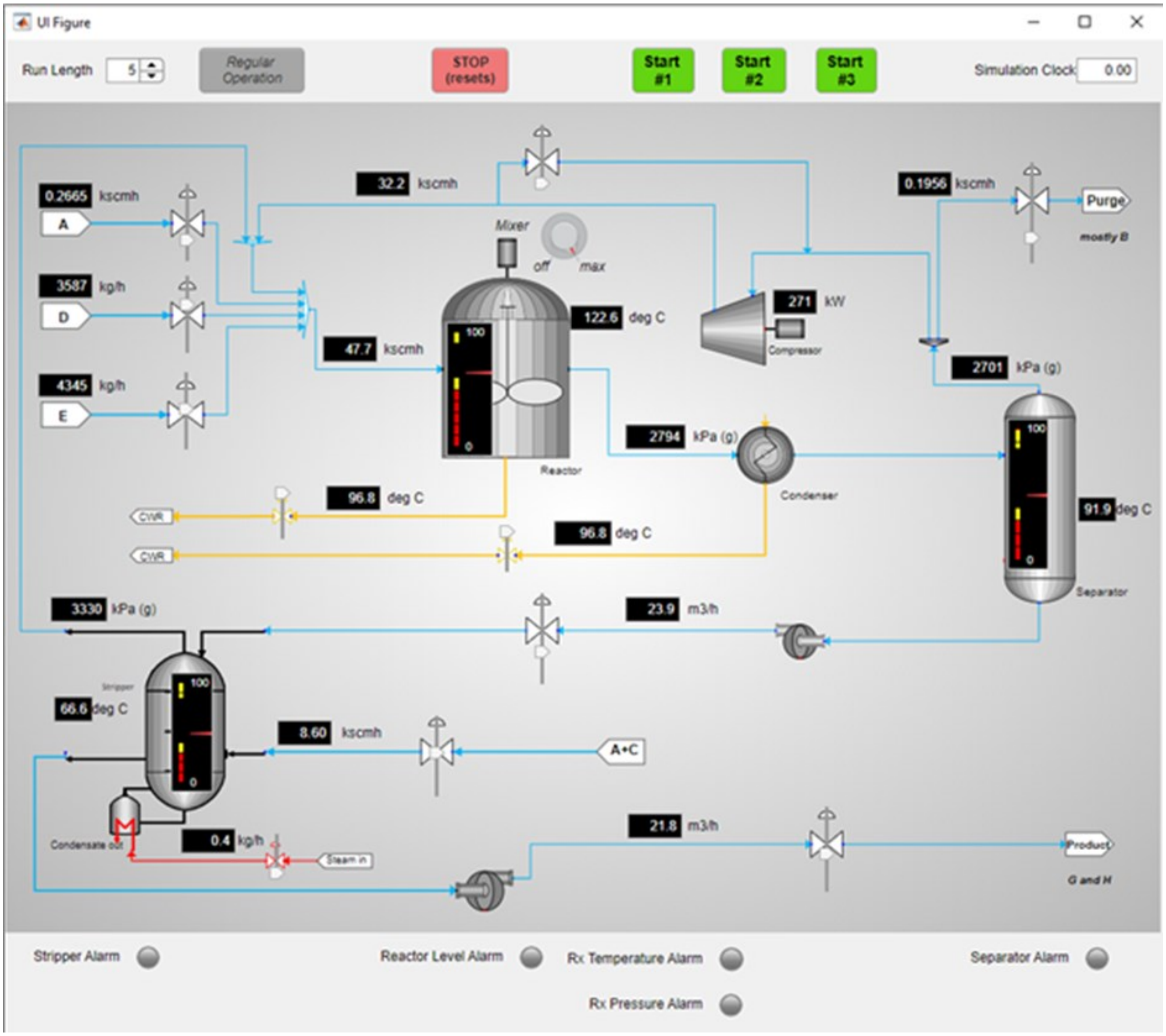


Figure 2-1 Interface used for the A tests, Participants 1 and 2
The graphics and symbols used in the GUI are used under Academic license.

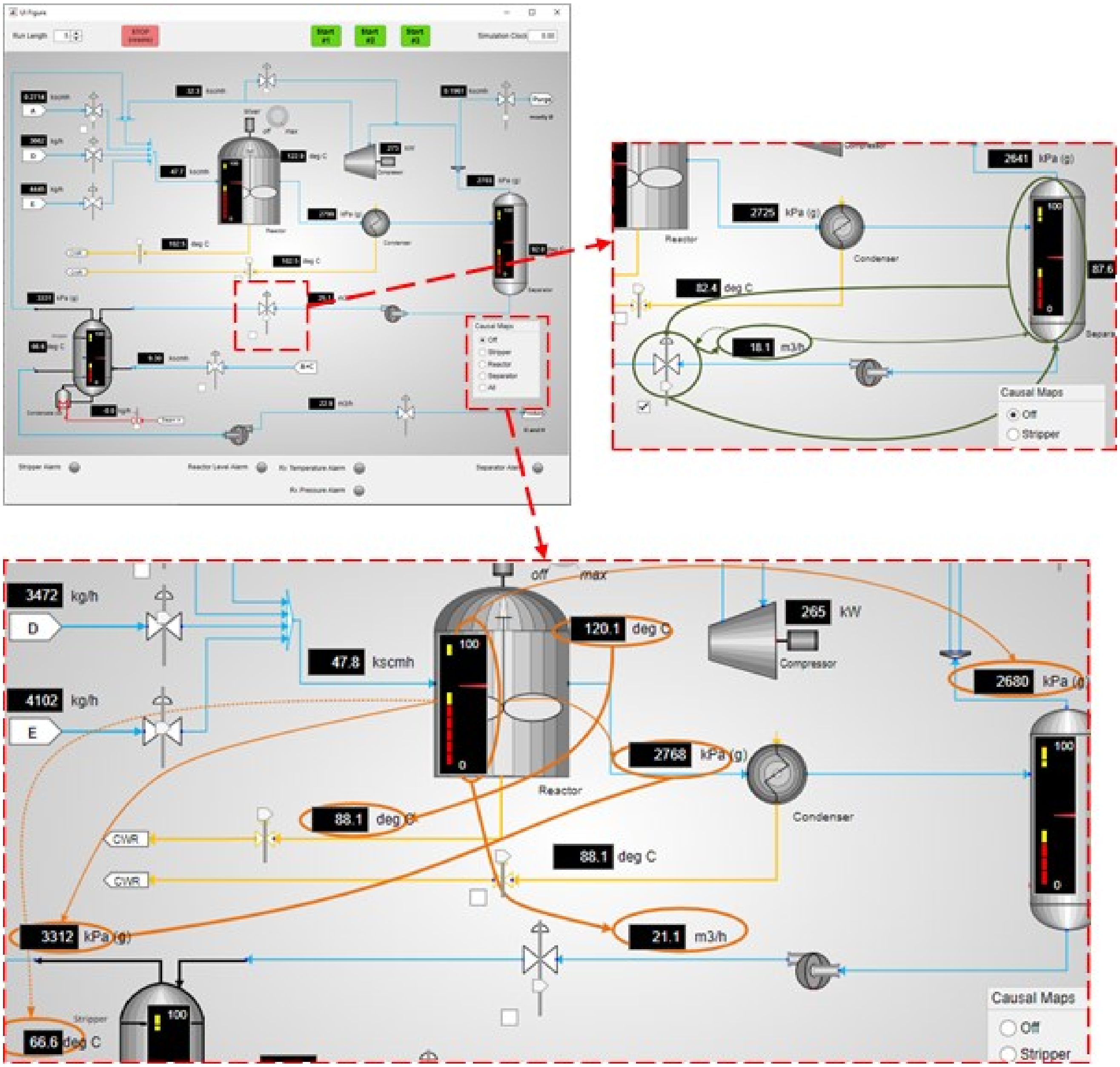


Figure 2-2: Interface used for the B tests, Participants 1 and 2

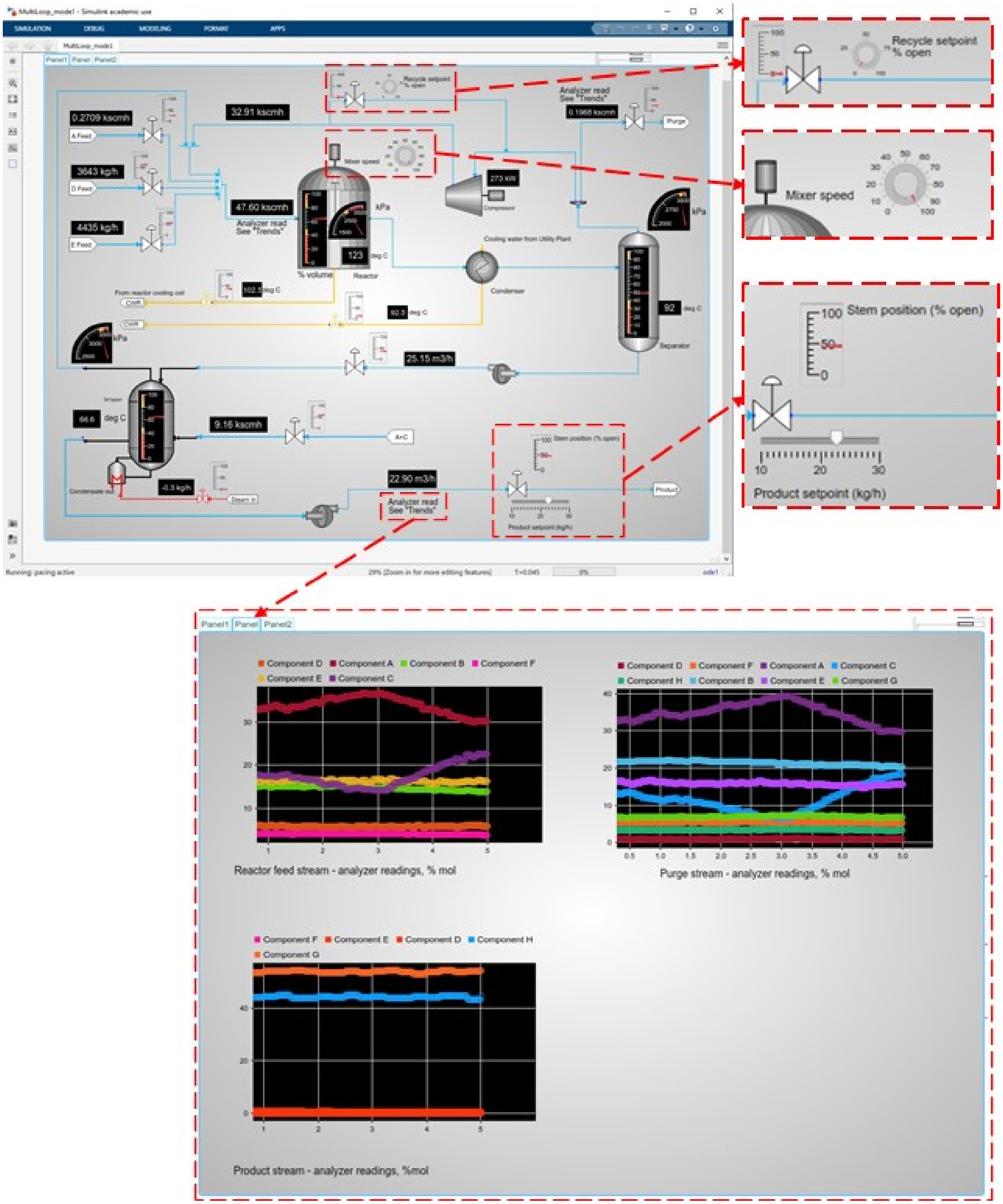


Figure 2-3: Modified interface used by Participant 3, tunable parameters circled and second tab showing trends for designated streams

Participants were prompted by the researcher to think aloud and encouraged to narrate their mental process for identifying the root cause of the fault and select an intervention to apply. Questions regarding the chemical process were answered, as were questions regarding the meaning of some symbols.

The B test followed the A-test format, and the adaptation was introduced. Toggles selected by participants caused a graphic overlay to appear and the causal map associated with the equipment or subsystem were overlaid on the screen. Users were prompted to reflect on the likely root cause and once they committed to a decision, the simulation was stopped. Following the "B" test, the "A" test was repeated. After completing the simulations, and a third set of NASA TLX questions, pilot participants were invited to provide reflections on the experiment overall. Each pilot test was recorded by the remote conference software and manually transcribed by researchers.

2.3.5 Participant selection

The GUI designed for this experiment relied on domain-specific visual representations of chemical process industry equipment. The symbols and ontology combine concepts of mass/energy conservation, thermodynamic and fluid dynamic principles using basic symbols and two-dimensional line diagrams. The target participants for the pilot study had verifiable familiarity with the symbology and ontology used in the development of the GUI. Three participants were contacted through personal networks to complete the pilot. Table 2-1 summarizes the participants' inclusion criteria and contributions to the data set.

Table 2-1: Qualitative data collected and participant profiles

Participant	**1**	**2**	**3**
Knowledge base	Mechanical engineering	Mechanical engineering	Materials engineering

Related experience	15 years	3 years	5 years
Industry	Chemical processing	Aerospace maintenance	Safety consulting
Root causes identified in each test run	3, 3, 3	1, 1, 1	1, 1, 1
Time in test	2:33	1:02	1:36

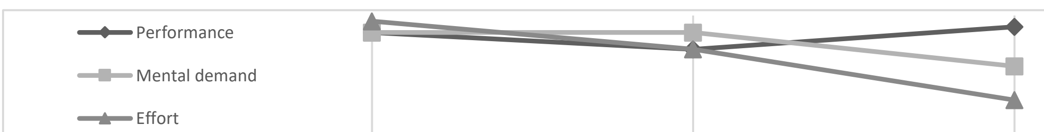
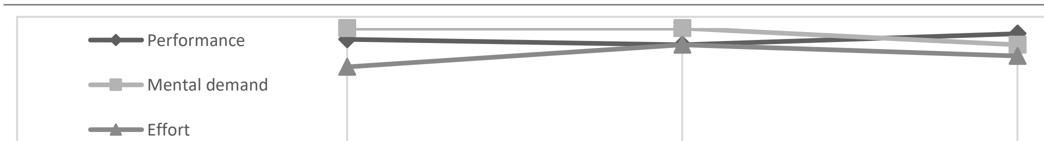
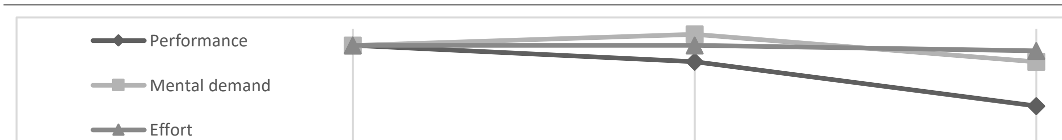
2.4 Result

Participants recorded their assessments of different types of task load using the NASA TLX scales and volunteered narrative self-assessments of their own performance. Participants all expressed disappointment that they had not determined the correct root cause, despite having no evidence that they had diagnosed it incorrectly. Participants all indicated that the physical demands of the test were not significant.

Table 2-2 summarizes the results of each pilot test; the diagnosis of the root causes and the number of correct inferences vocalized regarding the system's behavior. Table 2-2 shows the participants' evaluation of the mental demand perceived in each section, measured using the NASA TLX scales, used here to test their effectiveness with pilot participants. Numerical values are omitted from the axes to reflect the strictly qualitative nature of the observations. Included are the average number of interactions made with the causal maps in each section and the map levels most frequently accessed, a proxy measurement for controllability of the interface (distinct from controllability of the DCS). Participants 1 and 2 experienced frustrations with the speed of the test, and the GUI was replicated as closely as possible in Simulink prior to the third test.

Convergence between the first two participants in terms of performance and think-aloud themes was sufficient to develop preliminary results and the third participant performed the test using a modified GUI generating additional information and suggesting directions for interface improvement before a larger test pool is approached. Setpoints for the product valve, mixer and recycle valve could be altered in real-time by the user, which caused true-to-reality responses in the measured variables. Participant 3 recorded the most significant performance improvements suggesting that the ability to test and receive feedback from the interface about variable relationships is an important learning activity.

Table 2-2: Performance assessments, by participant, over three test sections. Charted responses to NASA TLX questions for Performance, Mental Demand and Effort

P1	Increased correct inferences, at levels of simple and complex causality. Average of 4 interactions with causal maps, subsystem level.
	
P2	Increased correct inferences at levels of steady state, simple and complex causality. Average of 7 interactions with causal maps, both levels
	
P3	Increased correct inferences, at levels of steady state, simple and complex causality. Average of 10 interactions with causal maps, both levels
	

The initial experimental design assumed that the “trouble” states were apparent through measurement/readings alone, but as more users interacted with the interface, it became clear that this was not the case. As one pilot participant described, “all I can tell is that the plant is compensating for something. I’m just waiting for the whole house of cards to fall down now.” Figure 2-4 shows the coding progression applied to qualitative data, and the classification of participant comments according to the degree of complexity of the relationship between variables.

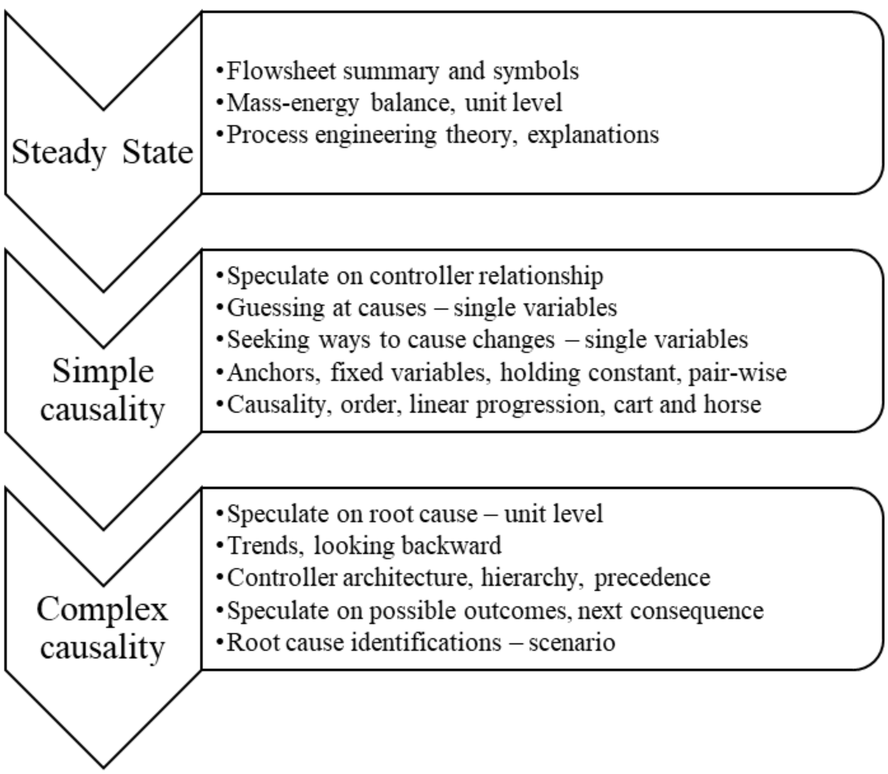


Figure 2-4 Primary codes and secondary codes (chevrons) applied to think-aloud narratives

Participants all looked for interactive options, like the valve causal maps. This reinforces the type of information that is sought by users, while not directly informing how it should be communicated.

During the development of the interface, preliminary testers evaluated the interface and experimental design informally to verify operation and basic parameters before pilot participants were involved. The preliminary testers did not have training or work experience in the subject domain but engaged readily with the test format and felt that they could have performed better if causal maps were associated with valves. These were interesting comments, particularly in the context of task loading. We originally assumed that participants would be overwhelmed by the number of subsystem maps. Given the relative lack of domain knowledge among preliminary testers, we did not expect them to converge with the study participants on this concept. Both preliminary testers and study participants most frequently engaged with the less-complicated controller maps to check basic assumptions. This activity appears to influence both knowledge acquisition and perception of load, which may not have been distinct without the application of two different evaluation techniques, the self-assessed task load index and the think-aloud problem solving. The codes that developed as transcripts were reviewed began to reflect the language of teaching and learning.

All three pilot participants expressed frustration with the implied meaning of the causal maps, which we did not explain until the end of the test. The relationships and directions did not align with participant understanding of either physical relationships or likely controller relationships. Participants 1 and 2 had few tools with which to experiment and test assumptions, leading to frustration and lower self-assessments regarding performance.

Table 2-3 summarizes the measurements of Unobtrusiveness, defined in this experiment as "the degree to which the causal maps interfered with root cause identification." Frustration and perceived temporal demand (rushing, pressure) as well as perceived mental demand were used as a proxy measurement for obtrusiveness. The number of incorrect inferences were counted for each participant and are used as a proxy measurement of obtrusiveness. If the objective of the interface is to improve the participants' understanding of the system, the number of correct inferences is relevant, but so is the number of incorrect inferences. The number of incorrect inferences made by all participants decreased in each section, and the number of correct inferences increased both in frequency and complexity. Participants indicated relatively high obtrusiveness in this context, but the reduction in incorrect inferences as they accessed the interface suggests that despite participant perceptions, the causal maps were effective in improving understanding.

Table 2-3: Unobtrusiveness, as measured by the number of incorrect inferences vocalized by participants in each section with participant scores for frustration, temporal demand, and mental demand

P1	Decreased number of incorrect inferences

P2	Decreased number of incorrect inferences

P3	Decreased number of incorrect inferences

All participants tested assumptions made at the simple causality level, with the separator and reactor temperature controller maps on the more complicated base maps for the reactor, stripper, and separator. The simulated problem was completely new to all three participants, as opposed to typical simulation tests where experienced operators evaluate interfaces for a chemical process they operate daily. Participants 1 and 2 spent similar amounts of time familiarizing themselves with the screen and mentally recording the ranges of readings as the simulator ran in "normal" mode before they initiated the first test. Participant 3 used the faster simulation speed to repeatedly cause the unit to fail by manipulating the mixer speed and recycle valve, effectively acquiring concrete knowledge on the simulator's internal "rules" and the relationships between at least two of the manipulated variables.

It appears that busy screens and multiple, unprioritized number displays lead to increased frustration and effort, but a busy display highlighting priority readings and giving indications of relative importance decreases the overall task load, particularly in terms of perceived mental and temporal demand. This seems intuitive, but this distinction is not made in many studies of operator load and cognitive state during abnormal situation management. The focus of many studies is simply to reduce the number of enunciated alarms and decrease the length of alarm lists (Wang et al., 2015, 2018; Zhu et al., 2014) as opposed to establishing what types of information are necessary to address the underlying cause of the disruption.

2.5 Discussion

The results of the pilot align with the framework developed by (Kluge et al., 2014) in linking cognition and learning environments for operators. They noted one of the cognitive requirements under "couplings and interconnections" as "simultaneously process the interplay of cross-coupled variables." Pilot participants talked through their thinking in describing the causal maps and moved between concepts of physically connected equipment (piping, vessels, and valves) and instruments connected to logic controllers and valve actuators. Participant 3, who was able to manipulate three of the twelve manipulated variables in the simulation spent the most time interacting with the normal mode and achieved the best performance results making more correct assessments of the root cause and making more correct inferences in each phase of the test. Some existing research links element interactivity and cognitive load (Hanham et al., 2017; Wouters et al., 2008) and further research into these subject areas will be helpful in continuing to develop both the interface and the experimental design.

The frustration and sense of "giving up" that both Participant 1 and 2 discussed in the third section suggests that without the ability to validate or verify their simple inferences regarding the system, there was an impediment to progressing to the complex causal level (Haji et al., 2016). Following the application of secondary codes to the narrative think-aloud transcripts, literature review focusing on barriers or impediments to learning linked cognitive load theory to the results of this test. Participants 1 and 2 had the widest variation in terms of hands-on experience in chemical processing facilities. The convergence of themes between both is an indicator that the observations in this area may be more broadly generalizable.

Nguyen et al., (2020) discussed learning paths in the context of online education resources. They synthesized the concepts of sequencing multiple learning objectives in proposing a framework for more effective student-led online instruction. The sequencing of learning objectives relied on both Bloom's taxonomies and Biggs' principles of constructive alignment. Of relevance, they noted the necessity for different types of assessment for students to continue progressing to higher levels of understanding.

The task set for participants in our experiment closely paralleled a student learning a complex subject independently. While not directly applicable to the design of the casual map adaptation, Nguyen et al., (2020) offered new insights for an improved experimental design that could more directly measure the effectiveness of the causal maps themselves, by structuring the simulation more intentionally as a learning activity.

The causal maps associated with valves in the system are simpler than the unit-level maps and appear to serve as a more significant bridge to understanding than was expected. (Kluge et al., 2014) discussed this idea in describing the learning objectives in three aspects contributing to complexity: couplings and interconnections, dynamic effects, and non-transparency. They identified acquisition in chunks and the acquisition of feedback about a decision as priority objectives in operator training programs. They offer further context to early results here in identifying "acquisition of situational cues" as objectives. Participants expressed relief when discovering the causal maps associated with valves, then realized that the relationships described were not control loops. Frustrated reactions followed, as the causality between variables both manipulated (valves) and responding (measurements) is not always intuitive. It was unclear to participants what the arrows and circles used in the adaptation were intended to mean, and the "Frustration" metric on the NASA TLX was higher for all users during the B test.

2.5.1 Procedure

The pilot study indicated that the relationships between display complexity and task load may not be as significant as the state of cognitive load that can be induced, when necessary information is not available, or confirmation of low-level inferences is not received. This is a potentially significant finding in the design of DCS interfaces in the chemical processing industry and supports the expansion of this experiment to a larger set of participants, applying a more carefully constructed evaluation structure, including more precise evaluations of cognitive load.

The pilot participants contributed valuable insight to the experimental design, demonstrating that in order to evaluate the effect of a novel adaptation fully, participants need to be offered a fair opportunity to learn how it works. Simulators with tunable parameters that accurately reflect real plant behavior can be an inexpensive and powerful tool, offering the opportunity to independently acquire knowledge on the relationships between variables at different levels of complexity. Improving the behavior and controllability of the causal maps, with more dynamic (indicating directions, moving signals) and modular (able to select one variable and path at a time, display others and move between) would provide further insight to this area and allow a better-quality study with a larger group. To progress to problem solving and the creation of new insights, participants demonstrated the necessity of feedback, both in the type of interactions with the causal map adaptations as well as in the design of the test. The experimental design for the expanded experiment will incorporate these findings and continue the development of an assessment methodology for user interface designs in the process industry.

2.6 Limitations

The application of the single-subject methodology in this initial test was not sufficient to extrapolate results. The experimental design categorized “root cause identification” as the task of interest, but that is a compound set of tasks that does not necessarily involve repetition of similar cognitive operations. To fully apply the single subject model, the behavior being evaluated must be repeated at a rate sufficient to observe stabilization in the user’s performance of the behavior or task. In this study, the task we sought to evaluate was a compound learning activity, and as a result did not meet the full standard for the single-subject model. As data were analyzed, it became apparent that the compound task of interest was the construction of a mental model of the process, which could be measured indirectly by the number and type of unprompted observations about the system as well as responses to questions from the researchers. Additional experimental design is necessary to develop a more rigorous and repeatable evaluation methodology, which will be the subject of continuing research.

This study was conceived to evaluate the potential for a timely application of analytic methods to existing technology; the interface developed for the process simulator reflects a moderate degree of automation and control for process plants currently in operation. Time series data for temperature, pressure in principal vessels are available, and time-delayed composition readings are acquired on principal streams. Methods involving conditional probability require the capability to store historical readings and — depending on the complexity of the system and the degree to which causal maps are integrated — processing power could limit real-time applications.

2.7 Conclusions

Participant responses to the causal maps applied during the pilot test were enlightening and suggest that with additional development an adaptive interface incorporating the transfer entropy relationships between variables in a complex system will be useful in improving operator understanding of couplings and dynamic effects in process facilities. Three participants completed the pilot study, which is not sufficient to develop broadly generalizable results or make attempts to quantify the benefits, but was sufficient to propose improvements to the experimental design and pursue further development and assessment of causal maps as real-time decision support for operations.

The application of the layered framework was unquestionably valuable in separating elements of the interface, and the value of some features, like the ability to toggle through increasingly complex causal maps in different combinations. The value of that specific feature would not likely have been demonstrated so clearly in a more traditional control systems experiment. The number of times the feature was accessed and the time spent looking at it were not as direct a measurement of utility as the number of new, correct inferences that were made during the B-test.

Causal maps derived through transfer entropy may be effective bases for improving operator understanding of complex causality, and the continued design and testing of the interface developed for this project could contribute to methods for design and evaluation in the chemical process industry.

Acknowledgements

We extend sincerest thanks to the participants in this study, generous with their time and insights. Thank you to Carrie Demmans Epp and our classmates for their thoughtful suggestions and peer reviews throughout the course.

Funding

This research is funded by NSERC RGPIN-2018-05170-Lefsrud and NSERC COVID-2020-06431 Lefsrud “Developing a Safety Management Systems (SMS) Model”.

Chapter 3

Chapter 3 of this thesis has been accepted for publication as Michelle Naef and Lianne Lefsrud, *"Smooth operator: aligning performance assessment methods with design and operating objectives"*, in the Journal of Loss Prevention in the Process Industries.

Contributions of the authors as follows:

Michelle Naef: Conceptualization, Methodology, Investigation, Formal analysis, Writing – Original Draft, Writing – Review & Editing

Lianne Lefsrud: Supervision, Funding Acquisition, Writing – Review & Editing

Research materials not submitted for journal publication, including the Research Ethics Board project and study instruments are included in Appendices 1 and 2.

Executive summary

As the pace of innovation in process automation continues to accelerate, the challenges of operationalizing these advances are of increasing interest. Reducing loss incidents in the process industry requires increasingly knowledgeable operators to access multiple sources of current and historic data through industrial informatics. To evaluate the effectiveness of new designs and design features, clear criteria for performance and performance improvements must be developed. The situational design model was developed to assess and develop problem solving skills, the higher cognitive processes necessary for high performing operators to employ in abnormal situations. The situational design model recognizes learner readiness and the disambiguation of "recurrent' and "non-recurrent" skills. In this paper, the situational design model is demonstrated in a workplace-analogous study to evaluate the effectiveness of Distributed Control System interface adaptation. This type of structured assessment can be used to tie design features more directly to the specific skills of interest and improve the value propositions for new features and designs. The demonstration in this study shows the possibilities for structured assessment and suggests next steps toward aligning business and operations objectives with design evaluation techniques.

3.1 Introduction

Process operator assessment has become a topic of increasing interest in the chemical process industry as development of automation technology continues to progress rapidly (Kluge et al., 2014). Contrary to optimistic predictions of business analysts (Venkatasubramanian, 2019a) automation has not significantly reduced the frequency of abnormal situations requiring intervention. Operations workforces have been reduced and the span of responsibility for operators has expanded, meaning the abnormal situations faced by individual operators are more varied and require a broader knowledge base to address (Castillo-Borja et al., 2017). "Competency of personnel" is often suggested as a contributing factor (Baybutt, 2016), but research into industry practices is discouraging; few formal or systematic assessment methodologies are applied to competence development in operators (P. T. Bullemer & Nimmo, 1998; P. T. Bullemer & Reising, 2015). When facility designs or design features are evaluated for effectiveness, there is often an implication that these changes will improve performance, but "performance" or more broadly competence in process operators is an ambiguous standard (Jennings, 2020; Kluge et al., 2014). If one of the design objectives for any innovative technological adaptation is to "improve performance" then reliable, repeatable, and industry-applicable methods to evaluate competence are necessary.

The strategies employed in the chemical process industry to assess operations competence were not designed to build expert problem solvers or evaluate higher level cognition. The pilot study undertaken by (Naef et al., 2022) sought to evaluate the value of a visual cue based on transfer entropy, overlaid on a typical operations interface, during a series of simulated operating tasks. It was difficult to identify a suitable assessment methodology to quantify operator understanding and determine whether the design feature had improved understanding, and attempting to project the value of the design feature to meaningful operating objectives was frustrating. That finding led to a review of competence assessment techniques in the process industry, which was necessarily connected to the training practices in use. There is an apparent need in the field of industrial training for a systematic, rigorous approach to competence, from the development of consistent training curricula through to the design and continuous improvement of assessment tools (Kluge et al., 2014; S. Lee et al., 2000; Ross, 2013; Stawarz & Sowerby, 1995). Current trends in informatic design reflect a desire to reduce overall cognitive load and focus user attention meaningfully evaluating the degree to which technological features foster better application of non-recurrent skills could unlock latent value in existing technology integration.

In this paper we demonstrate a practical application of the situation design model (Zhong & Xu, 2019) to evaluate the utility of a new design feature on the decision-making performance of participants interacting with a simulated DCS screen. The study seeks to answer how the situation design model can be applied to assess task performance improvements in a complex problem-solving activity. The demonstration experiment showed the feasibility of applying the situation design model and related assessment instruments in a workplace-sized group (n=32) of knowledgeable participants. Structuring the design evaluation using the situation design model offered the opportunity to disambiguate results and analyze performance from a variety of perspectives, beyond the simple "control/test" model. The study also highlighted challenges for the Chemical Process industry in adopting more complex frameworks for performance evaluation and indicates the additional rigor that can be applied to design activities promising "performance improvement" at the panel.

3.2 Literature review and context

The challenge of developing problem-solving skills is not unique to the chemical process industry, and one of the persistent challenges to problem solving research is the imprecise definition of "problem solving" consistently applied in the business world. Zhong and Xu (2019) revisited the definition gap observed in workplace-readiness surveys and suggested that a key distinction be made to reflect the different skills necessary to solve new problems (non-recurrent) versus well-understood or recognizable problems (recurrent). "Real life problem-solving relies on the interaction between the two systems." (Zhong & Xu, 2019) This application of information-processing theory aligns with key themes in the role of the chemical process operator.

Zhong & Xu, (2019) further detail that the development of expert problem solvers in any domain requires the mastery of necessary recurrent skills as well as the development of non-recurrent skills, and the educational techniques necessary to support growth in both dimensions are different, and they select assessment instruments and analytic techniques to measure and disambiguate results that are due to individual readiness, application of recurrent skill and application of non-recurrent skill. This presents a possible opportunity for the process industry, where critical reviews of the effectiveness of current operating procedures and routines have been published regularly (D. A. Kennedy & Whittaker, 2006; D. Kennedy & Whittaker, 2000; Park & Jung, 2015; Strobhar, 2014e). The development of non-recurrent skills is dependent on recurrent skills to some degree, and real improvement in non-recurrent skills requires that learners achieve mastery in the correct recurrent skills. Accurately measuring the impact of design changes to any element of the plant operating environment, including the control panel interface, requires ways to disambiguate these factors and consistently gauge performance to some standard.

As the roles for process operators grow to include more control loops and sub-processes (Strobhar, 2014d), the body of knowledge necessary to select and apply procedures is itself a significant learning effort (P. T. Bullemer, 2013). The concept of competently assessing an abnormal situation and being able to diagnose the cause, then select the appropriate procedural response is an example of successfully applying a "recurrent" skill. Process operators must also be able to diagnose an abnormal situation that is unique or sometimes a situation that appears routine but does not respond predictably to the procedural solution. Those abnormal events are examples of needing "non-recurrent" skill, requiring the operator to make decisions and judgments based on related knowledge.

Research regarding operator performance with respect to domain knowledge focuses on error reduction (Naef et al., 2022) and the assessments seldom reflect the complexity of the tasks undertaken. Even in simulator-based testing schemes, simple questions regarding defined procedures are weighted the same as questions that rely on detailed mental models to predict an outcome. Typical training programs in the process industry rely on didactic presentation of operating procedures, which may or may not be well-designed or effective (Park & Jung, 2015). Where formal assessments are applied, examinations tend to be conventional multiple choice or word-problem structures with short form responses, where the trainer and grader may not be a domain expert (Strobhar, 2014f). Doe et al., (2005) summarized the results of a human-centered interface design effort that quantified the benefits of improved operator performance in detailed, expert-judged scenarios; a paper that remains unique in the field but highlights the latent value in developing training scenarios and rigorous assessments.

3.2.1 Readiness – cognitive load and socio-emotional factors

Existing research designs in the process industry lack a clear tie between cognitive load and skill acquisition, by focusing on error. Many of the assessments are reductive, and to objectively quantify and analyze the results, the tasks assessed in "attention" studies on operators bear little resemblance to the roles they are required to take on during abnormal situations. (Strobhar, 2014b) summarizes the typical approaches to display decluttering, detailing the disordered approach to interface design that incorporates ideas about ergonomics, graphics design and a superficial approach to task load reduction that fails to differentiate between important information and "noise". There is a persistent challenge during the design of informatics that carries over into worker training due to the dependence on computer-based simulations (Brambilla & Manca, 2011). Operators must access information from many sources during abnormal situations and relying on recall for facts or statuses that can be displayed can tie up working memory rather than relieve load. The results of those studies do not offer insight into whether the "error" observed was due to insufficient mastery of necessary recurrent skills, insufficient willingness/readiness to complete the task presented, or insufficiency in non-recurrent skills when non-routine problems are presented.

The chemical processing industry has broadly recognized that cognitive load or stress is a factor in abnormal situation management (Das et al., 2017, 2018) and assessments seeking to improve problem-solving performance by reducing cognitive load have become common. Methodologies have been developed and are used to assess cognitive load in process operators (Hart, 2006; Hart & Staveland, 1988; Ikuma, Harvey, et al., 2014) and task loading has been recognized as a design factor worth considering. There is a drive in the design of process interfaces to "declutter" displays (B. P. T. Bullemer et al., 2014) and reduce the volume of inputs

(Strobhar & Harvey, 2011). Industrial design and training programs have superficially integrated the idea that cognitive load must be reduced to "free up" working memory and support the necessary non-recurrent skills for problem-solving. Kluge et al., (2014) offered a wide-ranging analysis of the field and concluded that few -if any- studies systematically interrogated conceptual understanding and cognitive load was largely used to explain the frequency of errors. "Errors" are defined inconsistently, the same way that knowledge tests applied in training and evaluation fail to account for complexity or difficulty of the task. The unfocused approach to cognitive load management is echoed in the often-ineffective assessment techniques applied in worker selection, training, and performance evaluation. A high personal stakes test is given, that might be scored by someone with little or no domain-specific knowledge, where errors of any type are weighted equally. Minor errors in recall for recurrent skills are weighted the same way as critical flaws in mental models used to generate solutions to novel problems. The assessment techniques do not offer deep insight into how well operators understand the complex systems they interact with, and without that insight, improvement to the training process is difficult to implement. The pilot study Zhong & Xu, (2019) presented included examples of assessment techniques and quantitative measures for evaluating performance. Key elements of the assessment included a self-readiness assessment and a problem-solving task that was developed by experts and scored independently by two domain experts. Further, the assessments were applied throughout the training cycle to assess individual readiness and with the primary objective of selecting appropriate learning activities. That change, moving toward a supportive educational model and using formative assessments might prove to be the most challenging element for industrial workplaces to adopt (Wilkins, 2011).

3.2.2 Comprehensive assessment model – recurrent and non-recurrent skills

Table 3-1 maps the cognitive requirements associated with the aspects of occupational competence Kluge et al., (2014) identified for process operators to the situation design model and learning taxonomy. The table deconstructs causality classifying causal relationships in complex systems as "complex, simple or steady state" based on the number of state variables involved and the number of time intervals under consideration. This causal complexity model served as one basis for the evaluation criteria used in the study. Further deconstruction of the aspects of complexity relied upon the ontological structure applied in asset management, conceptualizing nodes (typically assets, like a reactor or a single control loop), systems (connected nodes) and blocks (connected systems) as complimentary aspects of complexity to evaluate the quality and depth of thinking on the part of participants. Using these categories to deconstruct operating tasks allowed the re-grouping of several aspects of complexity and a more detailed examination of two complex areas, "couplings and interactions" as well as "dynamic effects." Additional detail on the development of Table 3-1 is included in Appendix C.

Table 3-1 Aspects of complexity mapped to situational design model and learning taxonomy with proposed adaptation

Aspect of complexity	Proposed complexity categorization	SDM	Learning taxonomy
Couplings and interconnections	Asset node data	Recurrent	Factual/descriptive
	Systems or single loops	Recurrent	Factual/descriptive Conceptual/procedural
	Blocks – combinations of systems	Recurrent	Conceptual/procedural Metacognitive
	Plant-wide, multiple blocks	Recurrent	Metacognitive
Dynamic effects and MPC and RTO implementation and multiple or conflicting goals	Steady state or single time interval	Recurrent	Factual/descriptive
	Simple causality - recurrent dynamic effects	Recurrent	Conceptual/procedural
	Simple causality – non-recurrent	Non-recurrent	Metacognitive

	Complex causality - recurrent	Recurrent	Conceptual/procedural
Multiple or conflicting goals	Complex causality – non-recurrent	Non-recurrent	Conceptual/procedural Metacognitive
Multiple or conflicting goals	Socio-emotional	Learner readiness	Task design and participant response
Crew-coordination complexity	Socio-emotional	Learner readiness	Assess impact or design task to eliminate
Non-transparency	Incomplete schema, limitation of existing design	Learner readiness	Task design and participant response, design evaluation
Alarm flooding	Task load, cognitive stress	Learner readiness	Measure during task, collect baseline, compare

This method requires a substantive knowledge base, to begin the development of problems. The concept of domain knowledge, and domain knowledge bases has been a topic of interest in the process industry, particularly in process automation, as an ontological knowledge base used to structure data storage offers many advantages to the development of analytic tools. Ontology knowledge bases, like MIMOSA (Machinery Information Management Open System Alliance) (Drever et al., n.d.) and OntoCAPE (Ontology for Computer Aided Process Engineering) (Morbach et al., 2007) offer a valuable structure and framework for design feature evaluation and the evaluation of operator competence – where competence is defined at least in part by an operator having expert-level domain knowledge and expert-level problem-solving skills within that domain. The hierarchical structure for assets and systems that was adopted for this study is based on the relationships captured in the OntoCAPE and MIMOSA Ontology Knowledge Bases.

The three axes for assessment in the Situational Design Model allow disambiguation in performance of complex tasks; higher proficiency in recurrent tasks or higher readiness is not a reflection of improved problem-solving performance. As a workplace model, for the purposes of design evaluation, the relationship between the design feature and the improvement in problem-solving skills is of interest.

3.3 Methods

This study was conceived as an operationalization of the methodology described above in the evaluation of design features related to control panel displays. The 2015 Tennessee Eastman simulation, built in Simulink (Bathelt et al., 2015), was used as the "plant model" for this experiment, and an industry-typical display screen was developed to display measured values, some trends and incorporate 5 of the 12 possible tunable parameters in the "MultiLoop_mode1" control scheme. The causal maps developed for the pilot study in (Naef et al., 2022), based on the hierarchical method described by (Suresh et al., 2019) were given to participants in the test group. The control group did not have access to the causal maps during the simulation task. The instruments used in the Zhong & Xu, (2019) situation design study were adapted to evaluate learner readiness and participant performance in recurrent and non-recurrent skills related to control panel interaction and model predictive control (MPC). Figure 3-1 is a screen capture of the user interface shown to participants for all tasks, during steady state operation. Figure 3-2 shows the trends available to users during the production run, captured at the same time interval. Users had the ability to alter the position and size of the trend screens. The complete marking rubric and instructions sent to markers is included as Appendix 3-A. The instructions sent to the markers include the participant task instructions, as well as the theoretical breakdown of the marking structure.

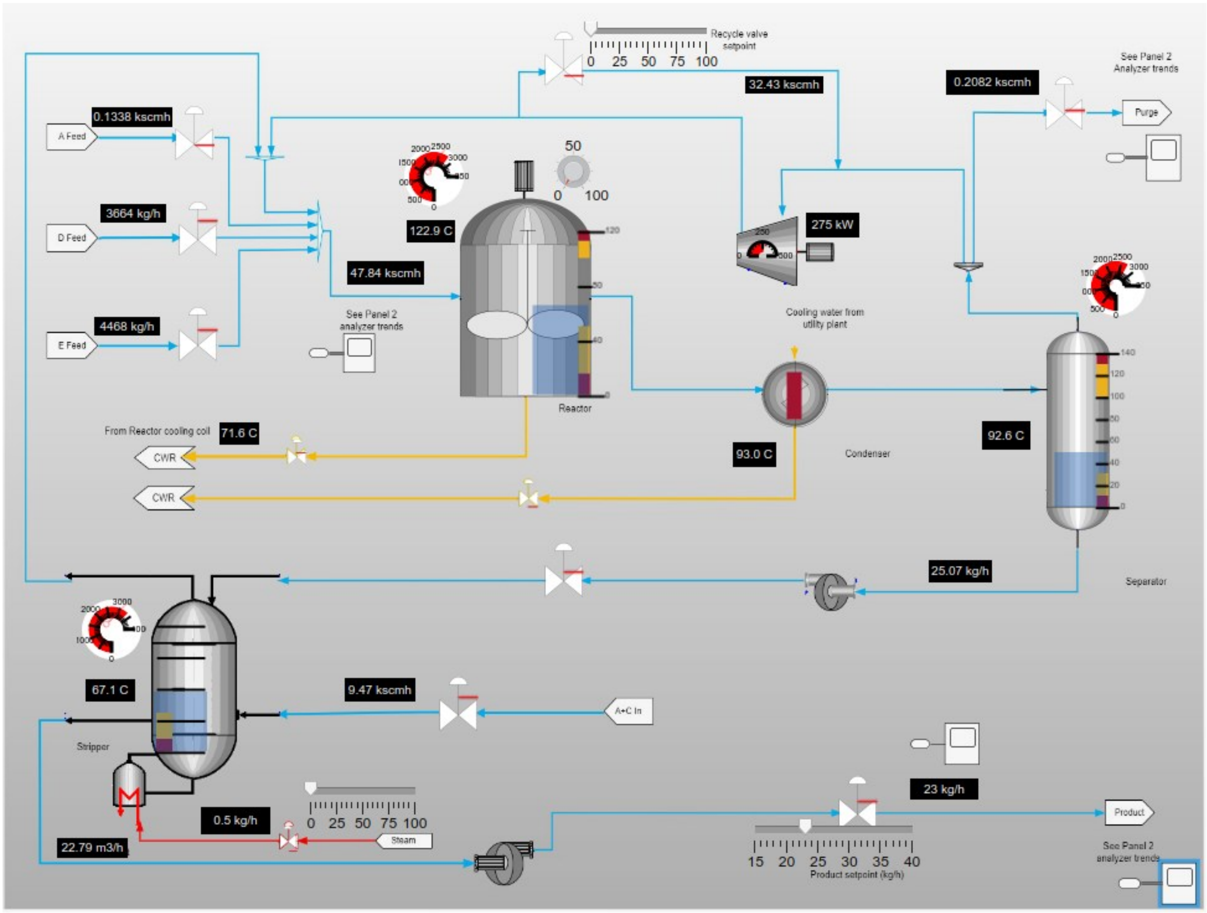


Figure 3-1 User interface shown to participants representing the revised Tennessee-Eastman Control Problem (2015) with four (4) tunable parameters. Screen captured at T=69 during a "normal" run of the simulation.

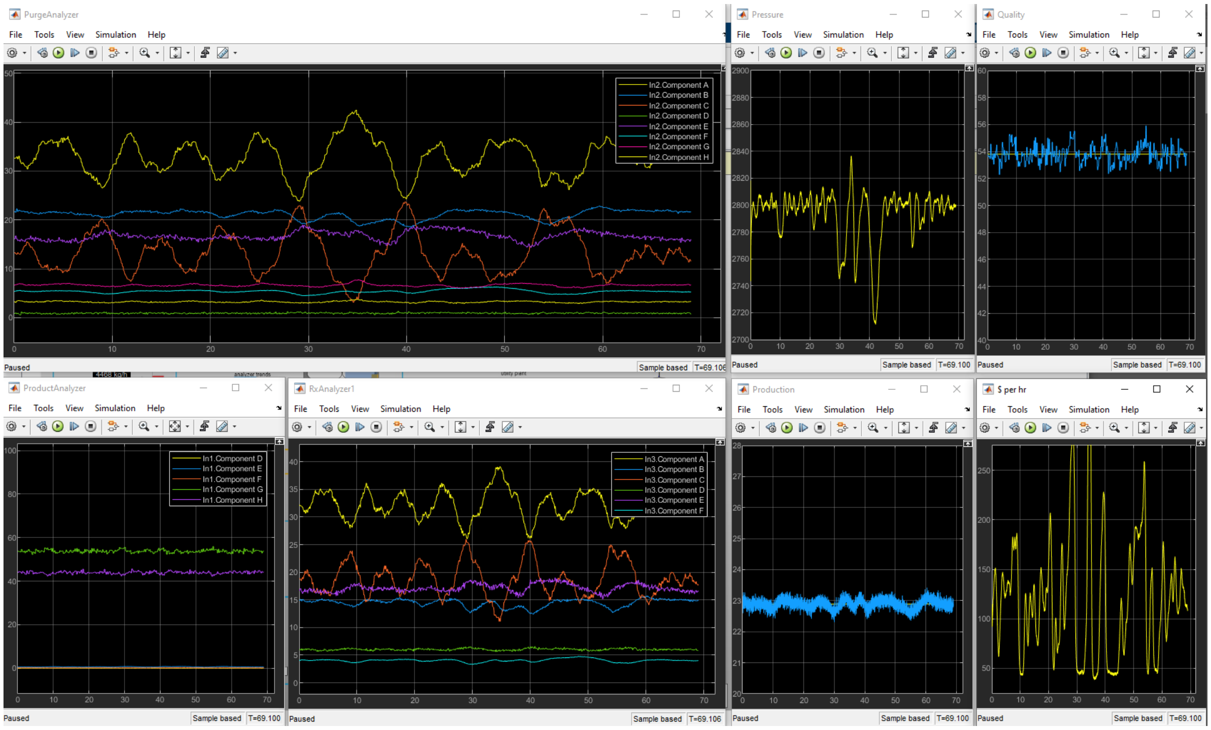


Figure 3-2 Trends available to users during production run simulations, shown here at full screen resolution for demonstration purposes. All trends were screen captured at T=69 during a "normal" production run

This study was initially conceived to evaluate design feature effectiveness in a simulation task over two similar sessions. The study design reflects workplace-typical constraints, including cohort size where n<100, intended to reflect typical crew sizes in operating facilities. Instruments and techniques designed around achieving statistical significance of results and broad generalizability were not selected for this study, as design implementations of technological innovations in the process industry tend to occur at the facility level, and most commonly at the unit or plant level. Participants were recruited from a pool of undergraduate and graduate students enrolled in Chemical Engineering. A total of 35 participants were recruited, and 32 completed both sessions of the study. 17 participants were enrolled in the Computer Process Control stream, and the balance of participants were from a variety of subspecialty areas. Participants were asked to complete the Learning Readiness survey using the Institution's course management platform, and then complete a 20-question multiple choice exam developed by the research team. Participants were then asked to open the plant simulation and answer 4 long-answer format questions about the simulation. The simulator was designed to toggle between "steady state" and "abnormal" operating modes, and participants were free to run as many simulations as they felt necessary. The study took place over two weeks in a university computer lab and was supervised by a member of the research team.

The multiple-choice exam was developed to evaluate knowledge in three areas and questions were adapted largely from industry and academic sources. Table 3-2 summarizes the content of the multiple-choice exam.

Table 3-2 Content tested by multiple choice exam, linked to aspect of complexity and learning taxonomy

Subject area	Aspect of complexity	Learning taxonomy	Number of questions
Process control theory	Asset node data	Factual/descriptive	6
	Systems or single loops	Conceptual/procedural	3
Industrial applications of control theory	Systems or single loops	Conceptual	8
Model predictive control	Systems or single loops	Factual	3

The simulation task was conceived as a root cause analysis or plant troubleshooting exercise, like the task given in the earlier pilot (Naef et al., 2021). The simulation task in the first session did not elicit the type of responses expected, and the task was modified for the second session where two of the four questions related to optimizing the simulation production over the run, and two questions related to an abnormal situation that presented as a plant-shut down due to high stripper level before the simulation run concluded. The root cause of the plant trip was a malfunctioning positioner on the overhead recycle valve, one of the tunable parameters in the simulator. The slider for setpoint showed a new position, but the simulator ran at a fixed set point of 50%. This scenario was selected to revisit several of the questions on the knowledge exam, which presented display errors for control loop elements. The eventual plant trip occurred on Stripper level, which is not a value directly connected to the overhead recycle measurement.

The responses to the tasks were marked following both sessions by three (3) graduate students in Chemical Engineering who were identified by their research or course work focus on control systems and plant engineering. Markers were briefed together but asked to mark the responses independently and provided with a detailed rubric developed using the methodology described above. Responses that included consideration of complex causality were awarded 5 marks, correct observations about simple causality were awarded 2 marks, and basic observations or reasoning involving individual components were awarded 1 mark. The rubric was modified to reflect the changes to questions prior to the second task completion.

3.4 Results

3.4.1 Learning readiness and knowledge exam

Survey completion rate was unexpected, with 27 of 35 session 1 participants entering any response, and responses sparsely populating. All 35 participants attempted and completed the subsequent 20-question knowledge exam, and all 35 opened and attempted the simulation task. As a result, the survey instructions were modified slightly for the second task, and anonymous responses were accepted. All session 2 participants were encouraged to complete the survey, and 24 opened the survey and completed at least one response. Of 32 respondents approached to participate, only one response indicated a greater-than-minimal level of knowledge about DCS interfaces or Model Predictive Control, and only 2 responses were recorded showing above-minimal experience with either DCS interfaces or MPC. There was no remarkable difference in responses between the first and second session, which occurred no more than five days apart and sometimes as close as the next day depending on participant availability.

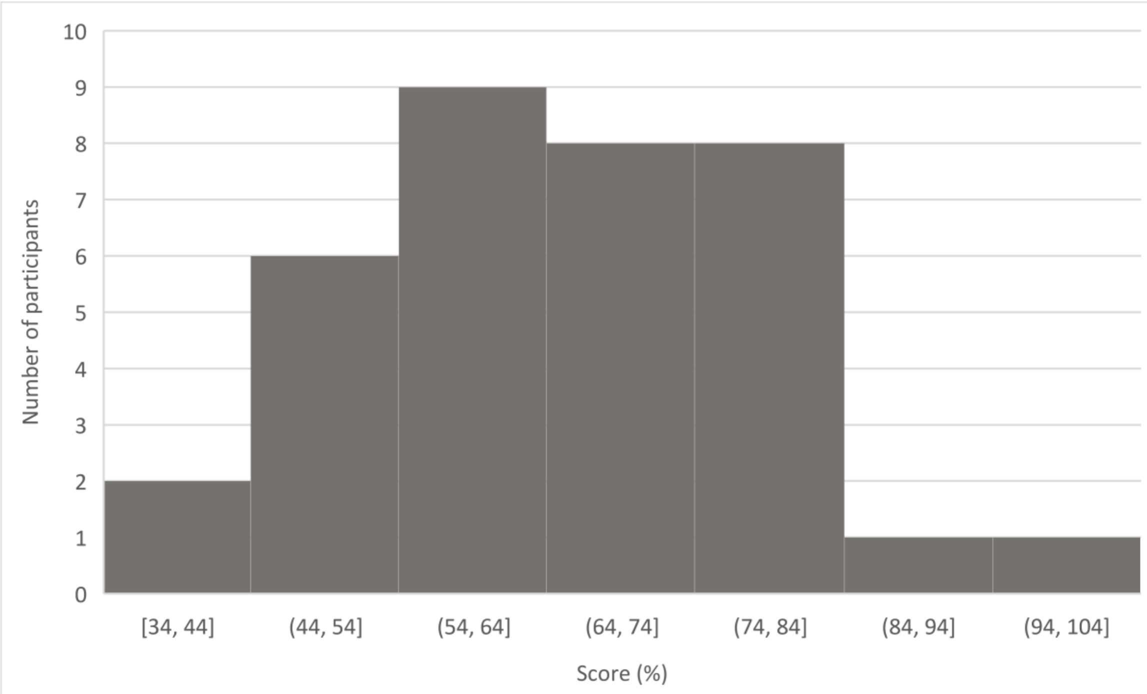


Figure 3-3 Participant scores on knowledge exam, 20 multiple choice questions

All participants completed the knowledge exam, and scores ranged from 34% to 95%, with a mean of 65%, distributed as shown in Figure 3-3. Combining the learning readiness survey results with the knowledge exam responses allowed the establishment of a threshold at 70%, where none of the participants scoring below 70% had reported any experience or knowledge of the subject beyond the minimal category.

3.4.2 Simulator task

The task was divided into two sections intended to encourage manipulation of the tunable parameters and interaction with the model predictive control element of the simulation. The first two questions were related to the Steady State simulation, and asked participants to use the tunable parameters to optimize production over a single run. The third and fourth questions related to the Abnormal Event simulation and asked participants to identify the root cause of the stripper level trip. The second simulator session forms the basis of the assessment of problem-solving skills with results from the first session simulator task dropped from the study. During the second session, there was no knowledge exam to complete, and participants were able to interact with the simulator immediately upon entering the lab, with encouragement to fill out the learning readiness survey.

Table 3-3 summarizes the results of the scored simulation task, grouping participants into categories of interest. The groups divided by experience threshold (combined knowledge test and readiness survey results) displayed weak trends in predicting problem-solving capability in the simulation tasks, with the mean scores higher in the "experienced" group.

Table 3-3 Simulator task results

Participant subgroup		Optimization tasks	Root cause analysis tasks
All participants (n=32)	Mean	10.9	9.6
	Standard deviation	5.1	6.1
	Range	[2.5,27.0]	[2.0,36.5]
Inexperienced cohort (n=21)	Mean	10	8.4
	Standard deviation	3.6	2.9
	Range	[3.5,17.5]	[4.0,13.5]
Experienced cohort (n=11)	Mean	12.5	11.9
	Standard deviation	7.1	9.4
	Range	[2.5,27.0]	[2.0,36.5]
Control Group (n=17)	Mean	12.6	10.9
	Standard deviation	5.5	7.6
	Range	[4.5,27]	[4.5,36.5]
Test Group (n=17)	Mean	9.2	8.3
	Standard deviation	4.1	3.9
	Range	[2.5,17.5]	[2,16]

The scores assigned did not explicitly indicate whether the task was completed successfully. This was of interest in the root cause analysis task, and further examination of participant responses was necessary. Ten (10) participants identified root causes in the correct subsystem, with responses identifying the recycle stream, recycle valve, and purge valve behavior included as "correct". Two (2) participants correctly identified the root cause as the recycle valve, with one participant determining the setpoint of the valve to be at 50%. The response indicated that the participant had compared the impact of the recycle valve on the reactor feed flow rates to the performance of the same subsystem in the Steady State simulation. Twelve (12) participants identified the reactor feed stream as the source of the problem, which was rated as partially correct, since the recycle stream immediately affects the controllers for feed flow. Two of those 12 participants identified a solution to the recycle valve malfunction in an unexpected way, both participants identifying the reactor level as an affected variable, with one reporting that the plant trip was avoided by increasing the reactor agitation speed. Identifying the reactor agitator as the source of the problem, meaning that their response was not captured as "correct" despite their action preventing the plant trip in the simulation.

3.4.3 Design feature evaluation

Performance among the control group was superior in terms of root cause identification, with 7 of 16 participants identifying the reactor feed as a problem and 7 of 16 identifying the recycle/purge subsystem as the source of the problem. The test group, who had access to the overlay which visually displayed causal relationships calculated using transfer entropy did not perform as well at the root cause identification, with 5 of 16 participants identifying the feed stream as the source of the instability and 3 of 16 correctly identifying the overhead subsystem as the root cause. There was no apparent correlation between non-recurrent skill performance and access to the design feature when data were grouped according to experience/knowledge. The design feature did not have the desired effect on participant understanding.

During both sessions, participants engaged in activities that were not predicted or controlled as part of the experimental design. In many cases, participants arrived at the same time, and sat together in the lab, in groups of two and three. There was no stated restriction on collaboration, and groups that arrived together worked together in a variety of ways, in some cases using one computer to hold the results of a previous run and comparing a new run on the second. Participants who did not collaborate with multiple machines developed ways to compare simulator runs, with screenshots and mobile phone photos appearing most frequently. Participants were asked what displays they would have liked to have seen, and few articulated that storing and comparing trend data between runs would be desirable, but almost all participants demonstrated the utility of that feature. Participants often collaborated in languages other than the language of the study, suggesting that providing instructions in English only was not optimal. While outside the scope of the study, the language fluency dimension is relevant in an increasingly international workforce with high worker mobility.

3.5 Discussion

This demonstration yielded insight into how assessment methodologies targeting cognition and understanding could be incorporated into operating workplaces and the chemical process industry in general. The situational design model offers a broad range of analytic possibilities to differentiate results and target supportive learning activities to improve performance. As a tool to evaluate design features, this model offers a more direct tie to business and operating objectives by directly measuring problem-solving performance as opposed to more indirect measures like attention or cognitive load. The situational design model was originally conceived to include supportive learning activities as well as an assessment framework. This study sought to examine the application of the assessment framework to the specific challenge of evaluating the effect of a design feature on the problem-solving skills of participants. Further extensions of this model into industrial competence and assessment could incorporate the supportive learning elements and offer a more systematic and repeatable approach to industrial training.

Table 3-4 Summary of key findings grouped by element of the situational design model

Key finding	Recommendation
Learning Readiness	
Limited response to surveys	Structured interviews and review of formal education
Survey responses misaligned with education and experience	Participant screening questions at enrollment or during the study
Recurrent skills	
Distribution of scores within expectation	Knowledge assessment effective; key utility as a threshold to group performance results
Learning phase associated with new interface	Formal training phase for simulator with guided tasks – tuning parameters and pulling trends
Existing interface barriers	Use of familiar simulation interfaces or formal training element to establish recurrent skills
Gap between knowledge test and procedure-based skills	Incorporate familiar abnormal conditions or optimization problems into the study design
Non-recurrent skills	
Assigning meaningful performance scores	Further develop the ontology-based dimensions of competence and align task outcomes with scores, avoid compound tasks
Eliciting meaningful responses to questions during simulation tasks	Pilot questions and simulation task, use subtasks for improved analysis of results, combine structured interview questions with task completion
Marker preparation and consistency	Formalized marker training and audit prior to study completion, further development of rubric

3.5.1 Learning readiness

The limited response to surveys was an important finding, as learner readiness surveys (Delahaye & Smith, 1995), particularly in the format presented in this study (Zhong & Xu, 2019), are frequent elements of adult education programs (Sloan & Scharff, 2022). The readiness surveys employed in this study were intended to gauge participant knowledge and experience, paired with a short exam testing basic concepts and simple applications of process control theory to establish learner readiness and recurrent skills. The sparseness of the survey responses, and the appearance of qualifying statements in the responses to the simulator tasks like "I'm not very well-versed in this…" support the idea that readiness and its relationship to motivation are important in comparing responses and task performance within the subject group and potentially against benchmarks. These data also suggest the instrument employed in the study was not optimal. This study was conceived to disambiguate the role of prior knowledge and experience when trained personnel are asked to perform a challenging task requiring application of non-recurrent skills. Objective measures of knowledge and experience play a role, but an attractive feature of the situation design model is the incorporation of subjective measures and self-assessment, broadly grouped under "Learning readiness" as dependent factors. Low completion rates and conflicting participant responses between the recruitment emails and informal interactions during the study make survey results challenging to apply to performance analysis. One possible cause for the sparse responses on the learner readiness surveys in the study is survey fatigue (Fass-Holmes, 2022). Another possibility is the persistent and concerning tendency for discouragement and "imposter syndrome" among the study population of university engineering students (Litzinger et al., 2005; Young et al., 2018).

Participants interacted with researchers during the sessions, and many freely offered statements about their experience or perceived readiness to address tasks. One participant completed the second session and reported that they felt "excited about what's to come" in their future as a Chemical Engineer. These responses were important in gauging the effectiveness of the study but were not captured by any of the instruments employed, which was itself instructive. This suggests that an informal interview would be more effective in collecting relevant data about knowledge, experience, and motivation. Reliance on self-directed, one-way tasks is a persistent criticism of workplace training (Albert & Hallowel, 2013; Ross, 2013), and because participant readiness is one of the three axes for assessment proposed in the situational design model, reliable methods of collecting complete data about readiness is of particular importance.

Kluge et al., (2014) identified organizational factors as an element of complexity affecting process operators, as well as competing objectives. This study effectively grouped those environmental influences under “Learning Readiness”. These factors are clearly relevant to participant performance in challenging simulation and operating tasks, but eliciting accurate, complete information about how environmental factors affect learner readiness is a more significant problem that was initially conceived (Baker & Lefsrud, 2019). Six questions using Likert scales were insufficient to capture the influences relevant to classifying participant experience, knowledge, readiness, and socio-emotional response to external factors. Additional insight was gained by collecting informal interaction throughout the study, as well as reviewing the participant responses to the recruitment letter; considerably more experience and interest were present in the sample population that was recorded on the readiness surveys. Further refinement of the readiness dimension is necessary to develop the situation design model fully to industrial performance assessment and design evaluation, but the results of this study support effort in that area. The readiness surveys in the Zhong & Xu, (2019) pilot were used to group participants and offer supportive learning activities suitable for their relative positions on the readiness scale. Despite relatively sparse responses, triangulation from informal comments and recruitment responses allowed a combined readiness/knowledge threshold to be established in the study population for improved analysis of results.

3.5.2 Disambiguating recurrent and non-recurrent skills

Participant interaction with the interface, particularly the creative methods employed to preserve data between runs, was valuable insight into necessary design features. Those insights were not captured effectively in the participant responses to questions during the task. Trend histories are common elements of control panel interfaces and would have been relatively simple to add to the study simulator. This type of gap, and similar basic interface elements affecting task performance could have been identified by thorough piloting prior to study completion. Participant questions and responses indicated that one of the principal problem-solving techniques employed during the task was to compare the two simulations and look for differences. This process is important, but because simple comparison allowed many participants to provide at least a partial answer to the task questions, more detailed prompts or subtasks are necessary to evaluate the degree to which participants understand system interaction and dynamic effects.

Further development of the problem-solving task should be pursued, avoiding the tendency toward behaviourist assessment techniques that evaluate conditioned responses in tasks, or measure outcomes only, with little insight into causal factors. To better integrate this method with existing workplace training methods, an assessment of specific recurring skills would be a good addition to the study methodology. Incorporating procedural competencies into the assessment methodology should reflect existing research into operating procedure effectiveness [43]–[46], which was outside the scope of this study. Tasks requiring the diagnosis of familiar problems and the selection of the correct response (procedural) (Strobhar, 2014e) would allow improved differentiation between recurrent and non-recurrent skills that a basic knowledge test is limited in assessing. Zhong & Xu, (2019) positioned “recurrent” and “non-recurrent” skills as orthogonal vectors; most learning and assessment tasks will have some combination of both elements. The skill design and marking rubric need to explicitly identify the elements of both and assessment outcomes need to be communicated such that there is separation between the two skill types.

The marking rubric employed for the simulation task was developed according to the structure mapped in Table 3-1, and scores were assigned cumulatively, increasing in value with the degree of complexity associated with the response. Markers for the task were graduate students focused on process control and marked submissions independently. The second simulator task had two major components, an optimization task on the stable operating condition, and the trouble-shooting task where the recycle valve positioner was rendered unresponsive. Scores for each task were effective in analyzing the results, where an averaged task score did not provide as much insight into performance. Performance scores between the two tasks were not consistent; performing well on the optimization task did not have a measurable correlation with performance in the troubleshooting task.

The rubric applied in the study awarded marks for the level of complexity applied in reasoning and did not distinguish between correct and incorrect identifications of the root cause. This type of marking scheme is often employed to penalize guessing, but when evaluating the effectiveness of a particular adaptation may bear reconsidering. Further, when marker scores were analyzed, it was determined that one of three markers adopted a point-award scheme inconsistent with the other two. Those scores were dropped from the analysis, and the scores awarded by two markers were averaged. Inter-rater reliability was assessed first by inspection, and further statistical analysis would be of questionable value due to the scatter in marks awarded. Because the marks awarded depended on the participant describing the complex causal relationships between systems, participants who identified the correct root cause did not necessarily score very well on the task.

3.6 Conclusions

The situation design model for the development of problem-solving skills, conceived with the idea that real world problems are themselves tangled and complex, shows promise to align the goals of the process industry with the training and assessment practices for its human employees. This study was effectively a second pilot, a follow-up to the experiment descried in (Naef et al., 2022). A shift in thinking is necessary for research into innovation and technology, a move away from behaviourist evaluations targeting simple cognitive operations and reductive examinations of repetitive behaviours with weak generalizability to workplace performance.

Readiness, broadened to include the concepts of task load and external stressors is an important dimension in analyzing participant performance and should not be ignored in design feature evaluation. Few education/assessment models incorporate the concept of readiness as thoroughly as the situation design model, and further refinement of this model could be valuable in analyzing and interpreting the results of competence and design evaluations in the process industry.

Further development of rubrics for assessing competencies in operators would offer considerable value to designers and firms struggling with technology integration and performance improvement. Focus on simulator-based training is necessary and recognition that existing measures of competence in recurrent skills do not typically evaluate whether the procedures themselves are correct or consistent. Design adaptations to user interfaces can be tested by applying the rubrics during the design and testing process, by shifting the design evaluation paradigm to reflect the role of informatics as an educational tool rather than strictly as industrial assets. A problem-solving focus aligns with the goals of facility owners and operators, who are driven to restore stable operation quickly and with minimum impact to people, the environment, and equipment.

Acknowledgements

We extend sincerest thanks to the markers, Sushmita Mittra, Devavrat Thosar and Shahdab Pathan and to all participants in the study.

Funding

This research is funded by NSERC RGPIN-2018-1509-Lefsrud "Developing Safety Management Systems (SMS) Models".

Appendix 3-A

3-A.1 Marking Instructions

Please review participant responses, in the excel file, each participant is a separate row. All four responses are included in their entirety. Please enter your question score in the column titled "score x" (to the immediate right of the response.)

For Task 2, two separate simulation files were used. The questions have been changed to encourage more thorough responses, but the rubric is similarly applied. Participants were asked to use tunable parameters to try and optimize the production cost

For each question, please apply each possible mark only once. For example, in question 1, if they describe multiple pieces of equipment, only 1 point is awarded for the first item "Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed" If they describe the continuously stirred reactor (1 point) and then describe connected systems and the control scheme (2 points) a total of 3 points is awarded for that statement. Similarly, if they describe the reactor (1 point) and give a detailed description of the control scheme for the Stripper and how it is linked to the downstream product analyzer trend (5 points), they receive 6 total points for their "Interaction with simulator and observations".

Please also note that participants do not review these marks, and the marks awarded are design evaluation scores…not reflective of the "performance" of the participant.

3-A.2 Marking rubric for each question

Question 1 – weight = 0.2

Use the SteadyState model file to observe at least one full production run (72 hours) and study the Display to get some ideas about the process. Describe the process in general terms, focusing on what you believe the production objective to be. Which products are the saleable or desired products? What makes up the purge/waste stream? Are there any parts of the process you want to know more about, or display elements that you think would improve understanding? Did you alter any of the tunable parameters to get an idea how the process works? Which ones?

Table A.1 Marking rubric for participant task, Question 1

Interaction with simulator and observations		
Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed		1
Any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)		2
Description of basic characteristics, single nodes or individual trends		2
Discussion of what might occur or what might have occurred, description of tunable parameter impact on process		2

Discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge		5
Description of a system or subsystem not exhibiting transient behaviour		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2
Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Interface Utility		
Description of a single measurement, variable or trend that appeared to be missing		1
Discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture		2
Analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset		5
Total		28

Question 2 – weight = 0.3

Using the SteadyState model, try to increase production (increase production valve setpoint) as much as possible while maintaining a 72 hour production run. Note that this may require tuning of additional parameter setpoints, like the agitator speed, stripping steam valve flow and altering the overhead recycle flow.

The setpoints can be changed during the run.

The original setpoint for production in stable operation was 22.75.

Can you improve the total production? What was your best production curve? Please screenshot the best production curve from your tests.

What was the impact of your changes on the production cost ($ per hr)?

Did the change you made to one parameter have an unexpected result on another? Did that suggest any relationship between elements that you did not foresee?

Table A.2 Marking rubric for participant simulation task, Question 2

Measuring decision utility		
An increase in production above 22.75 with no explanation of which parameters were changed, and no discussion of limits within the simulation on further increases		1
An increase in production above 22.75 for some length of the run with some indication that a hard limit was observed, but no explanation of why that was related to the production value		2
Detailed description of how one or more parameters allowed an increase in production above 22.75 for some length of a 72 hour run and/or description of how the production rate was effectively limited by other elements in the simulation		5
Supporting logic		
Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed		1
Any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)		2
Description of basic characteristics, single nodes or individual trends		2
Discussion of what might occur or what might have occurred, description of tunable parameter impact on process		2
Discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge		5
Causality		
Description of a system or subsystem not exhibiting transient behaviour – no causal relationship described		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2
Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Total		28

Question 3 – weight = 0.3

Using the "AbnormalState" model, please interact with the simulation and examine the trend data. What do you think is the root cause of the abnormal situation that you observed? Please offer a root cause and follow it with a brief description of your logic in arriving at that conclusion.

Please describe any additional tests, measurements or diagnostic activities that you would recommend to increase confidence in your result. If you do not believe any additional information is required, please briefly state why you are confident in your diagnosis.

Table A.3 Marking rubric for participant simulation task, Question 3

Interface utility		
Description of a single measurement, variable or trend that appeared to be missing		1
Discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture		2
Analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset		5
Interaction with simulation		
Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed		1
Any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)		2
Description of basic characteristics, single nodes or individual trends		2
Discussion of what might occur or what might have occurred, description of tunable parameter impact on process		2
Discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge		5
Causality		
Description of a system or subsystem not exhibiting transient behaviour		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2
Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Total		27

Question 4 – weight = 0.2

What would you recommend the plant owner/operator do to address the root cause of this abnormal situation? Would the problem present the same way in all operating modes?

Table 3-A.4 Marking rubric for participant simulation task, Question 4

Measuring decision utility		
Description of a single, directly causal factor without substantial discussion around additional tests or diagnostics		1
Description of a possible causal factor drawing dependent systems into the analysis, limited identification of additional steps or tests, simple recommendation that does not incorporate system-scale or multiple time-interval considerations		2
Detailed and well supported root cause with one or more linked subsystems or components.		5
Causal reasoning		

Description of a system or subsystem not exhibiting transient behaviour		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2
Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Total		16

Submission score: multiply question percentage by question weight, submission score will be a percentage

Table 3-A.5 Submission scoring instructions

	Participant score (A)	Total (B)	Weight (C)	Weighted scores ‘= (A / B) * C
Question 1		28	.2	
Question 2		28	.3	
Question 3		27	.3	
Question 4		16	.2	
			Submission score (total of weighted)	SUM(column)

Appendix 3-B

B.1 Instructions given to participants

Please review the questions, and then open the simulation file that goes with each question. The first 2 questions requires use of the "SteadyState" MultiLoop_mode1.slx file. The final two questions are based on the "AbnormalState" MultiLoop_mode1.slx file.

The objective of the quiz is to **use the display interface to better understand the plant process.** Please describe your thinking in your answer, noting which features on the display were useful in your conclusions (for example, the product analyzer trend was useful in explaining the oscillation in the feed valves).

The response fields are free-text, but point-form responses may be appropriate. Assume that the question graders are familiar with standard process control terminology and aim for clarity in responses. The technical content of your response is of interest.

You may start this quiz at any time. Interact with both simulation files as you see fit, these are labelled 'steady state' and 'abnormal situation'. There are tunable parameters (indicated with blue arrows in the image below) that allow you to alter some of the steady state characteristics, but altering these parameters may cause a "shut-down". There are trends available to monitor the gas concentrations and production rate, circled in red. The trends can be accessed by double-clicking the scope icons, and dismissed with the "x" in the top right corner. The simulation is set to run for 72 simulated hours (approximately 2 minutes), and you can alter the speed using "pacing".

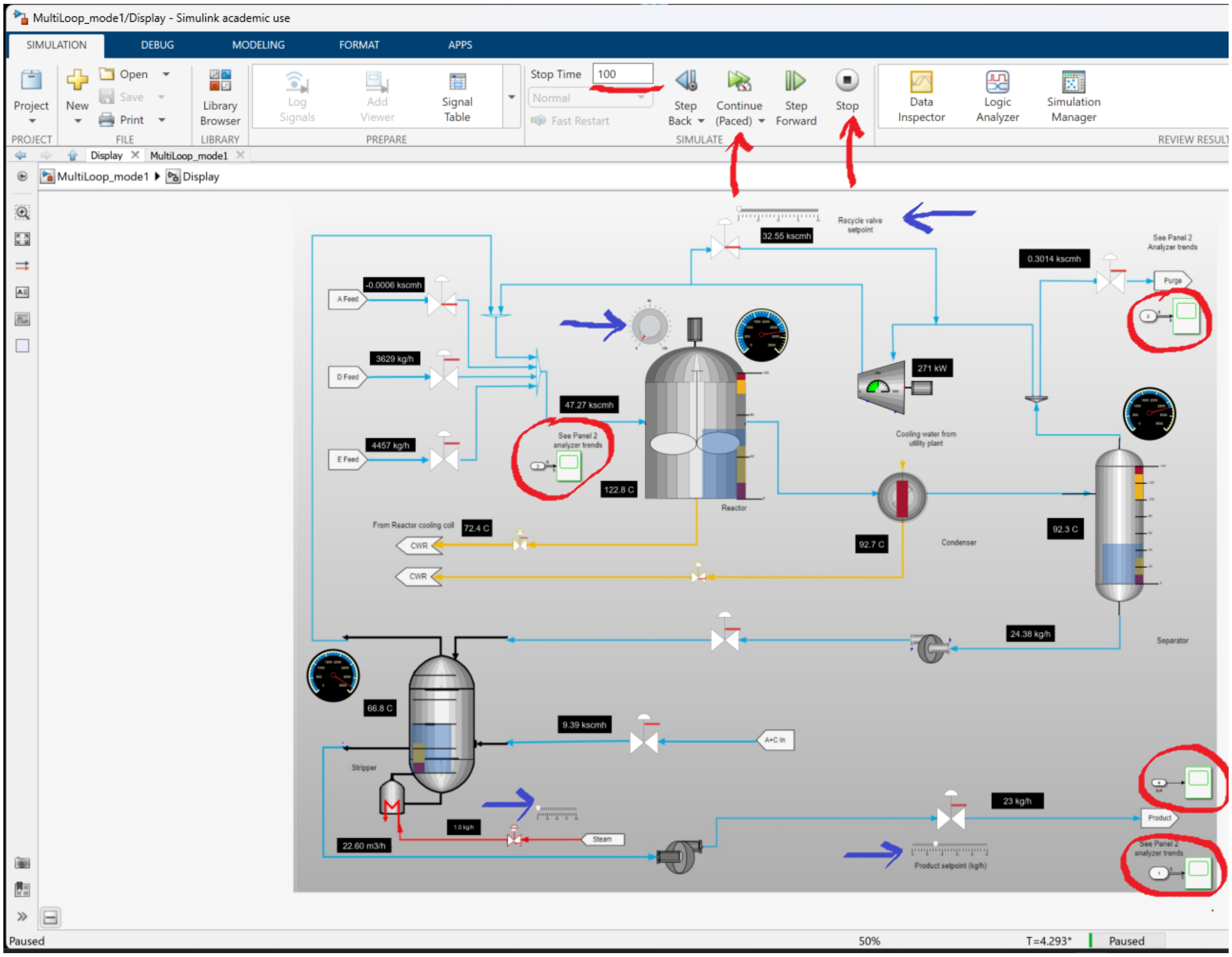


Figure 3-B.1 Screen capture from participant instructions, originally posted in the eClass experimental interface

Please notify the researcher if Simulink is not open on your computer.

Please note, the simulation files were prepared and validated by Bathelt et al. and cited/attributed according to their published terms and conditions.

The simulator and decentralized control scheme were not modified for the steady-state portion of this study. The Simulink objects added to the simulator for this study are largely cosmetic and consist of the "Display" layer and the addition of several signal/display blocks for the purposes of constructing a realistic control panel display only. The abnormal situation model **does not** represent the performance of the decentralized scheme developed by Bathelt et al. Complete details on the changes made for the abnormal situation model will be described and made available as appendices in any published works referring to this study following completion of all sessions.

Appendix 3-C

Table 3-C.1 Theoretical basis for rubric (summarized)

Aspects contributing to complexity	Reconstruction	Marks to be awarded
Couplings and interconnections	Acquisition of node data	1 – any discussion of process, control equipment that appears in the simulation, accurate description of what was observed
	Acquisition of couplings (simple causal links)	2 – any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)
	Identification of steady state characteristics (steady state mesh)	2 – description of basic characteristics, single nodes or individual trends
	Predict component effects in recurrent dynamic cases (alarm sequences, planned shut down)	2 – discussion of what might occur or what might have occurred, description of tunable parameter impact on process
	Elaborate, extrapolate and model complex causal relationships (downstream/knock on effects) perform root cause or fault detection on non-recurrent cases	5 – discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge
Dynamic effects	Steady state	1 – description of a system or subsystem not exhibiting transient behaviour
	Simple causality	2 – description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)
	Complex causality	5 – Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.
	Measuring decision utility	1 – description of a single, directly causal factor without substantial discussion around additional tests or diagnostics 2 – description of a possible causal factor drawing dependent systems into the analysis, limited identification of additional

		steps or tests, simple recommendation that does not incorporate system-scale or multiple time-interval considerations 5- detailed and well supported root cause with one or more linked subsystems or components.
Non-transparency	Incomplete schema Incomplete retrieval mechanism Incomplete interface Incorrect relationships implied	1 – description of a single measurement, variable or trend that appeared to be missing 2 – discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture 5- analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset
MPC and Real Time Optimization (RTO) implementation	Knowledge acquisition of methods, algorithms, conceptual contents of MPC and RTO Integration of insights into the mental model	2 – description of objective function (production cost), trend 5 – extension of objective function into any controller or loop behaviour
Alarm flooding	Prioritizing, simplifying or decoding unexpected alarm sequences	2 – logical discussion that shows evidence of prioritizing or sorting multiple alarms or rapidly changing measurement values.

Table constructed with extensive reliance on (Kluge et al., 2014)

Chapter 4

Chapter 4 of this thesis has been accepted for publication to the *Journal of Building Engineering* as Michelle Naef and Lianne Lefsrud, "*Application of engineering thinking for risk assessment in a Canadian elementary school*".

Contributions of the authors as follows:

Michelle Naef: Conceptualization, Methodology, Investigation, Formal analysis, Writing – Original Draft, Writing – Review & Editing

Lianne Lefsrud: Supervision, Funding Acquisition, Writing – Review & Editing

Executive summary

This paper outlines the application of research methodology for hazard assessment and adoption of controls that can be applied for aerosol hazards through a case study performed in an elementary school located in Alberta, Canada. Two classrooms with different educational activities planned were monitored for two weeks using multiple sensors to establish the impact of the activity on the accumulation of carbon dioxide in the space. Building operators were particularly concerned with the relative risks of choral singing as compared to traditional classroom activities during the COVID-19 pandemic, and little existing research supported decision-making in that area. The data collected in this study challenges the basis for public health controls in schools and demonstrates the feasibility of data collection and reporting outside the under-resourced public health departments. The classroom activities in this facility had little measurable impact on carbon dioxide levels. A challenge for public health officials during COVID-19 was in bridging a theoretical understanding of aerosol hazard transmission with the operationalization of that theory into effective risk controls for all facilities operating within a jurisdiction. The utility and effectiveness of risk controls should be re-evaluated routinely and supported by measurements and data as is feasible to collect, and this study demonstrates how researchers can bridge the gap between policy makers and regulated entities. Engineering thinking, and principles of industrial risk management and process control can be readily applied in any commercial building facility, and this study demonstrates some of the opportunities available to building operators and policy makers with the mass proliferation of affordable instruments with high utility for assessing building conditions.

4.1 Introduction

The COVID-19 pandemic, beginning December 2019 and extending into 2023 (Powell, 2022) creates an “unwelcome societal stress test” (Tufecki, 2021). As public knowledge about transmission has evolved, so do the prescriptive risk controls developed by public health authorities, which can make those controls unnecessarily controversial. Controversy is further amplified with retrospective examinations of pandemic performance in jurisdictions around the world. As Tufecki, (2021) highlights, people around the world were often confused and mistrustful of controls designed to protect against fomite transmission after research was released indicating COVID-19 was primarily transmitted as an aerosol. Moreover, a population with ready access to current scientific literature began to question public health decisions in real time with social media interaction playing a major role in the development of public health recommendations for the first time in history (Durand-Moreau et al., 2021). This resulted in an unhelpful combination of information and opinion overload (Berry et al., 2022), relatively inflexible prescriptive medical attitudes (Greenhalgh, 2021), and a flawed risk communication strategy (Huang et al., 2022). Concerns and confusion around built environments, in particular public buildings, became daily topics, and building operators were often overwhelmed by suggestions for filtration upgrades, measurement devices, and constant pressure by both public health authorities and a concerned public. This study examined the influence of classroom activities on CO_2 accumulation in a single facility to inform the risk management decisions of the building operators during the COVID-19 pandemic.

North American air health guidelines and design standards focus on ventilation rates - often communicated as air changes per hour - and prior studies have presented methods to calculate effective ventilation rates in school buildings (Aguilar et al., 2022; Petersen et al., 2016). The direct link between ventilation rates and exposure to airborne hazards has been questioned (Daisey et al., 2003; Godish, 1996) and the subject of many studies in schools (Daisey et al., 2003; Shendell, Prill, et al., 2004; Shendell, Winer, et al., 2004), which did not necessarily offer confidence to building operators during the COVID-19 pandemic when assessing resumption of certain activities. Tan et al., (2022) demonstrated the challenges to accurately modelling air flows, highlighting imperfect mixing and the unpredictable effect of obstructions, as well as turbulent mixing from a variety of factors. Multiple large-scale studies on classroom ventilation have been performed around the world, some of which were restricted to single-point measurements for airborne contaminants (Rosbach et al., 2016; van der Zee et al., 2017) or estimated averages for student generation of CO_2 (Shendell, Prill, et al., 2004). These methods did not offer high degrees of confidence in terms of estimating marginal risk relative to activity levels in specific classroom settings, due to questions surrounding mixing and the effect of specific ventilation systems, as discussed by Tan et al. Motuzienė et al., (2022) detailed further challenges for building operators in making day-to-day decisions surrounding occupancy and facility use, highlighting the limitations of occupancy models and the resulting gaps in building performance during operation. As an alternative approach O' Donovan & O' Sullivan (2023) presented a risk assessment method to building retrofit design that shared conclusions regarding the establishment of contaminant thresholds and the limitations of ventilation rates in terms of addressing risk from airborne hazards. A common concern of building operators and the public during the pandemic was that threshold ventilation rates did not address specific airborne hazards to the degree required for evidence-based decision making, resulting in the American Society of Heating, Refrigeration and Air-Conditioning Engineers releasing a position paper on the topic of indoor Carbon Dioxide in 2022 (American Society of Heating Refrigeration and Air-

Conditioning Engineers, 2022). Ferrari et al., (2022) summarized the challenges of air quality evaluation and offered a comprehensive review of ventilation improvement strategies for building operators, concluding that full-scale experiments would support design changes and retrofits where ventilation is insufficient. The gap observed in the research available at the time of study conception, even those studies which related to choral singing specifically (Bauer et al., 2022; Tamplin & Thompson, 2023) was that all studies noted ventilation as a driving factor in droplet suspension and aerosol contaminant accumulation. This study sought to operationalize air quality monitoring techniques for the purposes of facility-level risk assessment, when broad guidelines and recommendations were difficult to apply directly.

Methodologically, this case study demonstrates that measurement of a representative variable is accessible and can be performed cost-effectively, offering significantly more confidence to risk management over indirect guidelines and general performance thresholds. The calculation-based methods described in many prior studies relied on idealized conditions - often empty buildings or failing to account for partial height dividers and furnishings - and when evaluating the relative risks of specific activities or distributions of people throughout the space, the applicability was limited. The school administration that participated in this study had previously engaged the research team to assist in developing a risk management plan – to synthesize public health guidelines and develop practices for the initial re-opening of public school in September 2020 (Naef & Lefsrud, 2020) and were interested in the relative risk of choral classroom activities over traditional classroom activities. Pan et al., (2021) proposed that engineering thinking be applied to building ventilation and heating, ventilation, and air conditioning (HVAC) operation and highlighted the challenges and risk balances faced by many building operators. By answering their call for pragmatic guidance at all stages of building life-cycle, our approach represents an empirical, methodological, and theoretical contribution. This paper extends this research theoretically by detailing the application of engineering thinking to a building operation challenge, methodologically by demonstrating the approach recommended by Pan et al., (2021) and empirically in evaluating the impact of classroom activities on CO_2 accumulation in two classrooms. This study empowered a group of motivated, non-technical community leaders to collect data, analyze results, and communicate residual risk to their stakeholders around a controversial risk control applied to public schools – the elimination of choral singing and wind ensemble instruction. Six basic concept areas were identified using standard industrial and process engineering frameworks: 1) asset breakdown, 2) identification of mass and energy flows, 3) selection of measurement nodes, 4) specification of measurements, 5) data collection and analysis, and 6) decision-making.

4.2 Material and methods

The school community was particularly interested in assessing the marginal risks of instrumental and vocal music as compared to classroom activities that had been deemed "low risk" by public health officials. Two classrooms were selected based on planned activity levels over the three-week period under consideration. Classroom 36 was used by students from grades 1-6, led by the same choral teacher. Over the three-week period, some classes were engaged in seated pen-and-paper music theory exercises, while others were engaged in vocal ensemble performance. A range of "choral class" activities were performed and measurements collected continuously. Classroom 27 was used by a Grade 5 class, with one teacher leading most activity in what was characterized by school staff as a "traditional instruction" model. Specifically, students were usually seated at desks arranged in rows upon entering the classroom and instruction largely consisted of the teacher speaking and the students listening.

4.2.1 Asset Breakdown and Mass/Energy Flows

Assessing the facility using a standard ontological approach common in Process Engineering and Industrial Asset Management proved to be a valuable investment of time. In industrial asset management, the ontologies are designed to classify information about parts, machines and sub-systems that make up process facilities. The ontology can be described as a "schema", a repeatable method to break complicated systems into sub-components, based on the function and structure of the object under consideration (Gruber, 1993). Applying a standard industrial asset management schema to building mechanical systems facilitated discussion of both the experimental design and the analysis of results. Table 4-1 describes the asset modeling approach used in interactions with school administration and staff, and Figure 4-1 System schematic at two levels of complexityFigure 4-1 illustrates the concept from the top-most level and increasing in granularity to the classroom level, where decisions regarding sensor placement were made with support from classroom teachers. Energy flows are represented by red arrows, and mass flows by blue, but flows with combined heat and mass, such as the movements of people in and out of occupied spaces are represented by purple arrows.

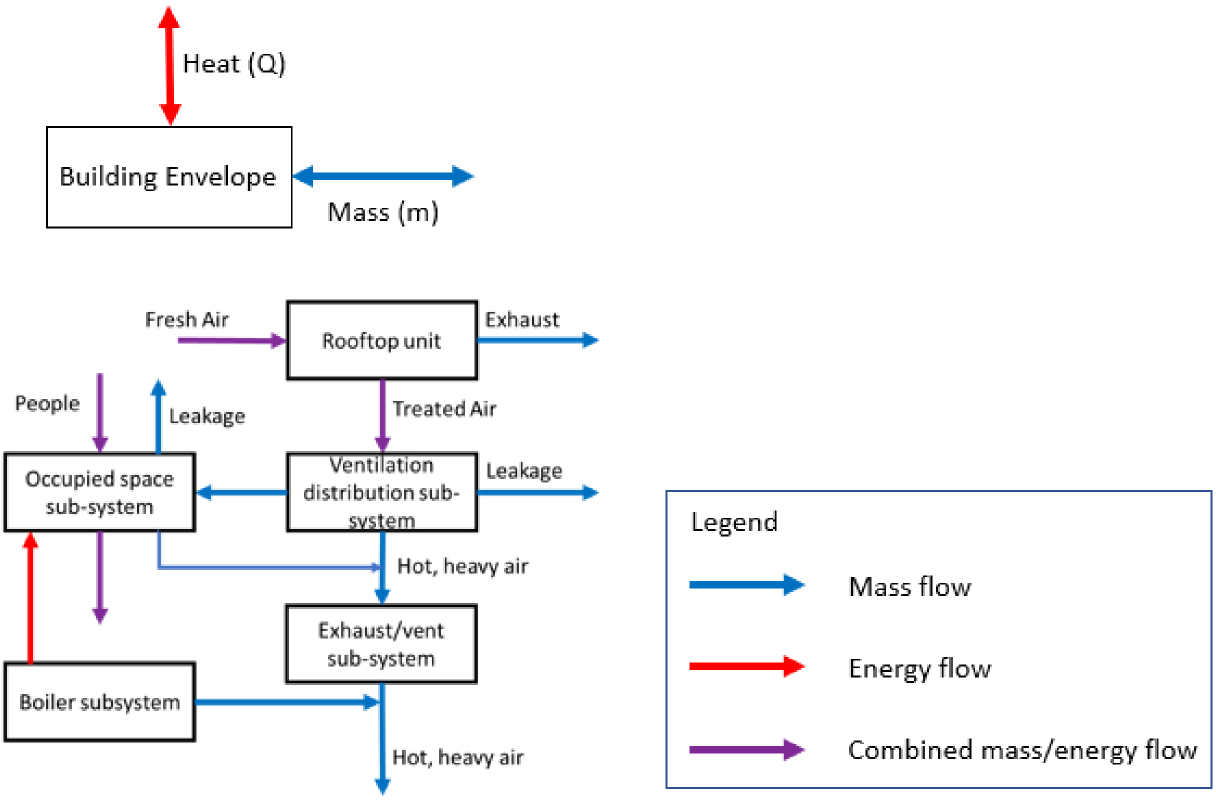


Figure 4-1 System schematic at two levels of complexity

Table 4-1 First level ontological breakdown of facility with mass and energy flows

Asset	Mass/Energy in	Mass/Energy out
Building envelope Properties: location, dimensions, insulation, occupancy	Fresh air (ventilation), Exhaust air, fuel, combustion exhaust, people in/out, envelope leakage (doors, windows)	Heat flows – people, boiler, envelope leakage (radiant and mass transferred), electronics
Highest level nodes (subsystems)		
Rooftop unit	Fresh air	Treated air, exhaust, losses
Ventilation distribution subsystem	Treated air	Hot, heavy air, leakage (duct)
Exhaust/vent subsystem	Hot, heavy air	Hot, heavy air, leakage (duct)
Boiler subsystem	Heat, Boiler Feedwater	Hot water, combustion exhaust, water losses
Occupied space subsystem	People	People, leakage (building envelope), hot heavy air

4.2.2 Measurement nodes and specification

Figure 4-2 illustrates the process control mechanisms in the classroom spaces, with the sub-system assignments according to the asset management ontology indicated with broken line borders. The mass-energy flow here represents the people entering the room, with a “hold time” indicating the occupied duration. This figure had a high degree of utility in focusing the conversations around which measurements were important in answering the research question; regardless of the performance of the building ventilation system, or its theoretical capacity, the measurement of interest was the accumulation of CO_2 in that space.

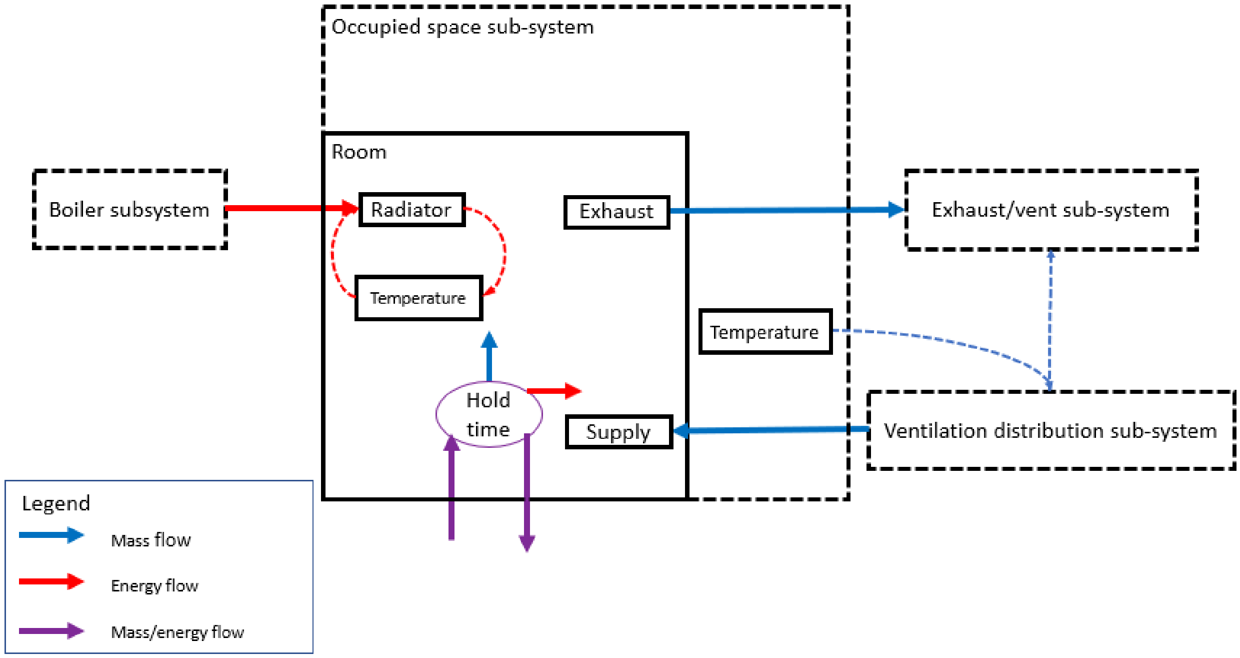


Figure 4-2 Process Control diagram of classroom, within the larger occupied space, mass and energy flows indicated

Deviceworx, a Canadian Internet of Things (IoT) company offered to assist in data collection with the generous loan of 30 air quality monitoring sensors to the school for a period of three weeks. The sensor capability and sampling frequency influenced the selection of measurement nodes. The Deviceworx units contained Sensiron CO_2 measurement elements (device id number SCD41-D-R2), applying photoacoustic non-dispersive infrared spectrometry (NDIR) (Palzer, 2020) with integrated temperature and humidity sensors. All CO_2 sensors underwent laboratory validation prior to assembly and the ventilation sensor assemblies were validated by Deviceworx prior to shipping. The complete manufacturer data sheet and manufacturer's calibration certification are included as supplemental materials, with the calibration certification. Table 4-2 summarizes the sensor specifications.

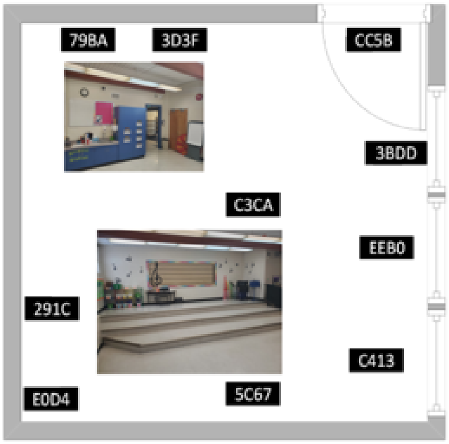
Table 4-2 Sensor specification and performance parameters, extracted from manufacturer's data sheet and product information. Complete Sensiron data sheet and product information is appended in the Supplemental Materials section

Parameter – CO_2 Sensing	Conditions	Value
CO2 output range	-	0 – 40,000 ppm
CO2 measurement accuracy	400 ppm – 5,000 ppm	± (40 ppm + 5% of reading)
CO2 repeatability	Typical	± 10 ppm
CO2 response time	τ63%, typical	60 s
Humidity measurement range	-	0 %RH – 100 %RH
Humidity accuracy (typ.)	15°C – 35°C, 20 %RH – 65 %RH	± 6% RH
	-10 °C – 60 °C, 0 %RH – 100%RH	± 9% RH
Humidity repeatability	Typical	± 0.4% RH
Humidity response time	τ63%, typical	90 s
Temperature measurement range	-	- 10°C – 60°C
Temperature accuracy (typ.)	15°C – 35°C	± 0.8°C
	-10°C – 60°C	± 1.5°C
Temperature repeatability	-	± 0.1°C
Default conditions of 25°C, 50 % RH, ambient pressure 1,013 mbar, default periodic measurement and 3.3 V supply voltage apply to values in the table above, unless otherwise stated.		

4.2.3 Specification of measurements

The use of CO_2 levels as a proxy for exhaled pathogens is well established in medicine and workplace safety (Cherrie et al., 2021; Du et al., 2019; van der Zee et al., 2017). One sensor was placed in the hallway of the school to act as a reference measurement. Deviceworx supplied 20 ventilation sensors, which sampled the temperature, relative humidity, and CO_2 levels every five minutes over a period of three weeks. While many ventilation studies plot air flow velocities to model air flows, and calculations using opening size and building tightness can be performed to assess the equivalent air changes per hour, these models do not necessarily provide accurate information about the amount of aerosol hazard in a space, particularly one that is fully occupied and contains furniture and partitions (Aguilar et al., 2022).

This study elected to use point measurements of CO_2 combined with relative humidity and temperature measurements at the same locations as a more direct proxy for the concentration of aerosol hazard in the space, and access to 8-10 sensors for each classroom allowed the inclusion of multiple heights and locations around the space. Sensors were mounted to ceiling and wall surfaces using 3M adhesive at entrances, near windows, near ventilation exhaust and supply nodes and at floor-level, breathing height and ceiling height. Figure 4-3 illustrates the placement of sensors relative to major room features, including major furnishings and layout. Detailed descriptions of sensor placement are included with the Device ID tags. Classroom 27 contained substantially more furnishings than the choral classroom, which does not contain any desks. Both classrooms have large banks of windows and at the time of the study, the aged windows and the air leakage from the window perimeters were likely a significant source of passive ventilation.



Classroom 36 – choral classroom

Device ID	Location description
3D3F	Near center vent
CC5B	Above door
79BA	Near left most vent
5C67	Rear most beam, center of room
291C	Cupboard, waist height.
C413	Mid window height rear most bulkhead
3BDD	Waist height, nearest bulkhead to door
C3CA	Central beam, near center of room
E0D4	Top step of riser, near floor
EEB0	Slightly lower, center bulkhead

Classroom 27 – traditional classroom

Device ID	Location description
93EA	Near left hand vent
F2BB	On window frame above heating unit, between 4th/5th window counted from left.
AC9C	Underside of front most beam, near center of room
5BE1	Side of teacher desk
EEB3	Underside of rearmost beam near windows
4B5A	Approximately 2' above floor, rear wall by sink
EB04	Cupboard door waist height
4F55	Chest height, side of cupboards over sink
CC99	Above door

Figure 4-3 Sensor placement and room layouts for Classrooms 36 and 27

Detailed sensor and app information are included as supplemental materials, the free xTag Explorer app downloads all readings from the last operating interval on each sensor through an encrypted Bluetooth connection and allows the user to save data points to a local file. Deviceworx supplied one tablet loaded with the xTag Explorer app that was used in the study, and the school supplied a second tablet that was easily configured to access sensor data and collect input from classroom teachers.

4.3 Analysis and results

The classroom activity had minimally observable impact on the CO_2 levels measured in the space. Students engaged in choral singing were exposed to similar concentrations of CO_2 as students in the traditional classroom, where desks were arranged in rows and public-health recommended distancing maintained.

The building mechanical configuration was identified on As-built construction drawings, validated in the field, by the research team. The building mechanical control system was accessed by the research team to evaluate the effect of fan speed on the temperature and CO_2 levels for one full occupied day June 8th, 2022. The control system display did not reflect the physical configuration of the mechanical equipment and the impact of changing fan speed through the control system had no measurable impact on the CO_2 levels in the classrooms under consideration. The building control system operated on a typical zone-based control scheme with hold temperatures in key locations set at the control panel. While individual rooms and offices contained heating elements and sometimes thermostats, the building temperature control was largely fixed and reverted to "automatic" after 1-2 hours of any locally initiated change. Figure 4-4 summarizes the temperature setpoint behavior measured at the hallway sensor over the three days showing the identical behavior between June 8th where fan speed was manipulated, and June 9th and 10th where the system operated without changes..

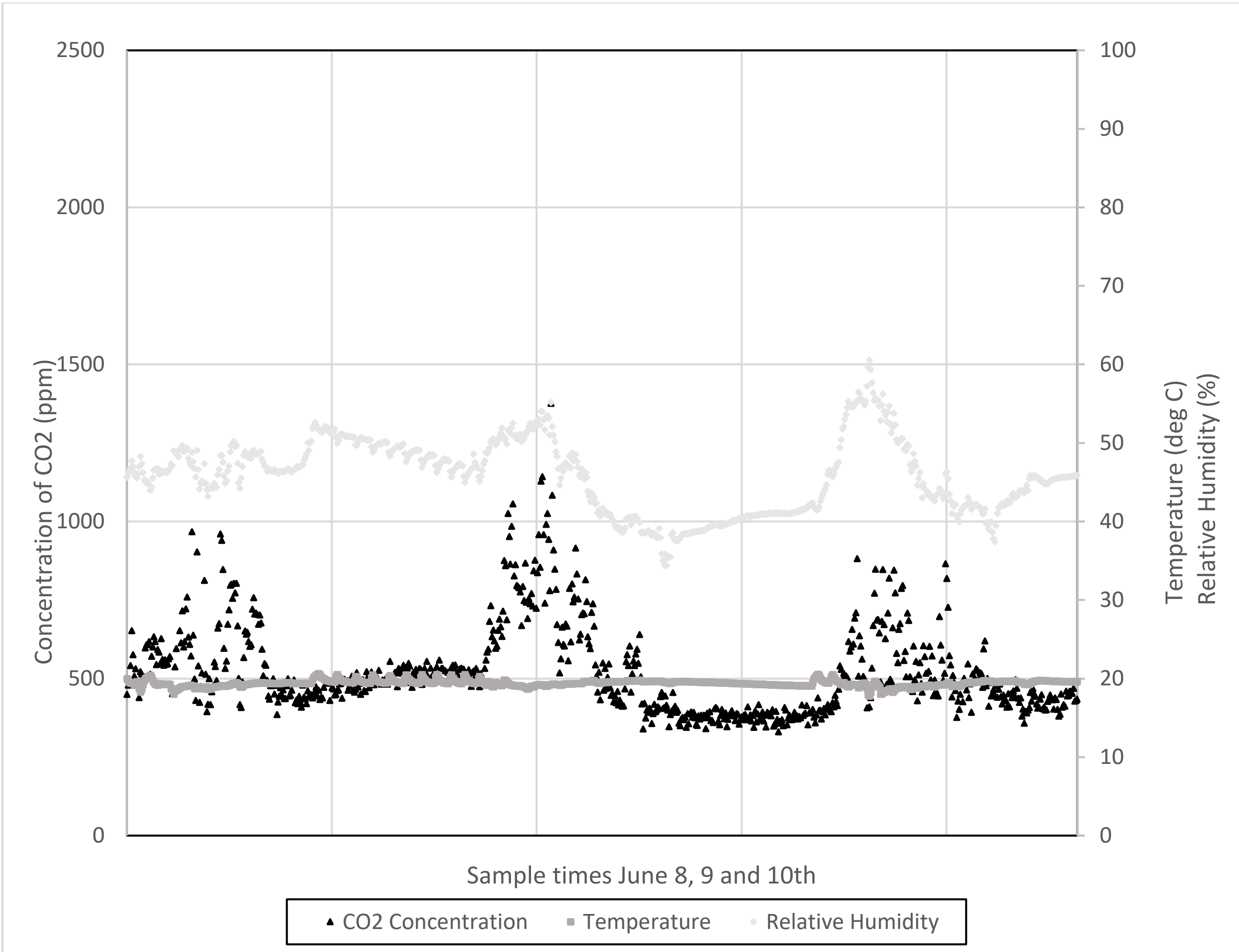


Figure 4-4 Detailed view of hallway sensor data when fan speed setpoints for Classroom 27 were manually set to 100%

The building has single-pane, wood framed windows and experiences typical temperature control challenges for aged public buildings, with overheating a problem in summer and difficulty retaining heat in winter. In May and June 2022, it was observable in both classrooms that passive ventilation via external windows and doors was significant and captured as “leakage” in the asset ontology. The study took place during the most challenging month for ventilation, in June which is the summer season. School is not in session during July and August, the hottest months of the year, and the temperature is typically cooler by September. The school does not have active air conditioning. The local weather was monitored by Environment Canada at the Blatchford weather station, and Figure 4-5 summarizes the outdoor conditions at the school over the study duration.

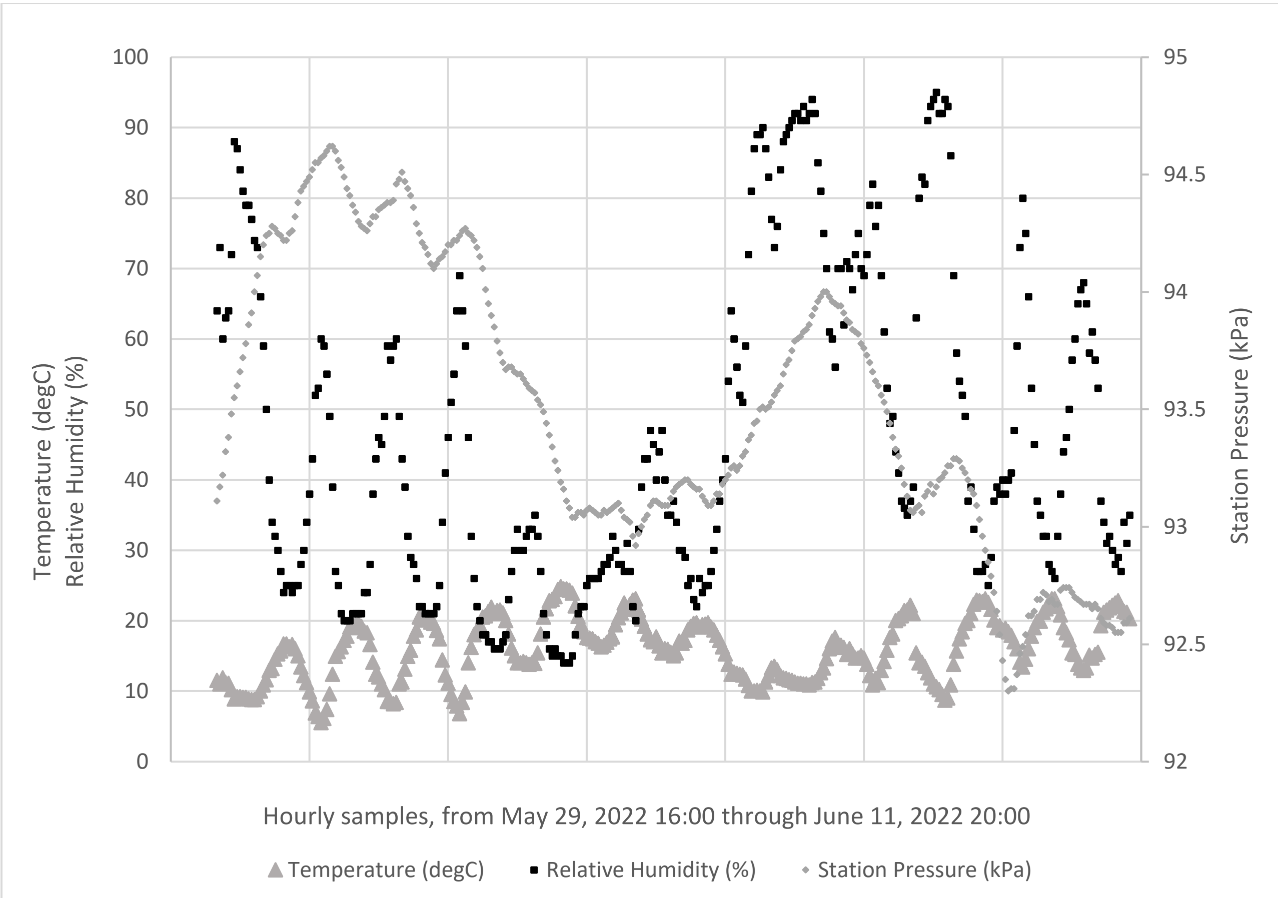


Figure 4-5 Extracted from the Environment and Climate Change Canada Historical Climate Data web site (https://climate.weather.gc.ca/index_e.html) on July 6, 2023. Climate data measured at the Edmonton Blatchford weather station located 6.6km from the school

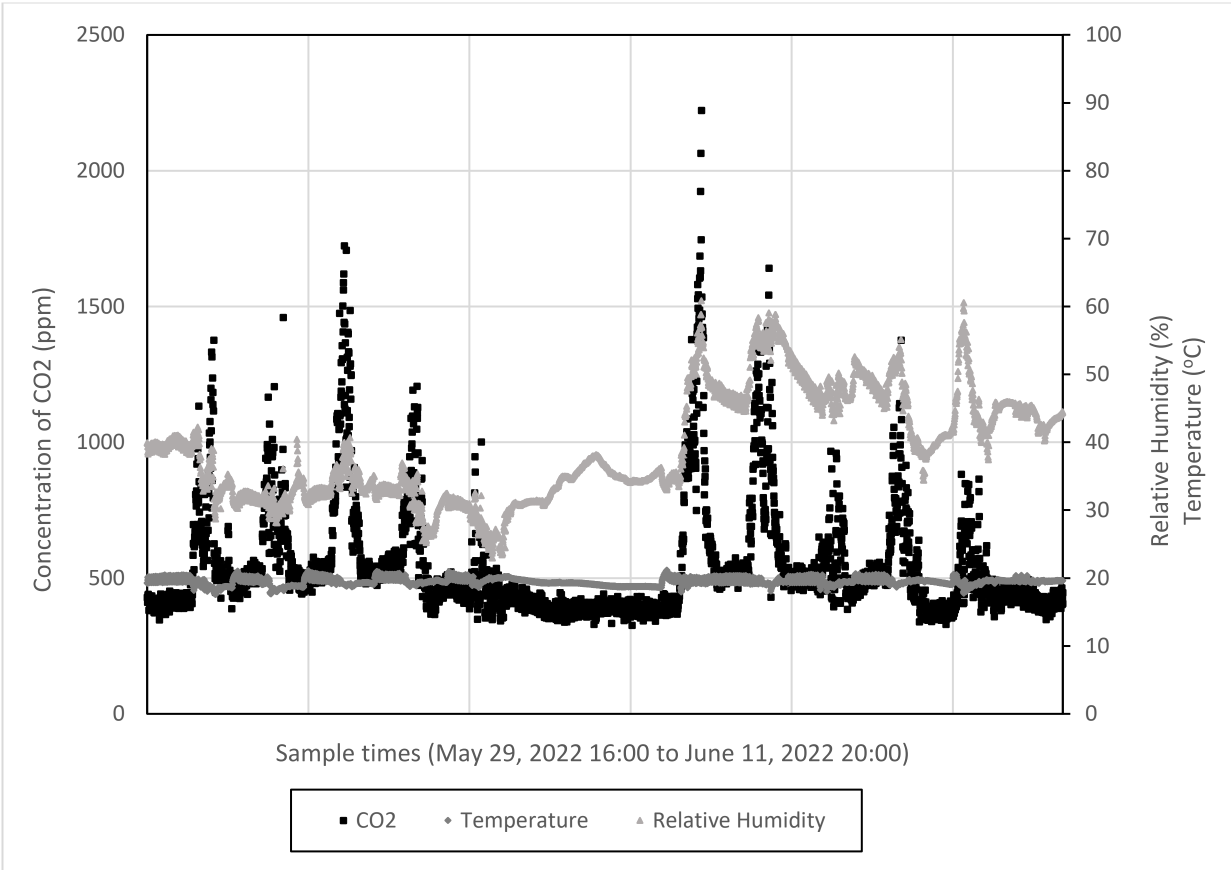


Figure 4-6 Sensor ID 5CDE located in the main hallway

The sensor data, summarized in Figure 4-6, Figure 4-7 and Figure 4-8 showed that mixing was more rapid and complete than was expected from assessment of the ventilation and operating parameters. In both classrooms, independent of the instructional activities in progress, two distinct operating states were observed. Once the classroom was occupied, and CO_2 levels cycled predictably upward until the room was vacated, when levels relatively rapidly returned to "unoccupied" levels. The reference sensor that was placed in the main hallway showed similar patterns.

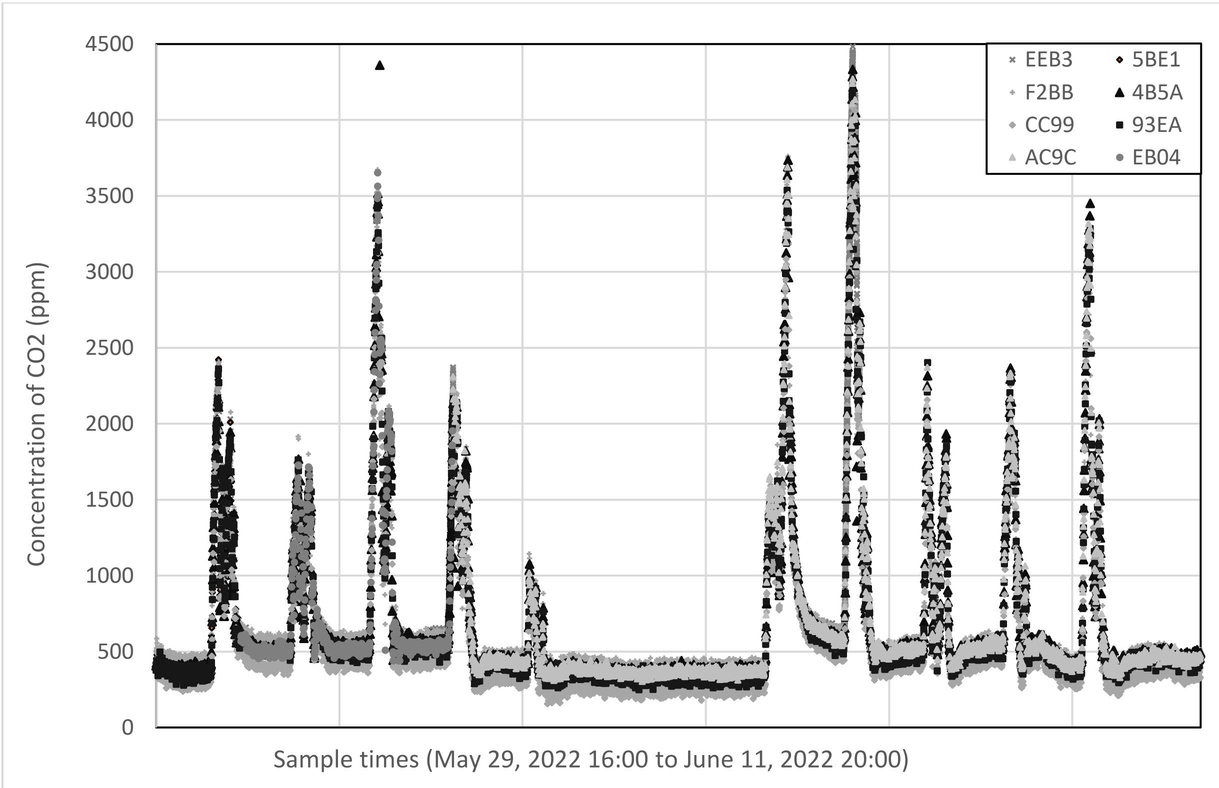


Figure 4-7 Classroom 27, traditional style grade 5 classroom over two weeks

4.3.1 Decision-making

After the first week of data collection, the school administration was curious about the findings and suggested several trials to be run to challenge the initial results. One of those trials included the manipulation of fan speeds, described above, to test a frequently-recommended risk control for public buildings and to examine the operation of the school's HVAC system. As shown in Figure 4-4, the increased fan speed had little impact on accumulated carbon dioxide levels, but it did cause overheating in the central office areas. The decision to return the system to automated control was supported by the information collected. To further test some of the often-recommended risk controls, administration developed a test around ensemble music. As instrumental music was still restricted under public health guidelines, adult volunteers played continuously in the otherwise vacant choral room to assess the change in CO_2 levels. Two volunteers simulated wind instrument playing by blowing through PVC tubing, seated directly beside two of the mounted sensors and the results were assessed. Using a sample interval of five minutes, the impact of these activities was not remarkable, and the choral classroom exhibited the typical cycling associated with any other period where the room was occupied. The public health restriction on instrumental music was lifted shortly after the completion of the study, but the results of the experiment reinforced the decision to stakeholders and increased the confidence in the school's decision to resume ensemble music classes for September 2022.

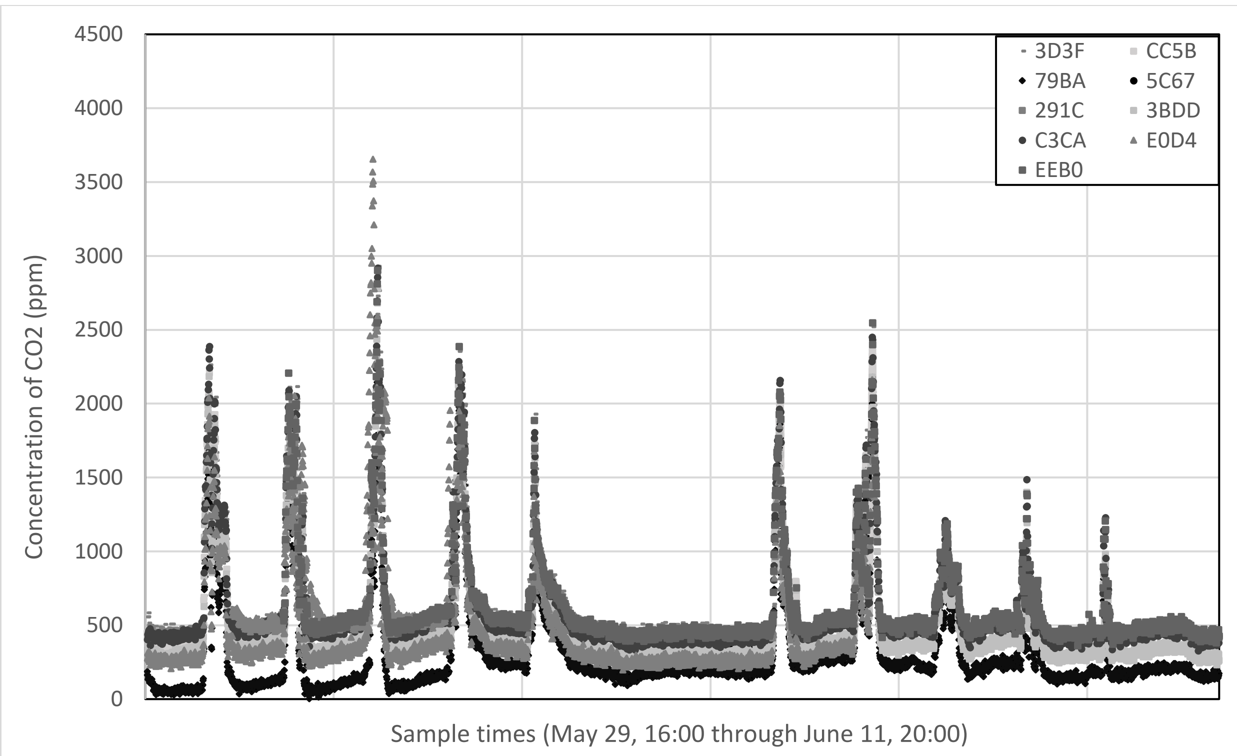


Figure 4-8 Concentration of CO2 in classroom 36, choral instruction over 2 weeks

4.4 Discussion

Cherrie et al., (2021) performed a systematic review of data collected between January 2020 and December 2020 (effectively the first year of the COVID-19 pandemic) and summarized the conclusions and limitations of studies documented in 35 papers. The results of their review identified many gaps in the risk controls recommended by regional and municipal public health authorities. Members of the public, particularly those accountable for the safety of members of the public and employees, continue to question the effectiveness of building ventilation and are actively seeking ways to assess and improve their facilities in terms of air quality and health. Few public resources are available that focus on empowering community leaders to implement effective risk management techniques when it comes to health and hygiene hazards (Shendell et al., 2021).

The sensor results collected in classrooms 27 and 36 shadow the trends of the reference sensor in the hallway, which demonstrates the cyclic nature of CO_2 concentrations in the entire facility, trending upward during “occupied” periods and predictably decreasing as the facility vacates. Use of ten sensors demonstrated more rapid and complete mixing of the air in the classroom than was expected, particularly in classroom 27, where there was more apparent congestion in the space due to furniture, shown in Figure 4-3. More detailed examination of daily results in both classroom spaces showed that CO_2 levels reached levels near “vacant” within 20 minutes of students leaving the classroom for recess and lunch periods, supporting the theory that vacating rooms and allowing “settle” time was a valid control for reducing exposure to exhaled contaminants in this facility, under these conditions. The return to average “vacant” levels is highly dependent on local factors, and use of multiple sensors over a relatively short test interval provided confidence in the length of “settle” time that would be effective. There was no significant increase in CO_2 generation in the choral music classroom, as compared to the traditional instruction classroom. Classroom 27 reached higher maximum concentrations of CO_2 than were observed in the choral classroom.

Shrestha et al., (2022) illustrate a related problem in building operation, which is that empirical data on exhaled contaminants simply does not match the design calculations and air flow models in all cases. They proposed alternative methods for calculating air changes following a field survey of naturally ventilated classrooms, which supports the approach of measuring locally to support field-based decisions. The relatively recent increased availability of high-quality, portable and relatively easy-to-use measurement devices suggests re-thinking some approaches to building operation and using more empirical data to establish process control strategies. Building on prior research performed in air quality assessment to develop facility-level study plans over short intervals can reduce reliance on airflow thresholds and traditional numerical models for risk assessment activities.

4.4.1 Evidence-based decision-making

The initial results of the risk management activities undertaken by the school administration in September 2020 demonstrated the value of basic industrial risk management theory and the layers of protection approach (LOPA), which was applied to evaluate the value of different non-mandatory controls. The school administration readily adopted the Risk Management process of "Identify-Assess-Plan-Control-Evaluate" and was enthusiastic to be able to "assess" the hazard in a more rigorous and direct manner than had been previously possible. The facility risk management plan had undergone several evaluation cycles and the school administration had made sensible alterations to risk controls when sufficient data was available to determine effectiveness. Practices like increased cleaning of classroom shared materials for example, was dropped as the public communication around COVID showed that the hazard was primarily aerosol. Some of the practices for sanitation impacted instructional activities, like requiring individually packaged and quarantined craft materials for primary school students as opposed to materials and tools that could be readily shared between groups of students.

The administration's response to the CO_2 study was equally enthusiastic and supported the decision to return choral singing to regular activity in the school. Public health restrictions on instrumental music were lifted before the 2022/2023 school year began, but school administration was prepared to request allowances from public health if restrictions had continued, based on the results of this study. There was no case to be made for instrumental music increasing the exposure risk in the occupied spaces of the facility; if traditional classrooms were safe to operate, then so too were group music lessons.

The school administration reported receiving 10-20 inquiries every week over the course of the COVID-19 pandemic, most attempting to sell devices promising to make schools safer. Products from hand sanitizer to air filtration units were offered, along with the increased communication from parents and staff, questioning different policies, and making suggestions to improve facility safety. Administration reported being generally overwhelmed by the sales inquiries but felt empowered by participating in the study to collect experimental data locally. Many of the risk controls around choral and instrumental music were the result of incident investigations from the early months of the COVID pandemic (Bauer et al., 2022; Nix et al., 2020), which represented risk scenarios that were simply not analogous to the day-to-day operation of an elementary/junior high school. Studies published in late 2021 (Schijven et al., 2021; Stockman et al., 2021) and 2022 (Bauer et al., 2022) have provided substantial additional theoretical and empirical basis for risk assessment, but at the time the risk controls were developed and rolled out, these resources were not available. This study demonstrates how research techniques can be mobilized and applied at the facility level to respond to new risks in the absence of large-scale studies. The layers of protection analysis has been a valuable tool in matching controls to risk scenarios, and identifying ways in which existing controls, like the practice of assessing students upon arrival for signs of illness, was offering facility-level protection that did not exist in the early incidents that had shaped the public health guidance. Public communication from community advisory boards and public health officials throughout the COVID-19 pandemic were often directive and paternalistic in tone, focusing on individual controls and emphasizing compliance behavior. The recommended controls varied widely between jurisdictions and, as a result, public confidence was predictably poor. This study demonstrates the effectiveness of basic, rigorous, and well-tested engineering practices to support risk management plans at the facility-level. With the widespread availability of both sensors and analytical capability, this type of study is accessible to facility owners and managers. The controls and protocols implemented in individual facilities should reflect the real consequences of the hazard in each facility itself, not the political decisions made at the regional level.

This study presents a re-usable method for facility-level studies that could foster increased public understanding and empower facility owners to implement realistic and effective risk management plans. The ontological breakdown used to communicate the building mechanical systems and simplify the movement and activity levels of students can be applied to any occupied facility used for any purpose and demonstrates the benefit of direct measurement over the challenges and expense of traditional ventilation models and projections. As part of a broader risk management plan, simple studies can be used to evaluate the effectiveness of controls and reduce uncertainty in both activity planning and spending decisions. The engineering approach summarized by Pan et al., (2021) was effective.

4.4.2 Limitations of the data set

There is considerable debate surrounding the "maximum" allowable level of CO_2 permitted in public spaces and workplaces. The American Society of Heating Refrigerating and Air-Conditioning Engineers (ASHRAE) issued a technical response to the Frequently Asked Question (FAQ) "What is the allowable level of carbon dioxide in an occupied space" (ASHRAE TC-04.03, ID 35) which provides qualified guidance to avoid concentrations exceeding 700ppm above the outdoor level. ASHRAE issued further guidance on this matter in a 2022 position paper (American Society of Heating Refrigeration and Air-Conditioning Engineers, 2022) summarizing the challenges and limitations of CO_2 as a single measurement of indoor air quality. In keeping with ASHRAE's positions on this topic, any potential risks associated with the levels recorded in this study are outside the scope of this analysis. Outdoor concentrations were not measured as part of this study, and further evaluations of indoor air quality would require measurement. The peak concentrations recorded in the classrooms during the months of May and June - particularly Classroom 27, where the single point maximum approached 4500ppm - may suggest room for improvement in ventilation, but this study was not designed to evaluate overall air quality or provide recommendations for improving building ventilation. The calibration performed on the sensors in this study does not allow extension of these data to that type of evaluation. It is noteworthy that the most recent ASHRAE standard for indoor air quality (American Society of Heating Refrigerating and Air-Conditioning Engineers & American National Standards Institute, 2022) does not establish a standard value for CO_2.

This study focused on comparing the CO_2 levels to a reference within the system, not benchmark values or exposure thresholds. Next steps would include extending the study to air quality and building health more broadly, which would require additional calibration and comparison to outdoor air. Combining the methods of previous works with the sensor technology and placement methodology could build on prior work (Hodgson et al., 2004; Rosbach et al., 2016; Shendell, Prill, et al., 2004; Shendell, Winer, et al., 2004) in classrooms and toward developing stronger policies for school building construction and operation. The objective of this study was to examine the marginal risk associated with choral classroom activities over traditional activities and the results in this school, during the study period supported the return to choral singing. It is important to note that these results are not transferable to other facilities nor broadly to any conclusions regarding air quality or building health in the school. Additionally, the role of passive ventilation in this school was noted, and further limits the generalizability of conclusions toward general air quality in the facility. Air egress at the window perimeters was observed with a portable anemometer and noted in risk management decisions, but similar spaces that do not experience the same degree of passive ventilation would not likely see CO_2 levels restored to reference ranges over the same time period.

4.5 Conclusions

The administrators and staff at the school adopted the principles and practice of risk management and provided valuable insight regarding the impact of different controls. They were able to communicate confidence in resuming both choral singing and later wind instrument ensemble classes after participating in this study, which demonstrated no significant change to CO_2 levels from choral singing as compared to traditional classroom activities. Further, the detailed examination of classroom data reinforced the intuition that vacating classrooms between courses, and at recess and lunch times reduced CO_2 levels effectively, promoting consideration of facility-use schedules to incorporate “people movement” as a ventilation lever. This increased confidence in school administration decisions around staggering break times and scheduling breaks throughout the day. The study data informed planning decisions for the 2022/2023 school year.

The COVID-19 pandemic may wane, but airborne hazards continue to be a global threat. Air quality assessments and empirical data collection can be feasibly performed at the facility level if research and industrial practices are made accessible. When region-wide policy statements are published, there can be gaps in understanding between the recommendation and the operationalization. Many studies seek to inform public policy or standard development with broad-based data collection, where this study sought to fill the opposite role, taking standard values and risk controls from public policy and operationalizing these concepts with data-driven decision-making. This study demonstrated that empirical data and thoughtful technical communication can demystify complex risk scenarios and empower building operators to make higher quality decisions for safe operation.

Funding

This research is funded by NSERC RGPIN-2018-05170-Lefsrud and NSERC COVID-2020-06431 Lefsrud "Developing a Safety Management Systems (SMS) Model".

Acknowledgements

The authors gratefully acknowledge the support of Deviceworx monitoring for donating the use of CO_2 monitors and Suzuki Charter School in Edmonton, Alberta, Canada for participating in this research.

Supplemental Materials

Deviceworx sensor data sheet

Sensiron product information and datasheet

Sensiron device calibration statement

Chapter 5

5.1 Conclusions

Decision making at human-process interfaces needs to be supported by the tools and aligned with operations goals to reduce the frequency of major loss incidents. The informatic tools that inform decision making are increasingly important in the causal networks that represent loss incidents, and ensuring that decision-makers can access current, accurate and relevant information during abnormal events is a design activity for engineers. To fully realize the latent value of information systems and historical data, these systems and data sources need to augment human decision-making at multiple organizational levels. These studies sought to examine the gaps between information systems and reality and interrogate how interactions with informatics in the process industry inform decision making.

The pilot study detailed in Chapter 2 applied qualitative research techniques to the evaluation of a design feature, causal maps visually representing a transfer entropy between variables in a process simulator. Participants were asked to narrate their decision-making process in diagnosing and responding to an abnormal event presented to them using a typical industrial DCS interface. Analyzing the results of that study revealed a gap in the techniques employed to evaluate operator effectiveness and led to the study presented in Chapter 3. The research question, "how can predictive analytics be integrated into interface designs to augment operator decision making" led to further questions. Participants in the pilot study made fewer incorrect causal inferences when using the interface augmentation, but the methodology employed to analyze the quality of their decisions, with and without the augmentation, was limited particularly from an industrial workplace applicability perspective. Collecting hours of narrative data for simulator testing is a significant data collection burden and the analysis of narrative test data is time consuming and difficult to compare between studies. A systematic assessment methodology, the situational design model, was adapted for industrial application. Building on prior work, particularly the dimensions of competence identified by Kluge et al., (2014) a scoring framework for operator performance during simulated optimization and fault diagnosis activities was developed. Recognition of stress-related influences on decision-making was reflected in the application of the NASA TLX tool during the initial pilot, but the concept of "Learning Readiness" was broadened for the second study to include psychosocial factors and individual self-assessment. The survey instrument applied in the large-scale study showed the limitations of that tool and challenges the widespread application of similar surveys, particularly in workplace applications.

The question "how can the situational design model be applied to assessing and developing operator decision making" led to the development of a study implementing the model with a group of chemical engineering students. The methodology demonstrated in Chapter 3 offered more dimensions of comparison for performance in the simulated task, and the application of a knowledge exam allowed participant results to be grouped into knowledge-related cohorts. This allowed improved disambiguation of the performance impact of the design feature, which could then be separated from the participant's individual knowledge level. The development of recurrent skills is recognized to be pre-requisite to the development of non-recurrent skills, and interactions with informatics, particularly simulators and other job supports serve to "teach" users about the mapped relationships between assets and the process. The results indicated that the knowledge exam was not sufficient to fully separate recurrent and non-recurrent skills, and a specific phase to introduce and test procedural knowledge for routine operations is recommended for improved utility. The design feature, causal maps represented visually, showed some promise in the initial pilot but did not contribute to improved decision making in the more detailed study.

Chapter 4 summarized the final risk management activity undertaken by the major stakeholders for a single commercial building facility as part of the COVID-19 response and demonstrated how asset management ontology knowledge bases can be applied to improve understanding of process systems in a risk management context. The carbon dioxide concentrations in classrooms were compared to evaluate the degree to which choral singing influenced the accumulation of airborne contaminants during occupied periods. The study results showed rapid mixing in the classrooms and showed no significant increase in accumulation due to choral singing, relative to the traditional instruction in the comparator classroom.

Facility information can be organized and presented using a variety of models; the asset management and process engineering ontologies employed as part of the COVID response strategy were effective in supporting industrial risk management approaches and in developing a facility-level study directly applicable to the risk-based decisions faced by the building operators. The application had active engagement and participation and the stakeholder groups were satisfied with the response. The ontology knowledge base was an important framework to ensure that decision-makers were talking about the same concept, asset or process flow using common terminology. Building operators gained confidence using the strategy and it was a valuable framework used to organize incoming questions, concerns, suggestions, and data variants. Facility operators reported overload, particularly with the volume of recommended products and suggestions from vendors and members of the public, that was alleviated by having a well-documented and detailed risk management plan in place. The asset management ontology employed to deconstruct the system and select measurement points proved to be an effective communication tool during the risk assessment and management processes.

The criteria that are used to evaluate the effectiveness of decision-support systems are interconnected and multi-variate (Adhitya et al., 2014; Doe et al., 2005; Golightly et al., 2018; Grozdanovic & Bijelić, 2019; Ikuma, Koffskey, et al., 2014) and continued synthesis of the measurement techniques and analytic methodologies of researchers in several focus areas will allow improved applicability to risk management activities. Applying the situational design model more effectively to the evaluation of operator understanding first requires a more thorough examination of evaluation techniques for recurrent skills in operation (van de Merwe et al., 2022), as well as increased focus on the task design for the simulator-based portion of the evaluation, ensuring that the features available in the simulator interface meet the intent of the task design. Vicente et al., (1998) demonstrated the balance that must be considered in implementing new technology, particularly in the control room, and the impact of having excessive degrees of freedom, one of the common trade-offs associated with increased display customization. Rajabiyazdi and Jamieson, (2020) summarized techniques evaluating transparency - how much information about underlying systems is optimal to communicate to operators - in a comparison of four models evaluating human-automation interaction. The type and relative complexity of the information necessary to solve operating problems is at least in part state dependent (Roth et al., 1997), and continuing to synthesize the results of prior studies will allow improved design and application of new technology. Balancing excessive degrees of freedom and increased interface complexity with the need for transparency is an important consideration in the design and delivery of new products but also in the design of experimental tasks. Measuring the direct improvements to understanding, through a systematic and well-designed evaluation of recurrent and non-recurrent skills can allow for more granular and focused analysis of the results of workplace studies, both practical and simulated. Multi-dimensional results of performance related studies can be more readily applied by designers and engineers in targeting feature selection for field applications.

Determining the optimal degree of customization in informatic systems includes consideration for collective decision-making, and how shared information sources influence those interactions (D. Chang & Chen, 2015; Goldstone & Janssen, 2005; Kortschot et al., 2018; Trappey et al., 2015). Multi-agent decision-making and collective cognition are further dimensions of interest in this area, as the group interactions in the Chapter 3 study were surprising and of relevance to the evaluation of participant performance. The evaluation of the dynamics of high performing teams (Engome Tchupo & Macht, 2023; Yim & Seong, 2015; Zenk et al., 2010) parallel current trends in artificial intelligence research examining multi-agent evaluations (Ferber, 1999; Seng & Srinivasan, 2010), and the critical role interpersonal communication plays in understanding complex systems is an important element to consider for industrial applications, where individuals seldom work in real isolation from interpersonal and institutional decision-making influences. In pursuing further refinement of the situational design model, the task design for evaluation of both recurrent and non-recurrent skills is an area where collective decision-making and task distribution could be incorporated to better represent real operating environments. The learning readiness dimension is also likely influenced by team dynamics and expanding and refining the "willingness" assessment to include psychosocial factors more explicitly would be a valuable integration of research from many disciplines.

These studies were motivated by two broad questions, why major loss incidents continue to occur in the process industry despite technological advances and why those technological advances do not seem to be achieving their potential in industry. In exploring these questions, an opportunity in assessment was identified to better support engineers and innovators, in establishing the real benefit of the information systems and automation structures they offer for use in industry.

5.2 Limitations

The work detailed in this dissertation was necessarily time-limited, and the development of the instruments used in the situational design study is an area for immediate improvement. Due to time constraints, the knowledge exam was assembled using publicly question banks and developed using a parallel format to the pilot experiment (Zhong & Xu, 2019) but the exam was not broadly tested to evaluate utility. Similarly, the learning readiness exam was limited to "self assessment" of knowledge and experience, but an evaluation of task loading would have provided additional dimensions for comparison and the pilot study in Chapter 2 demonstrated the utility of the NASA Task Load Index in eliciting meaningful participant responses. The marking rubrics provided to the three markers were developed independently and focused on the complexity of the causal relationships identified but failed to adequately capture the accuracy of the causal relationship identified. The evaluation structure for the problem-solving element needs additional refinement, but further developing the dimensions of complexity identified by (Kluge et al., 2014) for process operators shows promise as a framework to apply in a structured assessment.

The ontology knowledge bases leveraged throughout the work, including the one assembled as a communication and knowledge transfer tool for the school administration and stakeholders in Chapter 4 remains a combination of published methods and original adaptations. Continuing to pursue a rigorously defined and publicly available asset management ontology knowledge base would have tremendous value in standardizing approaches across industries and allowing the results of risk management efforts to be more directly compared.

5.3 Recommendations for future studies

Further investigating the ways informatics influence risk-based decision making requires additional refinement of the evaluation methodology presented here, but the structure and analytic techniques show promise in not only evaluating the quality of decisions but developing supportive activities for training and improved informatic design. The following are suggestions for the direction of future studies and work:

- Deconstructing effective training programs and classifying learning elements according to common dimensions of competence would provide a valuable resource to industrial producers and allow the effective elements to be identified and replicated.
- Mapping the information accessed and utilized in risk-based decisions outside operations would be of value. Risk-based decisions are made by more workers than panel operators, and informatics support those roles, while significant "noise" reaches decision-makers from a variety of sources.
- Repositioning "human factors" to the axis of "learning readiness" in the situational design model, and more consistent recognition of the differences between "recurrent" and "non-recurrent" skills in training and assessment would provide more opportunities for effective assessment of human decision-making performance in the process industry.
- Risk-based decisions are not limited to selecting courses of action during abnormal events. Decision-making or problem-solving skills are employed in root cause analysis

and incident reporting. Compare the accuracy and utility of incident report inputs using OKB system breakdowns and structured prompts versus traditional input formats.

- Further exploration of the connections between human cognition and machine learning/artificial intelligence, in particular comparisons for analytic performance using ontology-based data storage versus existing typical storage structures.
- Detailed examination of human performance literature and team performance in particular, and further refinement of the assessment methodology to examine the roles of collective cognition and group interaction

List of References

Adhitya, A., Cheng, S. F., Lee, Z., & Srinivasan, R. (2014). Quantifying the effectiveness of an alarm management system through human factors studies. *Computers and Chemical Engineering*, *67*, 1–12. https://doi.org/10.1016/j.compchemeng.2014.03.013

Agar, N. (2014a). Truly Human Enhancement. In *Truly Human Enhancement*. https://doi.org/10.7551/mitpress/9780262026635.001.0001

Agar, N. (2014b). Truly Human Enhancement. In *Truly Human Enhancement*. The MIT Press. https://doi.org/10.7551/mitpress/9780262026635.001.0001

Aguilar, A. J., de la Hoz-Torres, M. L., Costa, N., Arezes, P., Martínez-Aires, M. D., & Ruiz, D. P. (2022). Assessment of ventilation rates inside educational buildings in Southwestern Europe: Analysis of implemented strategic measures. *Journal of Building Engineering*, *51*. https://doi.org/10.1016/j.jobe.2022.104204

Albert, A., & Hallowel, M. R. (2013). Revamping occupational safety and health training: Integrating andragogical principles for the adult learner. *Australasian Journal of Construction Economics and Building, Vol 13, Iss 3, Pp 128-140 (2013) VO - 13*, *3*, 128. https://doi.org/10.5130/ajceb.v13i3.3178

American Society of Heating Refrigerating and Air-Conditioning Engineers, & American National Standards Institute. (2022). Ventilation and Acceptable Indoor Air Quality. In *ANSI/ASHRAE Standard 62.1-2022*. www.ashrae.org/technology.

American Society of Heating Refrigeration and Air-Conditioning Engineers. (2022). *ASHRAE Position Document on Indoor Carbon Dioxide*. www.ashrae.org

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Baker, K. E., & Lefsrud, L. (2019). Organizational myopia: how organizations create complexity, ambiguity, and uncertainty to blind their risk management efforts. *Society for Risk Analysis (SRA) Benelux Conference*.

Baker, K. E., Macciotta, R., Hendry, M. T., & Lefsrud, L. M. (2020). Combining safety approaches to bring hazards into focus: An oil sands tailings case study. *Canadian Journal of Chemical Engineering*.

Basoglu, N., Daim, T., & Kerimoglu, O. (2007). Organizational adoption of enterprise resource planning systems: A conceptual framework. *Journal of High Technology Management Research*, *18*(1), 73–97. https://doi.org/10.1016/j.hitech.2007.03.005

Bathelt, A., Ricker, N. L., & Jelali, M. (2015). Revision of the Tennessee eastman process model. *IFAC-PapersOnLine*, *28*(8), 309–314. https://doi.org/10.1016/j.ifacol.2015.08.199

Bauer, K., Hardege, R., Neumann, S., Schwarze, R., Fuchs, M., & Heinrich Pieper, L. (2022). How Safe is Singing Under Pandemic Conditions? - CO2-Measurements as Simple Method for Risk Estimation During Choir Rehearsals. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2022.05.003

Baybutt, P. (2016). Insights into process safety incidents from an analysis of CSB investigations. *Journal of Loss Prevention in the Process Industries*, *43*, 537–548. https://doi.org/10.1016/j.jlp.2016.07.002

Berry, G., Parsons, A., Morgan, M., Rickert, J., & Cho, H. (2022). A review of methods to reduce the probability of the airborne spread of COVID-19 in ventilation systems and enclosed spaces. *Environmental Research*, *203*. https://doi.org/10.1016/j.envres.2021.111765

Blunden, H., Logg, J. M., Brooks, A. W., John, L. K., & Gino, F. (2019). Seeker beware: The interpersonal costs of ignoring advice. *Organizational Behavior and Human Decision Processes*, *150*, 83–100. https://doi.org/10.1016/J.OBHDP.2018.12.002

Brambilla, S., & Manca, D. (2011). Recommended features of an industrial accident simulator for the training of operators. *Journal of Loss Prevention in the Process Industries*, *24*(4), 344–355. https://doi.org/10.1016/j.jlp.2011.01.009

Bruch, E., & Feinberg, F. (2017). Decision-Making Processes in Social Contexts. *Annual Review Sociology*, *43*, 207–227. https://doi.org/doi:10.1146/annurev-soc-060116-053622

Buddaraju, D. (2011). *Performance of control room operators in alarm management* [Louisiana State University and Agricultural and Mechanical College]. https://digitalcommons.lsu.edu/gradschool_theses

Bullemer, B. P. T., Reising, D. V. C., Tolsma, M., & Laberge, J. C. (2014). Plant Assesses Alarm Displays. *Chemical Processing Progress*.

Bullemer, P. T. (2013). Properly Handle Abnormal Situations. *Chemical Processing Progress*.

Bullemer, P. T., & Nimmo, I. (1998). Tackle abnormal situation management with better training. *Chemical Engineering Progress*, *94*(1), 43–54.

Bullemer, P. T., & Reising, D. V. C. (2015). Managing Human Reliability: An Abnormal Situation Management Historical Perspective. *Mary Kay O'Connor Process Safety Center*

2015 International Symposium, 30. http://www.asmconsortium.net/Documents/Managing Human Reliability_MKOC15_Bullemer and Reising_Final.pdf

Cameron, D., Clausen, C., & Morton, W. (2002). *Dynamic Simulators for Operator Training* (pp. 393–431). https://doi.org/10.1016/S1570-7946(02)80019-0

Cameron, I., Mannan, S., Németh, E., Park, S., Pasman, H., Rogers, W., & Seligmann, B. (2017). Process hazard analysis, hazard identification and scenario definition: Are the conventional tools sufficient, or should and can we do much better? *Process Safety and Environmental Protection*, *110*, 53–70. https://doi.org/10.1016/j.psep.2017.01.025

Castillo-Borja, F., Vázquez-Román, R., Quiroz-Pérez, E., Díaz-Ovalle, C., & Mannan, M. S. (2017). A resilience index for process safety analysis. *Journal of Loss Prevention in the Process Industries*, *50*, 184–189. https://doi.org/10.1016/J.JLP.2017.06.017

Chang, C. C., & Johnson, T. (2020). Integrating heuristics and think-aloud approach to evaluate the usability of game-based learning material. *Journal of Computers in Education*, *8*(1), 137–157. https://doi.org/10.1007/s40692-020-00174-5

Cheng, H., Wu, J., Liu, Y., & Huang, D. (2019). A novel fault identification and root-causality analysis of incipient faults with applications to wastewater treatment processes. *Chemometrics and Intelligent Laboratory Systems*, *188*(March), 24–36. https://doi.org/10.1016/j.chemolab.2019.03.004

Cherrie, J. W., Cherrie, M. P. C., Smith, A., Holmes, D., Semple, S., Steinle, S., Macdonald, E., Moore, G., & Loh, M. (2021). Contamination of Air and Surfaces in Workplaces with SARS-CoV-2 Virus: A Systematic Review. *Annals of Work Exposures and Health*, *65*(8), 879–892. https://doi.org/10.1093/annweh/wxab026

Chiang, I.-C. A., Jhangiani, R. S., & Price, P. C. (2015). Single-Subject Research. In *Research Methods in Psychology* (2nd Canadi). Opentext BC.

Chu, F., Fu, Y., & Liu, S. (2019). Organization is also a “life form”: Organizational-level personality, job satisfaction, and safety performance of high-speed rail operators. *Accident Analysis & Prevention*, *125*, 217–223. https://doi.org/10.1016/J.AAP.2019.01.027

Colombo, S., & Golzio, L. (2016). The Plant Simulator as viable means to prevent and manage risk through competencies management : Experiment results. *Safety Science*, *84*, 46–56. https://doi.org/10.1016/j.ssci.2015.11.021

Dahiya, R., & Dahiya, P. (2018). Why do we keep repeating major accidents? *Loss Prevention Bulletin*, *259*, 19–23. https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=129148608&site=ehost-live

Dai, Y., Wang, H., Khan, F., Zhao, J., El-Halwagi, M., & Ng, K. M. (2016). Abnormal situation management for smart chemical process operation. *Current Opinion in Chemical Engineering*, *14*, 49–55. https://doi.org/10.1016/j.coche.2016.07.009

Daisey, J. M., Angell, W. J., & Apte, M. G. (2003). Indoor air quality, ventilation and health symptoms in schools: An analysis of existing information. In *Indoor Air* (Vol. 13, Issue 1, pp. 53–64). https://doi.org/10.1034/j.1600-0668.2003.00153.x

Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. *Computers in Industry*, *123*, 103298. https://doi.org/10.1016/j.compind.2020.103298

Das, L., Iqbal, M. U., Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2018). Toward Preventing Accidents in Process Industries by Inferring the Cognitive State of Control Room Operators through Eye Tracking. *ACS Sustainable Chemistry and Engineering*, *6*(2), 2517–2528. https://doi.org/10.1021/acssuschemeng.7b03971

Das, L., Srinivasan, B., & Srinivasan, R. (2017). Cognitive Behavior Based Framework for Operator Learning : Knowledge and Capability Assessment through Eye Tracking. In *27th European Symposium on Computer Aided Process Engineering* (Vol. 40). Elsevier Masson SAS. https://doi.org/10.1016/B978-0-444-63965-3.50498-0

Delahaye, B. L., & Smith, H. E. (1995). The validity of the learning preference assessment. *Adult Education Quarterly*, *45*(3), 159–173.

Dindar, M., Järvelä, S., & Järvenoja, H. (2020). Interplay of metacognitive experiences and performance in collaborative problem solving. *Computers and Education*, *154*. https://doi.org/10.1016/j.compedu.2020.103922

Doe, K., Bloom, C., Reising, D. V. C., Bullemer, P., Coppard, D., DeMaere, T., & Errington, J. (2005). Establishing Human Performance Improvements and Economic Benefit for a Human-Centered Operator Interface: An Industrial Evaluation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(23), 2036–2040. https://doi.org/10.1177/154193120504902308

Downs, J. J., & Vogel, E. F. (1993). A plant-wide industrial process control problem. *Computers and Chemical Engineering*, *17*(3), 245–255. https://doi.org/10.1016/0098-1354(93)80018-I

Drever, J., Naughton, H., Nagel, M., Löhr, A., & Buderath, M. (n.d.). *Implementing MIMOSA Standards*.

Du, C.-R., Wang, S.-C., Yu, M.-C., Chiu, T.-F., Wang, J.-Y., Chuang, P.-C., Jou, R., Chan, P.-C., & Fang, C.-T. (2019). Effect of ventilation improvement during a tuberculosis outbreak in underventilated university buildings. *Indoor Air*, *30*, 422–432.

Duan, P., Yang, F., Chen, T., & Shah, S. L. (2013). Direct Causality Detection via the Transfer Entropy Approach. *IEEE Transactions on Control Systems Technology*, *21*(6), 2052–2066. https://doi.org/10.1109/TCST.2012.2233476

Durand-Moreau, Q., Mackenzie, G., Adisesh, A., Straube, S., Chan, X. H. S., Zelyas, N., & Greenhalgh, T. (2021). Twitter Analytics to Inform Provisional Guidance for COVID-19 Challenges in the Meatpacking Industry. In *Annals of Work Exposures and Health* (Vol. 65, Issue 4, pp. 373–376). Oxford University Press. https://doi.org/10.1093/annweh/wxaa123

Eizenberg, S., Shacham, M., & Brauner, N. (2006). Combining HAZOP with dynamic simulation-Applications for safety education. *Journal of Loss Prevention in the Process Industries*, *19*(6), 754–761. https://doi.org/10.1016/j.jlp.2006.07.002

Eljack, F., & Kazi, M. K. (2016). Process safety and abnormal situation management. *Current Opinion in Chemical Engineering*, *14*, 35–41. https://doi.org/10.1016/j.coche.2016.07.004

Elsawah, S., McLucas, A., & Mazanov, J. (2017). An empirical investigation into the learning effects of management flight simulators: A mental models approach. *European Journal of Operational Research*, *259*(1), 262–272. https://doi.org/10.1016/j.ejor.2016.10.011

Engome Tchupo, D., & Macht, G. A. (2023). Entropy for team communication pattern recognition. *Applied Ergonomics*, *111*. https://doi.org/10.1016/j.apergo.2023.104038

Fass-Holmes, B. (2022). Survey Fatigue-What Is Its Role in Undergraduates' Survey Participation and Response Rates? *Journal of Interdisciplinary Studies in Education*, *11*(1), 56–73. https://ojed.org/jise

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*.

Ferrari, S., Blázquez, T., Cardelli, R., Puglisi, G., Suárez, R., & Mazzarella, L. (2022). Ventilation strategies to reduce airborne transmission of viruses in classrooms: A systematic review of scientific literature. In *Building and Environment* (Vol. 222). Elsevier Ltd. https://doi.org/10.1016/j.buildenv.2022.109366

Gehman, J., Lefsrud, L., Lounsbury, M., & Lu, C. (2016). Perspectives on energy and environment risks with implications for canadian energy development. *Bulletin of Canadian Petroleum Geology*, *64*(2), 384–388. https://doi.org/10.2113/gscpgbull.64.2.384

Gerbec, M. (2013). Supporting organizational learning by comparing activities and outcomes of the safety-management system. *Journal of Loss Prevention in the Process Industries*, *26*(6), 1113–1127. https://doi.org/10.1016/j.jlp.2013.04.009

Gharahbagheri, H., Imtiaz, S., & Khan, F. (2017). Combination of KPCA and causality analysis for root cause diagnosis of industrial process fault. *Canadian Journal of Chemical Engineering*, *95*(8), 1497–1509. https://doi.org/10.1002/cjce.22852

Glaser, B. (2016). The Grounded Theory perspective: Its origins and growth. *Grounded Theory Review*, *15*(1), 4–9. http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=116403201&site=ehost-live

Godish, T. (1996). Relationships between ventilation and indoor air quality: A review. *Indoor Air*, *6*(2), 135–145. https://doi.org/10.1111/j.1600-0668.1996.00010.x

Golightly, D., & Dadashi, N. (2016). How do principles for human-centred automation apply to disruption management decision support? *2016 IEEE International Conference on Intelligent Rail Transportation, ICIRT 2016*, *285243*, 150–156. https://doi.org/10.1109/ICIRT.2016.7588725

Golightly, D., Kefalidou, G., & Sharples, S. (2018). A cross-sector analysis of human and organisational factors in the deployment of data-driven predictive maintenance. *Information Systems and E-Business Management*, *16*(3), 627–648. https://doi.org/10.1007/s10257-017-0343-1

Greenhalgh, T. (2021). *Miasmas , mental models and preventive public health : some philosophical reflections on science in the COVID-19 pandemic*.

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic Problem Solving: A New Assessment Perspective. *Applied Psychological Measurement*, *36*(3), 189–213. https://doi.org/10.1177/0146621612439620

Grozdanovic, M., & Bijelić, B. (2019). Impact of Human, workplace and indoor environmental risk factors on operator’s reliability in control rooms. *Human and Ecological Risk Assessment*. https://doi.org/10.1080/10807039.2018.1501659

Gruber, T. R. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *In International Journal Human-Computer Studies*, *43*, 907–928.

Haji, F. A., Cheung, J. J. H., Woods, N., Regehr, G., de Ribaupierre, S., & Dubrowski, A. (2016). Thrive or overload? The effect of task complexity on novices' simulation-based learning. *Medical Education*, *50*(9), 955–968. https://doi.org/10.1111/medu.13086

Halim, S. Z., & Mannan, M. S. (2018). A journey to excellence in process safety management. *Journal of Loss Prevention in the Process Industries*, *55*(April), 71–79. https://doi.org/10.1016/j.jlp.2018.06.002

Hanham, J., Leahy, W., & Sweller, J. (2017). Cognitive Load Theory, Element Interactivity, and the Testing and Reverse Testing Effects. *Applied Cognitive Psychology*, *31*(3), 265–280. https://doi.org/10.1002/acp.3324

Hart, S. (2006). NASA TLX - Task Load Index; 20 years later. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, 904–908.

Hart, S., & Staveland, L. (1988). Development of NASA-TLX: Results of empirical and theoretical workload. In *Human mental workload*. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16243365

Hendershot, D. C. (2011). Process safety: Natural disasters and process safety incidents - Some common perspectives. *Journal of Chemical Health and Safety*, *18*(2), 30–31. https://doi.org/10.1016/j.jchas.2011.01.007

Hodgson, A. T., Shendell, D. G., Fisk, W. J., & Apte, M. G. (2004). Comparison of predicted and derived measures of volatile organic compounds inside four new relocatable classrooms. *Indoor Air, Supplement*, *14*(8), 135–144. https://doi.org/10.1111/j.1600-0668.2004.00315.x

Hu, J., Zhang, L., Ma, L., & Liang, W. (2010). An integrated method for safety pre-warning of complex system. *Safety Science*, *48*(5), 580–597. https://doi.org/10.1016/j.ssci.2010.01.007

Huang, J., Jones, P., & He, X. (2022). Masks, ventilation and exposure time: A web-based calculator of indoor COVID-19 infection risk. *Frontiers in Built Environment*, *8*. https://doi.org/10.3389/fbuil.2022.986923

Ikuma, L. H., Harvey, C., Taylor, C. F., & Handal, C. (2014). A guide for assessing control room operator performance using speed and accuracy, perceived workload, situation awareness, and eye tracking. *Journal of Loss Prevention in the Process Industries*, *32*, 454–465. https://doi.org/10.1016/j.jlp.2014.11.001

Ikuma, L. H., Koffskey, C., & Harvey, C. M. (2014). A Human Factors-Based Assessment Framework for Evaluating Performance in Control Room Interface Design. *IIE Transactions on Occupational Ergonomics and Human Factors*, *2*(3–4), 194–206. https://doi.org/10.1080/21577323.2014.990589

Jain, P., Pasman, H. J., Waldram, S., Pistikopoulos, E. N., & Mannan, M. S. (2018). Process Resilience Analysis Framework (PRAF): A systems approach for improved risk and safety management. *Journal of Loss Prevention in the Process Industries*, *53*, 61–73. https://doi.org/10.1016/J.JLP.2017.08.006

Jennings, M. (2020). The oil and gas industry, the competence assessment of Offshore Installation Managers (OIMs) and Control Room Operators (CROs) in emergency response, and the lack of effective assessment of underpinning technical knowledge and understanding. *Journal of Loss Prevention in the Process Industries*, *65*. https://doi.org/10.1016/j.jlp.2020.104090

Kanse, L., Parkes, K., Hodkiewicz, M., Hu, X., & Griffin, M. (2018). Are you sure you want me to follow this? A study of procedure management, user perceptions and compliance behaviour. *Safety Science*, *101*(August 2016), 19–32. https://doi.org/10.1016/j.ssci.2017.08.003

Kennedy, D. A., & Whittaker, J. (2006). *A Renewed Focus on Procedures : Conformance versus Innovation*.

Kennedy, D., & Whittaker, J. (2000). Engineering procedures manuals—benefit or liability? *EMJ - Engineering Management Journal*, *12*(1), 9–14. https://doi.org/10.1080/10429247.2000.11415058

Klarborg, B., Lahrmann, H., Nielsagerholm, Tradisauskas, N., & Harms, L. (2012). Intelligent speed adaptation as an assistive device for drivers with acquired brain injury: A single-case field experiment. *Accident Analysis and Prevention*, *48*, 57–62. https://doi.org/10.1016/j.aap.2011.05.004

Kluge, A. (2008). Performance assessments with microworlds and their difficulty. In *Applied Psychological Measurement* (Vol. 32, Issue 2, pp. 156–180). https://doi.org/10.1177/0146621607300015

Kluge, A., Nazir, S., & Manca, D. (2014). Advanced Applications in Process Control and Training Needs of Field and Control Room Operators. *IIE Transactions on Occupational Ergonomics and Human Factors*, *2*, 121–136. https://doi.org/10.1080/21577323.2014.920437

Kutsch, E., & Hall, M. (2010). Deliberate ignorance in project risk management. *International Journal of Project Management*, *28*(3), 245–255. https://doi.org/10.1016/j.ijproman.2009.05.003

Landman, R., & Jämsä-Jounela, S. L. (2016). Hybrid approach to casual analysis on a complex industrial system based on transfer entropy in conjunction with process connectivity information. *Control Engineering Practice*, *53*, 14–23. https://doi.org/10.1016/j.conengprac.2016.04.010

Lapkin, A. A., Voutchkova, A., & Anastas, P. (2011). A conceptual framework for description of complexity in intensive chemical processes. *Chemical Engineering and Processing: Process Intensification*, *50*(10), 1027–1034. https://doi.org/10.1016/j.cep.2011.06.005

Lee, C. B. (2010). The interactions between problem solving and conceptual change: System dynamic modelling as a platform for learning. *Computers and Education*, *55*(3), 1145–1158. https://doi.org/10.1016/j.compedu.2010.05.012

Lee, C. B., Hanham, J., & Leppink, J. (2019). Instructional Design Principles for High-Stakes Problem-Solving Environments. In *Instructional Design Principles for High-Stakes Problem-Solving Environments*. Springer Singapore. https://doi.org/10.1007/978-981-13-2808-4

Lee, S., Jeong, I., & Il, M. (2000). Development of evaluation algorithms for operator training system. *Computers & Chemical Engineering*, *24*(2–7), 1517–1522. https://doi.org/10.1016/S0098-1354(00)00557-3

Lee, S. T., Kim, S. Y., & Gilmore, D. (2017). Human-In-The-Loop Evaluation of Human-Machine Interface for Power Plant Operators. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 34–39. https://doi.org/10.0/Linux-x86_64

Lefsrud, L. (2019a). *Lecture Notes CME 694 : Advanced Topics in Quantitative Risk Analysis*.

Lefsrud, L. (2019b). Veils of Ambiguity : Professionals ' Categorization of Oil and Organization Studies. In *Organization Studies*.

Lefsrud, L., Fox, K. A., Cooper, D. J., & Taminiau, Y. (2019). *Veils of Ambiguity: Professionals' Categorization of Oil and Gas Reserves*.

Lei, Y., Lin, J., He, Z., & Zuo, M. J. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, *35*(1–2), 108–126. https://doi.org/10.1016/j.ymssp.2012.09.015

Li, Y., Ahlstrom, D., & Ashkanasy, N. M. (2010). A multilevel model of affect and organizational commitment. *Asia Pacific Journal of Management*, *27*(2), 193–213. https://doi.org/10.1007/s10490-010-9193-9

Lithoxoidou, E., Vafeiadis, T., Krinidis, S., Ioannidis, D., Tzovaras, D., Ziogou, C., & Voutetakis, S. (2018). Malfunction diagnosis in industrial process systems using data mining for knowledge discovery. *2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017 - Proceedings*, *2018-Janua*, 454–461. https://doi.org/10.1109/ICE.2017.8279920

Litzinger, T. A., Wise, J. C., & Lee, S. H. A. (2005). Self-directed learning readiness among engineering undergraduate students. *Journal of Engineering Education*, *94*(2), 215–221. https://doi.org/10.1002/j.2168-9830.2005.tb00842.x

Lu, Y., Liu, C., Wang, K. I. K., Huang, H., & Xu, X. (2020). Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robotics and Computer-Integrated Manufacturing*, *61*(August 2019), 101837. https://doi.org/10.1016/j.rcim.2019.101837

Lucke, M., Chioua, M., Grimholt, C., Hollender, M., & Thornhill, N. F. (2018). Online Alarm Flood Classification Using Alarm Coactivations. *IFAC-PapersOnLine*, *51*(18), 345–350. https://doi.org/10.1016/j.ifacol.2018.09.324

Macchi, M., Roda, I., Negri, E., & Fumagalli, L. (2018). Exploring the role of Digital Twin for Asset Lifecycle Management. *IFAC-PapersOnLine*, *51*(11), 790–795. https://doi.org/10.1016/j.ifacol.2018.08.415

Manca, D., Brambilla, S., & Colombo, S. (2013). Advances in Engineering Software Bridging between Virtual Reality and accident simulation for training of process-industry operators. *Advances in Engineering Software*, *55*, 1–9. https://doi.org/10.1016/j.advengsoft.2012.09.002

Manca, D., Nazir, S., Lucernoni, F., & Colombo, S. (2012). Performance Indicators for the Assessment of Industrial Operators. In *Computer Aided Chemical Engineering* (Vol. 30, Issue June). Elsevier B.V. https://doi.org/10.1016/B978-0-444-59520-1.50143-3

Marsden, E. (2018). *The Heinrich / Bird safety pyramid Pioneering research has become a safety myth*. Risk Engineering. https://risk-engineering.org/concept/Heinrich-Bird-accident-pyramid

Marshall, P., Hirmas, A., & Singer, M. (2018). Heinrich's pyramid and occupational safety: A statistical validation methodology. *Safety Science*, *101*(September 2017), 180–189. https://doi.org/10.1016/j.ssci.2017.09.005

Mathew, A., Bever, K., Purser, M., & Ma, L. (2012). Bringing the MIMOSA OSA-EAI into an Object-Oriented World. In *Engineering Asset Management and Infrastructure Sustainability* (pp. 633–646). Springer London. https://doi.org/10.1007/978-0-85729-493-7_49

Maume-Deschamps, V., Rullière, D., & Said, K. (2015). *A risk management approach to capital allocation*.

Mazmanian, M., Orlikowski, W. J., & Yates, J. (2013). The Autonomy Paradox: The Implications of Mobile Email Devices for Knowledge Professionals. *Organization Science*, *24*(5), 1337–1357. https://doi.org/10.1287/orsc.1120.0806

McSween, T., & Moran, D. J. (2017). Assessing and Preventing Serious Incidents with Behavioral Science: Enhancing Heinrich's Triangle for the 21st Century. *Journal of Organizational Behavior Management*, *37*(3–4), 283–300. https://doi.org/10.1080/01608061.2017.1340923

Miller, C. A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. *Human Factors*, *49*(1), 57–75. https://doi.org/10.1518/001872007779598037

Mishra, P., Roger, J. M., Rutledge, D. N., Biancolillo, A., Marini, F., Nordon, A., & Jouan-Rimbaud-Bouveresse, D. (2020). MBA-GUI: A chemometric graphical user interface for multi-block data visualisation, regression, classification, variable selection and automated pre-processing. *Chemometrics and Intelligent Laboratory Systems*, *205*(August), 104139. https://doi.org/10.1016/j.chemolab.2020.104139

Miskinis, C. (2019). *The Dangers of Digital Twins and What Threats It Can Impose*. https://www.challenge.org/insights/digital-twin-risks/

Mitchell, W., Shah, S. L., & Shook, D. (2005). Challenges in the detection, diagnosis and visualisation of controller performance data. *Computing and Control Engineering*, *16*(4), 30–34. https://doi.org/10.1049/cce:20050406

Mkpat, E., Reniers, G., & Cozzani, V. (2018). Process safety education: A literature review. *Journal of Loss Prevention in the Process Industries*, *54*(February), 18–27. https://doi.org/10.1016/j.jlp.2018.02.003

Morbach, J., Yang, A., & Marquardt, W. (2007). OntoCAPE-A large-scale ontology for chemical process engineering. *Engineering Applications of Artificial Intelligence*, *20*(2), 147–161. https://doi.org/10.1016/j.engappai.2006.06.010

Morgan, M. G. (2002). *Risk Communication : A Mental Models Approach*. Cambridge University Press. http://login.ezproxy.library.ualberta.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=120190&site=ehost-live&scope=site

Motuzienė, V., Bielskus, J., Lapinskienė, V., Rynkun, G., & Bernatavičienė, J. (2022). Office buildings occupancy analysis and prediction associated with the impact of the COVID-19 pandemic. *Sustainable Cities and Society*, *77*. https://doi.org/10.1016/j.scs.2021.103557

Moura, R., Beer, M., Patelli, E., Lewis, J., & Knoll, F. (2017). Learning from accidents: Interactions between human factors, technology and organisations as a central element to validate risk studies. *Safety Science*, *99*, 196–214. https://doi.org/10.1016/j.ssci.2017.05.001

Naderpour, M., Lu, J., & Zhang, G. (2015). A human-system interface risk assessment method based on mental models. *Safety Science*, *79*, 286–297. https://doi.org/10.1016/j.ssci.2015.07.001

Naef, M., Chadha, K., & Lefsrud, L. (2021). Decision support for process operators. *CCEC 2021 - 71st Canadian Chemical Engineering Conference*.

Naef, M., Chadha, K., & Lefsrud, L. (2022). Decision support for process operators: Task loading in the days of big data. *Journal of Loss Prevention in the Process Industries*, *75*. https://doi.org/10.1016/j.jlp.2021.104713

Naef, M., & Lefsrud, L. (2020). Teaching PSM: COVID-19 and Alberta Public School Re-entry. *CCEC 2020 - 70th Canadian Chemical Engineering Conference*. https://www.alberta.ca/covid-19-orders-and-legislation.aspx#toc-0

Nakai, A., Kajihara, Y., Nishimoto, K., & Suzuki, K. (2017). Information-sharing system supporting onsite work for chemical plants. *Journal of Loss Prevention in the Process Industries*, *50*, 15–22. https://doi.org/10.1016/J.JLP.2017.08.011

Ngo, T., & Kunkel, J. (2020). *Exploring Mental Models for Transparent and Controllable Recommender Systems : A Qualitative Study*. 183–191.

Nguyen, P. H., Tangworakitthaworn, P., & Gilbert, L. (2020). Individual learning effectiveness based on cognitive taxonomies and constructive Alignment. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, *2020-Novem*, 1002–1006. https://doi.org/10.1109/TENCON50793.2020.9293733

Nix, J., Jers, H., & Ternström, S. (2020). Acoustical, psychoacoustical, and pedagogical considerations for choral singing with COVID-19 health measures. *Source: The Choral Journal*, *61*(3), 30–39. https://doi.org/10.2307/27034960

Normanyo, E., Husinu, F., & Agyare, O. R. (2014). Developing a Human Machine Interface (HMI) for Industrial Automated Systems using Siemens Simatic WinCC Flexible Advanced Software. *Journal of Emerging Trends in Computing and Information Sciences*, *5*(2), 134–144.

O' Donovan, A., & O' Sullivan, P. D. (2023). The impact of retrofitted ventilation approaches on long-range airborne infection risk for lecture room environments: design stage methodology and application. *Journal of Building Engineering*, *68*. https://doi.org/10.1016/j.jobe.2023.106044

Ottewell, S. (2011). Operator Training Gains Ground. *Chemical Processing Progress*.

Ottewell, S. (2016). Training Simulators Make Real Difference. *Chemical Processing Progress*.

Padovano, A., Longo, F., Nicoletti, L., & Mirabelli, G. (2018). A Digital Twin based Service Oriented Application for a 4.0 Knowledge Navigation in the Smart Factory. *IFAC-PapersOnLine*, *51*(11), 631–636. https://doi.org/10.1016/j.ifacol.2018.08.389

Palzer, S. (2020). Photoacoustic-based gas sensing: A review. *Sensors*, *20*(9). https://doi.org/10.3390/s20092745

Pan, Y., Du, C., Fu, Z., & Fu, M. (2021). Re-thinking of engineering operation solutions to HVAC systems under the emerging COVID-19 pandemic. *Journal of Building Engineering*, *43*. https://doi.org/10.1016/j.jobe.2021.102889

Paramythis, A., Weibelzahl, S., & Masthoff, J. (2010). Layered evaluation of interactive adaptive systems: Framework and formative methods. *User Modeling and User-Adapted Interaction*, *20*(5), 383–453. https://doi.org/10.1007/s11257-010-9082-4

Park, J., & Jung, W. (2015). A systematic framework to investigate the coverage of abnormal operating procedures in nuclear power plants. *Reliability Engineering and System Safety*, *138*, 21–30. https://doi.org/10.1016/j.ress.2015.01.013

Patle, D. S., Ahmad, Z., & Rangaiah, G. P. (2014). Operator training simulators in the chemical industry: review, issues, and future directions. *Reviews in Chemical Engineering*, *30*(2). https://doi.org/10.1515/revce-2013-0027

Peng, D., Gu, X., Xu, Y., & Zhu, Q. (2015). Integrating probabilistic signed digraph and reliability analysis for alarm signal optimization in chemical plant. *Journal of Loss Prevention in the Process Industries*, *33*, 279–288. https://doi.org/10.1016/j.jlp.2015.01.005

Perry, P. (2016). Human factors in health and safety. *Health and Safety Questions and Answers*, *September*, 131–140. https://doi.org/10.1680/hasqaa.60746.131

Petersen, S., Jensen, K. L., Pedersen, A. L. S., & Rasmussen, H. S. (2016). The effect of increased classroom ventilation rate indicated by reduced CO2 concentration on the performance of schoolwork by children. *Indoor Air*, *26*(3), 366–379. https://doi.org/10.1111/ina.12210

Powell, A. (2022). Is pandemic finally over? We asked the experts. *The Harvard Gazette*. https://news.harvard.edu/gazette/story/2022/10/is-pandemic-finally-over-we-asked-the-experts/

Rajabiyazdi, F., & Jamieson, G. A. (2020). A Review of Transparency (seeing-into) Models. *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 302–308.

Rebello, S., Yu, H., & Ma, L. (2019). An integrated approach for real-time hazard mitigation in complex industrial processes. *Reliability Engineering and System Safety*, *188*(February), 297–309. https://doi.org/10.1016/j.ress.2019.03.037

Revell, K. M. A., & Stanton, N. A. (2016). Mind the gap – Deriving a compatible user mental model of the home heating system to encourage sustainable behaviour. *Applied Ergonomics*, *57*, 48–61. https://doi.org/10.1016/j.apergo.2016.03.005

Rosbach, J., Krop, E., Vonk, M., van Ginkel, J., Meliefste, C., de Wind, S., Gehring, U., & Brunekreef, B. (2016). Classroom ventilation and indoor air quality—results from the FRESH intervention study. *Indoor Air*, *26*(4), 538–545. https://doi.org/10.1111/ina.12231

Ross, M. (2013). Improve Operator Training. *Chemical Processing Progress*.

Roth, E. M., Mumaw, R. J., Vicente, K. J., & Burns, C. M. (1997). Operator monitoring during normal operations: vigilance or problem-solving. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, 158–162.

Sandvik, H. (2008). Tree thinking cannot taken for granted: Challenges for teaching phylogenetics. *Theory in Biosciences*, *127*(1), 45–51. https://doi.org/10.1007/s12064-008-0022-3

Schijven, J., Vermeulen, L. C., Swart, A., Meijer, A., Duizer, E., & de Roda Husman, A. M. (2021). Quantitative microbial risk assessment for airborne transmission of sars-cov-2 via breathing, speaking, singing, coughing, and sneezing. *Environmental Health Perspectives*, *129*(4), 1–11. https://doi.org/10.1289/EHP7886

Seel, N. M. (2006). Mental Models and Complex Problem Solving: Instructional Effects. In *Handling Complexity in Learning Environments: Theory and Research* (pp. 43–66).

Seng, Y., & Srinivasan, R. (2010). Engineering Applications of Artificial Intelligence Multi-agent based collaborative fault detection and identification in chemical processes. *Engineering Applications of Artificial Intelligence*, *23*(6), 934–949. https://doi.org/10.1016/j.engappai.2010.01.026

Shanqing, Y. (2012). *Proactive Monitoring in Process Control using Predictive Trend Display* [Doctorate, Nanyang Technological University]. https://hdl.handle.net/10356/49971

Shendell, D. G., Aggarwal, J., Campbell, M. L. F., Gonzalez, L. N., Kaplun, E., Koshy, K., & Mackie, T. I. (2021). Fall 2020 covid-19 needs assessment among new jersey secondary school educational professionals. *International Journal of Environmental Research and Public Health*, *18*(8). https://doi.org/10.3390/ijerph18084083

Shendell, D. G., Prill, R., Fisk, W. J., Apte, M. G., Blake, D., & Faulkner, D. (2004). Associations between classroom CO2 concentrations and student attendance in Washington and Idaho. *Indoor Air*, *14*(5), 333–341. https://doi.org/10.1111/j.1600-0668.2004.00251.x

Shendell, D. G., Winer, A. M., Weker, R., & Colome, S. D. (2004). Evidence of inadequate ventilation in portable classrooms: Results of a pilot study in Los Angeles County. *Indoor Air*, *14*(3), 154–158. https://doi.org/10.1111/j.1600-0668.2004.00235.x

Sherer, P. D. (2019). The researcher as detective: Another way of looking at professions and organizations. *Journal of Professions and Organization*, *6*(1), 91–102. https://doi.org/10.1093/jpo/joy020

Shook, D., Khalili, R., Grenier, M., & Lariviere, L. (2003). Achieving business results through IT: Insuring controller performance benefits with web-enabled monitoring at Noranda. *CIM Bulletin*, *96*(1067), 57–59.

Shrestha, M., Rijal, H. B., Kayo, G., & Shukuya, M. (2022). An investigation on CO2 concentration based on field survey and simulation in naturally ventilated Nepalese school buildings during summer. *Building and Environment*, *207*. https://doi.org/10.1016/j.buildenv.2021.108405

Shu, Y., Ming, L., Cheng, F., Zhang, Z., & Zhao, J. (2016). Abnormal situation management : Challenges and opportunities in the big data era. *Computers and Chemical Engineering*, *91*, 104–113. https://doi.org/10.1016/j.compchemeng.2016.04.011

Silaipillayarputhur, K. (2018). Process safety management in manufacturing industries a review. *International Journal of Engineering & Technology*, *7*(2), 540–543. https://doi.org/10.14419/ijet.v7i2.9252

Sloan, J. A., & Scharff, L. F. (2022). Student Self-Assessment: Relationships between Accuracy, Engagement, Perceived Value, and Performance. *Journal of Civil Engineering Education*, *148*(3). https://doi.org/10.1061/(asce)ei.2643-9115.0000063

Stawarz, A. C., & Sowerby, B. (1995). Cost effective operator training. *Computers & Chemical Engineering*, *19*, 459–464. https://doi.org/10.1016/0098-1354(95)87079-2

Stockman, T., Zhu, S., Kumar, A., Wang, L., Patel, S., Weaver, J., Spede, M., Milton, D. K., Hertzberg, J., Toohey, D., Vance, M., Srebric, J., & Miller, S. L. (2021). Measurements and Simulations of Aerosol Released while Singing and Playing Wind Instruments. *ACS Environmental Au*, *1*(1), 71–84. https://doi.org/10.1021/acsenvironau.1c00007

Strobhar, D. A. (2014a). 4 System Demands/Automation. In *Human Factors in Process Plant Operation* (pp. 25–36). https://www.ebsco.com/terms-of-use

Strobhar, D. A. (2014b). 6 Interface. In *Human Factors in Process Plant Operation* (pp. 51–76). https://www.ebsco.com/terms-of-use

Strobhar, D. A. (2014c). 7 Selection and Training. In *Human Factors in Process Plant Operation* (pp. 77–88). https://www.ebsco.com/terms-of-use

Strobhar, D. A. (2014d). 8 Job and Organizational Design. In *Human Factors in Process Plant Operation* (pp. 89–102). https://www.ebsco.com/terms-of-use

Strobhar, D. A. (2014e). 9 Procedures/Job Aids. In *Human Factors in Process Plant Operation* (pp. 103–112). https://www.ebsco.com/terms-of-use

Strobhar, D. A. (2014f). *Human Factors in Process Plant Operation*. Momentum Press. http://login.ezproxy.library.ualberta.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=799855&site=eds-live&scope=site

Strobhar, D. A., & Harvey, C. M. (2011). How Many Alarms Can an Operator Handle ? *Chemical Processing Progress*.

Suresh, R., Sivaram, A., & Venkatasubramanian, V. (2019). A hierarchical approach for causal modeling of process systems. *Computers and Chemical Engineering*, *123*, 170–183. https://doi.org/10.1016/j.compchemeng.2018.12.017

Susarev, S. V., Bulkaeva, E. A., Sarbitova, Y. V., & Dolmatov, D. S. (2017). Training simulators development technique for oil and gas industry automation control systems. *Proceedings of 2017 IEEE 2nd International Conference on Control in Technical Systems, CTS 2017*, 207–210. https://doi.org/10.1109/CTSYS.2017.8109527

Svenson, O. (1989). On expert judgements in safety analyses in the process industries. *Reliability Engineering and System Safety*, *25*(3), 219–256. https://doi.org/10.1016/0951-8320(89)90095-1

Tamplin, J., & Thompson, Z. (2023). How health-focused choirs adapted to the virtual world during the COVID-19 pandemic – An international survey. *Arts in Psychotherapy*, *82*. https://doi.org/10.1016/j.aip.2023.101997

Tan, S., Zhang, Z., Maki, K., Fidkowski, K. J., & Capecelatro, J. (2022). Beyond well-mixed: A simple probabilistic model of airborne disease transmission in indoor spaces. *Indoor Air*, *32*(3). https://doi.org/10.1111/ina.13015

Thacker, I., & Sinatra, G. (2019). Visualizing the Greenhouse Effect: Restructuring Mental Models of Climate Change Through a Guided Online Simulation. *Education Sciences*, *9*(1), 14. https://doi.org/10.3390/educsci9010014

Theodore, L., & Dupont, R. R. (2022). Introduction to Environmental and Health Risk. In *Chemical Process Industries* (pp. 9–20). CRC Press. https://doi.org/10.1201/9781003283454-3

Tufecki, Z. (2021). 5 Pandemic Mistakes We Keep Repeating. *The Atlantic*. https://www.theatlantic.com/ideas/archive/2021/02/how-public-health-messaging-backfired/618147/

van de Merwe, K., Mallam, S., & Nazir, S. (2022). Agent Transparency, Situation Awareness, Mental Workload, and Operator Performance: A Systematic Literature Review. *Human Factors*. https://doi.org/10.1177/00187208221077804

van der Zee, S. C., Strak, M., Dijkema, M. B. A., Brunekreef, B., & Janssen, N. A. H. (2017). The impact of particle filtration on indoor air quality in a classroom near a highway. *Indoor Air*, *27*, 291–302. https://doi.org/10.1111/ina.12308

Venkatasubramanian, V. (2003). Abnormal events management in complex process plants : challenges and opportunities in intelligent supervisory control. *(FOCAPO2003) Fault Diagnosis, Abnormal Events Management, Process Hazards Analysis*, 117–132.

Venkatasubramanian, V. (2011). Systemic Failures: Challenges and Opportunities in Risk Management in Complex Systems. *AIChE Journal*, *57*(1), 2–9. https://doi.org/DOI 10.1002/aic.12495

Venkatasubramanian, V. (2019a). The Promise of Artificial Intelligence in. *Chemical Engineering IT*, *65*(2). https://doi.org/10.1002/aic.16489

Venkatasubramanian, V. (2019b). The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal*, *65*(2), 466–478. https://doi.org/10.1002/aic.16489

Vicente, K. J., Mtmiaw, R. J., & Roth, E. M. (1998). More about operator monitoring under normal operations: the role of workload regulation and the impact of control room technology. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, 229–233.

Wang, H., Khan, F., & Abimbola, M. (2018). A new method to study the performance of safety alarm system in process operations. *Journal of Loss Prevention in the Process Industries*, *56*(May), 104–118. https://doi.org/10.1016/j.jlp.2018.08.017

Wang, H., Khan, F., & Ahmed, S. (2015). Design of scenario-based early warning system for process operations. *Industrial and Engineering Chemistry Research*, *54*(33), 8255–8265. https://doi.org/10.1021/acs.iecr.5b02481

Whitley, C. T., Rivers III, L., Mattes, S., Marquart-Pyatt, S. T., Ligmann-Zielinska, A., Olabisi, L. S., & Du, J. (2018). Climate-induced migration: using mental models to explore aggregate and individual decision-making. *Journal of Risk Research*, *21*(8), 1019–1035. https://doi.org/10.1080/13669877.2017.1281331

Wilkins, J. R. (2011). Construction workers' perceptions of health and safety training programmes. *Construction Management and Economics*, *29*(10), 1017–1026. https://doi.org/10.1080/01446193.2011.633538

Woodcock, B., & Au, Z. (2013). Human factors issues in the management of emergency response at high hazard installations. *Journal of Loss Prevention in the Process Industries*, *26*(3), 547–557. https://doi.org/10.1016/j.jlp.2012.07.002

Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, *78*(3), 645–675. https://doi.org/10.3102/0034654308320320

Yang, F., Duan, P., Shah, S. L., & Chen, T. (2014). Capturing connectivity and causality in complex industrial processes. In *SpringerBriefs in Applied Sciences and Technology* (Issue 9783319053790). https://doi.org/10.1007/978-3-319-05380-6

Yim, H. Bin, & Seong, P. H. (2015). A Quantitative Team Situation Awareness Measurement Method Considering Technical and Nontechnical Skills of Teams. *Nuclear Engineering and Technology*, *48*(1), 144–152. https://doi.org/10.1016/j.net.2015.09.007

Young, A. M., Wendel, P. J., Esson, J. M., & Plank, K. M. (2018). Motivational decline and recovery in higher education STEM courses. *International Journal of Science Education*, *40*(9), 1016–1033. https://doi.org/10.1080/09500693.2018.1460773

Zenk, L., Stadtfeld, C., & Windhager, F. (2010). How to analyze dynamic network patterns of high performing teams. *Procedia - Social and Behavioral Sciences*, *2*(4), 6418–6422. https://doi.org/10.1016/j.sbspro.2010.04.051

Zhang, Z., & Dong, F. (2014). Fault detection and diagnosis for missing data systems with a three time-slice dynamic Bayesian network approach. *Chemometrics and Intelligent Laboratory Systems*, *138*, 30–40. https://doi.org/10.1016/j.chemolab.2014.07.009

Zhong, L., & Xu, X. (2019). Developing real life problem-solving skills through situational design: a pilot study. *Educational Technology Research and Development*, *67*(6), 1529–1545. https://doi.org/10.1007/s11423-019-09691-2

Zhu, J., Shu, Y., Zhao, J., & Yang, F. (2014). A dynamic alarm management strategy for chemical process transitions. *Journal of Loss Prevention in the Process Industries*, *30*, 207–218. https://doi.org/10.1016/j.jlp.2013.07.008

Appendices

The following appendices contain supplemental documents and records supporting the manuscripts in Chapters 2, 3 and 4.

Appendix 1 contains the study summary submitted to the University of Alberta Research Ethics Board and approval for the study described in Chapter 3.

Appendix 2 contains the assessment instruments employed in the study described in Chapter 3.

Appendix 3 contains the manufacturer's data sheet for the environmental monitoring sensors used in the study described in Chapter 4.

Appendix 1

1.0 Ethics Board Approval

Research Ethics Board approval transferred from arise.ualberta.ca Pro00128584:

Notification of Approval

Date: March 2, 2023

Study ID: Pro00128584

Principal Investigator: Lianne Lefsrud

Study Title: Demonstrating the effectiveness of the situational design model in evaluating user interface design features

Approval Expiry Date: Friday, March 1, 2024

Sponsor/Funding Agency: NSERC - Natural Sciences And Engineering Research Council

RSO Managed Funding:

Project ID	Title	Grant Status	Sponsor	Project Start Date	Project End Date	Purpose
RES0045685	Developing a Safety Management Systems (SMS) Model	Awarded	Dscvry Launch Supp Erly Career	4/1/2018	3/31/2022	Grant
RES0038828	Developing a Safety Management Systems (SMS) Model	Awarded	Discov Grt Individual	4/1/2018	3/31/2025	Grant

Thank you for submitting the above study to the Research Ethics Board 2. Your application has been reviewed and approved on behalf of the committee.

Approved Documents:

Recruitment Materials
Recruitment letter
Consent Forms
Informed Consent
Questionnaires, Cover Letters, Surveys, Tests, Interview Scripts, etc.
Troubleshooting Task Description
DCS and MPC Knowledge Exam
Learning Readiness Survey

Any proposed changes to the study must be submitted to the REB for approval prior to implementation. A renewal report must be submitted next year prior to the expiry of this approval if your study still requires ethics approval. If you do not renew on or before the renewal expiry date, you will have to re-submit an ethics application.

Approval by the REB does not constitute authorization to initiate the conduct of this research. The Principal Investigator is responsible for ensuring required approvals from other involved organizations (e.g., Alberta Health Services, Covenant Health, community organizations, school boards) are obtained, before the research begins.

Sincerely,

Carol Boliek, Ph.D
Associate Chair, Research Ethics Board 2

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Section 2 – Study Summary content transferred in full from arise.ualberta.ca Pro00128584. Heading formats and numbering preserved from the Arise system.

2.1 Study Objectives and Design

1.0 Provide a lay summary of your proposed research which would be understandable to general public

Modern chemical processes rely on distributed control systems to make repetitive and routine adjustments to maintain steady-state operation. Operators are still required to "supervise the (system) supervisor" and intervene when variables exceed pre-programmed parameters to avert major incidents. Research in human-computer interaction and advanced process control has often focused on data-driven methods for fault detection as distinct from operator effectiveness, without examining the interaction. In this research, we are developing an evaluation framework for 1) UX (user experience) for data visualization/interface and 2) user/operator learning from the UX interface. We expand on the situational design model, piloted by researchers in Educational Psychology, to demonstrate the effect of a visual design feature on a standard operator interface screen. The situational design model evaluates improvements in problem-solving skills. The objective of the study is to determine whether a visual adaptation to the standard screen results in improved problem-solving performance during a routine process operating task.

2.0 Provide a full description of your research proposal outlining the following*:*

Purpose

Hypothesis

Justification

Objectives

Research Method/Procedures

Plan for Data Analysis

Purpose

To demonstrate the efficacy of a systematic evaluation framework in determining the value of design features on a control system interface.

Hypothesis

Design features can be evaluated effectively if improvements in problem-solving skills can be measured

Justification

There are no systematic methods readily available to evaluate the degree to which a digital systems interface improves problem solving skills. There is demonstrated need for such a methodology as the chemical processing industry increases the integration of automated control systems and process safety incidents continue to identify "human error" as a significant causal factor. A common theme in incident investigation recommendations is that additional training is required, but industrial training practices do not have a ready catalog of standard assessment methodologies or instructional practices. This study tests a new method that will allow "operator understanding" to be rigorously measured and contextualized within the industrial knowledge base for the chemical process industry. Because industrial informatics have grown into one of the primary knowledge-transfer tools, the degree to which design features improve (or reduce) operator understanding of complex industrial processes is relevant and the lack of such a framework makes rigorous design evaluation more difficult.

Objectives

1. Evaluate the effectiveness of a visual design feature on a typical operator interface by measuring the degree to which the feature improved the operator's understanding of the system represented on the interface

2. Demonstrate the utility of a rigorous assessment methodology for design evaluations in the chemical process industry

Research Method/Procedures

The procedures have been adapted from a pilot study published by Zhong and Xu (2019) 10.1007/s11423-019-09691-2

30 volunteer participants will be selected from a pool of University students, the entry criteria being successful completion of Chemical Engineering 464, a mandatory fourth year undergraduate course in chemical engineering design. The potential participants include current undergraduate students and graduate students. The recruitment letter will be attached. The entry criteria ensures that all participants have demonstrated competence in several necessary recurrent skill groups, including process flow symbology, basic fluid mechanics, state and phase behavior for Newtonian and non-Newtonian fluids and basic control system design and representation. Participants will be offered a small honorarium.

The participant pool will be separated into two groups, one acting as a control accessing a typical interface and the other accessing the visually augmented interface.

Participants will be asked to complete the following activities.

Learning Readiness Survey 1

A self-assessment survey using Likert scales will be completed prior to the exam evaluating dimensions of learner readiness

Exam 1

The participants will be engaged twice, the first interaction will include a 20 question multiple-choice exam intended to establish a baseline knowledge level in the subject areas relevant a chemical process operations problem. Participants will be provided with feedback on their responses, and incorrect selections will be directly linked to supportive learning activities according to the situational design model.

Task 1

Following the multiple-choice exam, which will be timed to allow participants ample opportunity to review the feedback information, they will access a simple process simulation of a Model Predictive Control scheme. They will have the ability to manipulate variables in both a "steady" and "fault" state. Their objective will be to identify the fault and propose a possible solution.

Two graders will independently evaluate participant solutions to Task 1, and structured feedback will be provided.

Participants will have access to the simulator for the period between tasks, and encouraged to interact with it, but interaction is not necessary. All participant interaction with the simulator will be logged in a limited manner, the participant's identification number and the results of the simulations they chose to run will be retained. Participants will also have access to their exam and feedback results, with session data logged.

Learning Readiness Survey 2

The identical readiness survey will be provided to participants.

Exam 2

A similar multiple choice exam will be given to participants, with questions selected from the same pool as Exam 1. Students may answer new questions or repeat previous ones. Feedback will be provided in the same manner as before.

Task 2

A similar process simulation task will be provided to participants, with slight variations to some parameters. The system will be the same, and the interface will be identical. Participants will again be asked to identify a fault in the process control scheme and propose a solution.

Participants will be released following the second task.

Plan for Data Analysis

Solutions will be marked by 2 independent graders using the same rubric and scoring scale. Averaged scores will be used for the analysis of results excepting interrater reliability.

Interrater reliability will be assessed using Cohen's Kappa statistic

The exams and task scores will be evaluated using the Wilcoxian signed-rank test pulling out recurrent and non-recurrent skills, problem-solving (fault detection and solution generation), knowledge/exam skills, time spent with the simulator between tests

Spearman's rank-order test will be applied to further analyze the correlation between recurrent and non-recurrent skills as well as the correlation between exam performance and task performance

The reliability of the learner readiness evaluations will be analyzed using Cronbach's Alpha.

Correlations and scores from the signed-rank tests will be compared between the control and study groups.

3.0 Describe procedures, treatment, or activities that are above or in addition to standard practices in this study area (eg. extra medical or health-related procedures, curriculum enhancements, extra follow-up, etc):
The practices proposed for this study are common in the fields of education and psychology. Surveys using Likert scales, multiple choice exams and a scenario-based task will be provided to participants. The study will take place over two weeks, to allow participants to demonstrate the acquisition of recurrent skills which are necessary to differentiate responses during the task portion which demonstrate the application of non-recurrent skills.

4.0 If the proposed research is above minimal risk and is not funded via a competitive peer review grant or industry-sponsored clinical trial, the REB will require evidence of scientific review. Provide information about the review process and its results if appropriate.

5.0 For clinical trials, describe any sub-studies associated with this Protocol.

2.2 Research Methods and Procedures
Some research methods prompt specific ethical issues. The methods listed below have additional questions associated with them in this application. If your research does not involve any of the methods listed below, ensure that your proposed research is adequately described in Section 2.1: Study Objectives and Design or attach documents in the Documentation Section if necessary.

1.0 This study will involve the following (select all that apply)
Internet-based Interaction with Participants (excluding internet surveys or data collection over internet without human interaction)

Surveys and Questionnaires (including internet surveys)

NOTE 1: Select this if you are directly collecting health information as part of your protocol OR will be conducting a chart/record review/reviewing health data secondarily. This includes anonymized or identifiable health information.

NOTE 2: Select this option if this research ONLY involves analysis of blood/tissue/specimens originally collected for another purpose but now being used to answer your research question. If you are enrolling people into the study to prospectively collect specimens to analyze you SHOULD NOT select this box.

NOTE 3: This section is intended to reflect the secondary use of non-health data. Do NOT select this if you are using data that originally came from health sources, i.e., anonymized administrative data.

2.4 Internet-based Interaction with Human Participants

1.0 Internet-based Research

1.1 Will your interaction with participants occur in private internet spaces (eg. members only chat rooms, social networking sites, email discussions, etc)?
No

1.2 Will these interactions occur in public space(s) where you will post questions initiating and/or maintaining interaction with participants?
No

2.0 Describe how permission to use the site(s) will be obtained, if applicable:
I intend to use the University eClass platform (in sandbox mode) to host the simulator and the evaluation. I will obtain permission from the CME department to use the platform for this purpose.

Participants will be assigned anonymous identifiers on their results, so that their CCIDs are not linked to their performance.

3.0 If you do not plan to identify yourself and your position as a researcher to the participants, from the onset of the research study, explain why you are not doing so, at what point you will disclose that you are a researcher, provide details of debriefing procedures, if any, and if participants will be given a way to opt out, if applicable:

4.0 How will you protect the privacy and confidentiality of participants who may be identified by email addresses, IP addresses, and other identifying information that may be captured by the system during your interactions with these participants?
I intend to assign participants an eClass "sandbox" class created strictly for the study. Depending on how IT services are able to support this project, the research team may not have access to the identity of the participants at all, if the anonymized profiles can be assigned directly by the department to a list of volunteer participants.

2.9 Surveys and Questionnaires (including Online)

1.0 How will the survey/questionnaire data be collected (i.e. collected in person, or if collected online, what survey program/software will be used etc.)?
eClass will be used to collect survey, exam and task data.

2.0 Where will the data be stored once it's collected (i.e. will it be stored on the survey software provider servers, will it be downloaded to the PI's computer, other)?
Secure department drive.

3.0 Who will have access to the data?
The Principal Investigator and the Co-Investigator

4.0 If you are using a third party research tool, website survey software, transaction log tools, screen capturing software, or masked survey sites, how will you ensure the security of data gathered at that site?
University of Alberta eClass will be used.

Appendix 2

The following survey was transferred from the original format in e-Class. The following text appeared on the course description page, under "Learning Readiness Survey" in eClass.

Learning Readiness Survey

Please choose the option that best describes how you feel.

For each question, assume that "5" is the level of knowledge necessary for an engineer or EIT to immediately begin work on a new problem in that area without requiring review of the theory or general practice elements.

Please note, survey has been adapted from:

Zhong, L., & Xu, X. (2019). Developing real life problem-solving skills through situational design: a pilot study. Educational Technology Research and Development, 67(6), 1529–1545. https://doi.org/10.1007/s11423-019-09691-2

How much knowledge do you have regarding distributed control system interfaces or control panels?

- (1)Have seen representations in textbooks or photos
- (2)
- (3)
- (4)
- (5)Understand the function and purpose of objects on the interface

How much knowledge do you have on the topic of model predictive control?

- (1)Familiar with the term
- (2)
- (3)
- (4)
- (5)Can describe the theoretical elements including equations and standard variable names

What level of experience do you have interacting with DCS interfaces or control panels?

- (1)None, experience is with theory only

- (2)
- (3)
- (4)
- (5)>10 hours experience with one or more interfaces whether operating or in design capacity

What level of experience do you have with model predictive control?

- (1)Familiar with the term
- (2)
- (3)
- (4)
- (5)>10 hours experience operating (including troubleshooting) or designing model predictive control schemes

What level of interest do you have in the operation or design of DCS interfaces?

- (1)Very Little
- (2)
- (3)
- (4)
- (5)Very much

What level of interest do you have in model predictive control, whether in design, operation OR maintenance?

- (1)Very little
- (2)
- (3)
- (4)
- (5)Very much

Thank you for completing the learner readiness survey. Please proceed to Exam 1.

Marking Instructions - simulator task 1

Please review participant responses, in the excel file, each participant is a separate row. All four responses are included in their entirety. The marking rubric is replicated in excel on a separate sheet, but presented here for ease of review. There is a final column that will total the scores, please enter your question score in the column titled “score x” (to the immediate right of the response.)

Note that for Task 1, the simulator did not work as intended. There was no difference between abnormal and steady state. This was detected in the first session and confirmed in the third, but after discussion with the PI it was decided to continue with the experiment and give consistent experiences to all participants. Participants were advised to describe the ways they determined the simulation was the same. As a result the marking for this task may be challenging. It would be equally correct for participants to identify “display” as the root cause of a problem, as well as to simply state that there is no abnormal condition. All participants were advised to explain their logic, and the marking scheme can still be applied. The second session will be conducted with two distinct simulation files, and the “abnormal condition” is dramatic and materializes very quickly.

For each question, please apply each possible mark only once. For example, in question 1, if they describe multiple pieces of equipment, only 1 point is awarded for the first item “Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed” If they describe the continuously stirred reactor (1 point) and then describe connected systems and the control scheme (2 points) a total of 3 points is awarded for that statement. Similarly, if they describe the reactor (1 point) and give a detailed description of the control scheme for the Stripper and how it is linked to the downstream product analyzer trend (5 points), they receive 6 total points for their “Interaction with simulator and observations”.

Please also note that participants do not review these marks, and the marks awarded are design evaluation scores…not reflective of the “performance” of the participant.

Marking rubric for each question

Question 1 – weight = 0.2

Describe the ways you recognized the "abnormal situation" mode to be different from the "steady state" operating mode. Please be specific, reference equipment by tag number and concisely describe transient behavior (for example: the temperature readings on the reactor began to oscillate with increasing amplitude and frequency)

Interaction with simulator and observations		
Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed		1

Any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)		2
Description of basic characteristics, single nodes or individual trends		2
Discussion of what might occur or what might have occurred, description of tunable parameter impact on process		2
Discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge		5
Description of a system or subsystem not exhibiting transient behaviour		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2
Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Interface Utility		
Description of a single measurement, variable or trend that appeared to be missing		1
Discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture		2
Analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset		5
Total		28

Question 2 – weight = 0.5

What do you think is the root cause of the abnormal situation that you observed? Please offer a root cause and follow it with a brief description of your logic in arriving at that conclusion. As with the previous question, please be specific and concise in your response, and assume that markers have knowledge of process control terminology.

Please describe any additional tests, measurements or diagnostic activities that you would recommend to increase confidence in your result. If you do not believe any additional information is required, please briefly state why you are confident in your diagnosis.

Measuring decision utility		

Description of a single, directly causal factor without substantial discussion around additional tests or diagnostics		1
Description of a possible causal factor drawing dependent systems into the analysis, limited identification of additional steps or tests, simple recommendation that does not incorporate system-scale or multiple time-interval considerations		2
Detailed and well supported root cause with one or more linked subsystems or components.		5
Supporting logic		
Detailed and well supported root cause with one or more linked subsystems or components.		
Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed		1
Any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)		2
Description of basic characteristics, single nodes or individual trends		2
Discussion of what might occur or what might have occurred, description of tunable parameter impact on process		2
Discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge		5
Description of a system or subsystem not exhibiting transient behaviour		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2
Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Interface utility		
Description of a single measurement, variable or trend that appeared to be missing		1
Discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture		2
Analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset		5
Total		36

Question 3 – weight = 0.1

Consider the simulation run as a whole. Did any of the trends or behaviors you observed stand out as anomalous or unexpected, besides those you identified in your root cause logic?

Interface utility		
Description of a single measurement, variable or trend that appeared to be missing		1

Discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture		2
Analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset		5
Interaction with simulation		
Any discussion of process, control equipment that appears in the simulation, accurate description of what was observed		1
Any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)		2
Description of basic characteristics, single nodes or individual trends		2
Discussion of what might occur or what might have occurred, description of tunable parameter impact on process		2
Discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge		5
Description of a system or subsystem not exhibiting transient behaviour		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2
Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Total		27

Question 4 – weight = 0.2

What would you recommend the plant owner/operator do to address the root cause of this abnormal situation? Would the problem present the same way in all operating modes?

Measuring decision utility		
Description of a single, directly causal factor without substantial discussion around additional tests or diagnostics		1
Description of a possible causal factor drawing dependent systems into the analysis, limited identification of additional steps or tests, simple recommendation that does not incorporate system-scale or multiple time-interval considerations		2
Detailed and well supported root cause with one or more linked subsystems or components.		5
Causal reasoning		
Description of a system or subsystem not exhibiting transient behaviour		1
Description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)		2

Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.		5
Interface utility		
Description of a single measurement, variable or trend that appeared to be missing		1
Discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture		2
Analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset		5
Total		24

Submission score: multiply question percentage by question weight, submission score will be a percentage

	Participant score (A)	Total (B)	Weight (C)	Weighted scores '= (A / B) * C
Question 1		28	.2	
Question 2		36	.5	
Question 3		27	.1	
Question 4		24	.2	
			Submission score (total of weighted)	SUM(column)

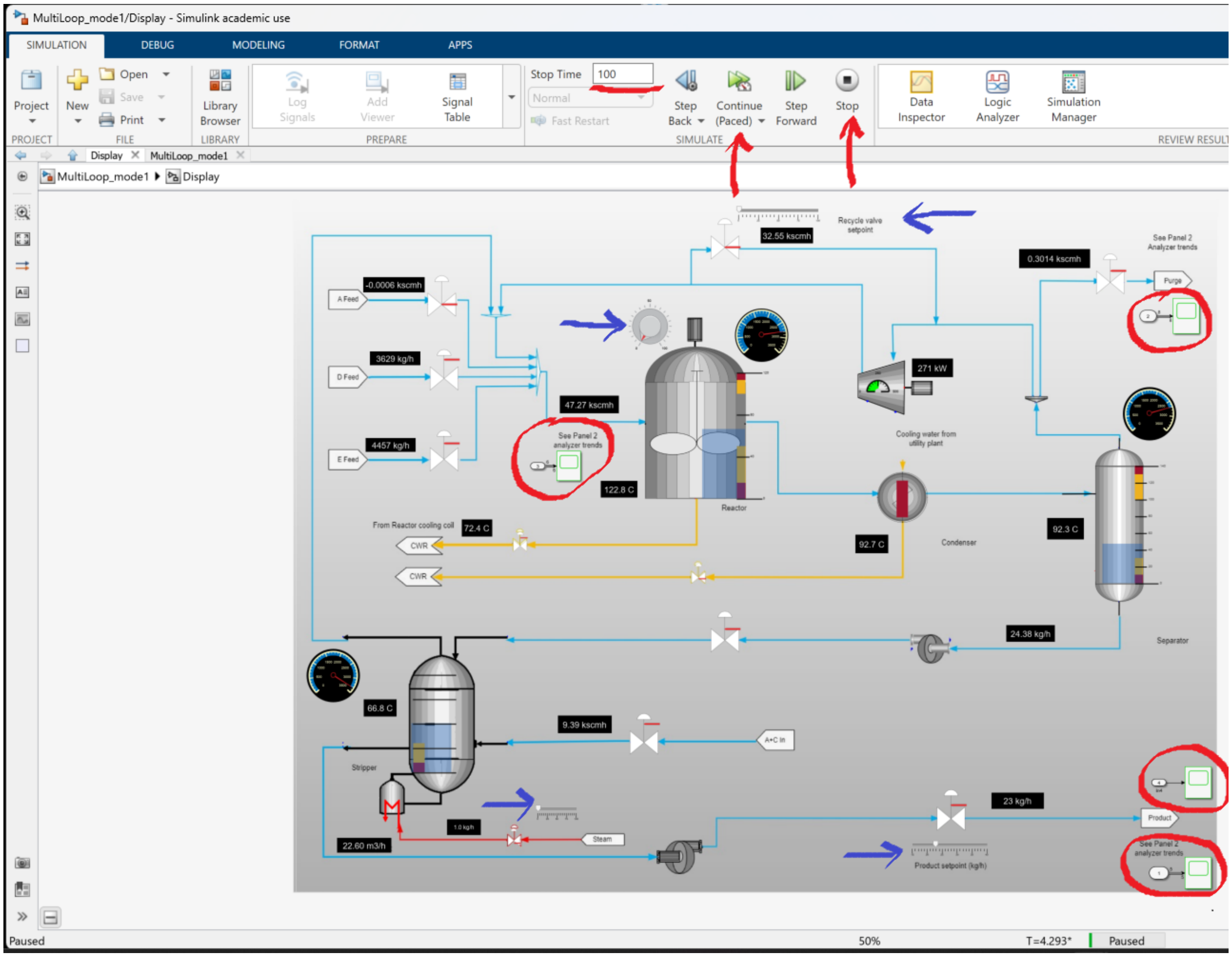
Instructions given to participants

After working with the process simulation, please complete the following 4 questions.

The objective of the quiz is to ***identify the root cause of the "abnormal situation".*** As part of communicating your understanding of the material presented, you are also requested to **suggest sensible steps to obtain further information** about the cause and/or **recommend a solution to the issue**.

The response fields are free-text, but point-form responses may be appropriate. Assume that the question graders are familiar with standard process control terminology and aim for clarity in responses. The technical content of your response is of interest.

You may start this quiz at any time. Interact with the simulation as you see fit, and note that it can be run in "steady state" and "abnormal situation" modes. There are tunable parameters (indicated with blue arrows in the image below) that allow you to alter some of the steady state characteristics, but altering these parameters may cause a "shut-down". There are trends available to monitor the gas concentrations and production rate, circled in red. The trends can be accessed by double-clicking the scope icons, and dismissed with the "x" in the top right corner. The simulation is set to run for 100 simulated hours (approximately 2 minutes), and you can alter the speed using "pacing".



Please notify the researcher if Simulink is not open on your computer.

Please note, the simulation files were prepared and validated by: Bathelt, A., Ricker, N. L., & Jelali, M. (2015). Revision of the Tennessee Eastman process model. *IFAC-PapersOnLine*, *28*(8), 309–314. https://doi.org/10.1016/j.ifacol.2015.08.199 and cited/attributed according to their published terms and conditions.

The simulator and decentralized control scheme were not modified for the steady-state portion of this study. The Simulink objects added to the simulator for this study are largely cosmetic and consist of the "Display" layer and the addition of several signal/display blocks for the purposes of constructing a realistic control panel display only. The abnormal situation model **does not** represent the performance of the decentralized scheme developed by Bathelt et al. Complete details on the changes made for the abnormal situation model will be described and made available as appendices in any published works referring to this study following completion of all sessions.

Theoretical basis for rubric (summarized)

Aspects contributing to complexity	Reconstruction	Marks to be awarded
Couplings and interconnections (Moray,1997; Wickens & Hollands, 2000; Vicente, 2007; Kluge, 2008)	**Acquisition of node data**	1 – any discussion of process, control equipment that appears in the simulation, accurate description of what was observed
	Acquisition of couplings (simple causal links)	2 – any description of a link between a single element and a trend or measurement, description of connected nodes (ie: reactor feeds condenser, purge valve fluctuations reflected in purge gas analyser trends)
	Identification of steady state characteristics (steady state mesh)	2 – description of basic characteristics, single nodes or individual trends
	Predict component effects in recurrent dynamic cases (alarm sequences, planned shut down)	2 – discussion of what might occur or what might have occurred, description of tunable parameter impact on process
	Elaborate, extrapolate and model complex causal relationships (downstream/knock on effects) perform root cause or fault detection on non-recurrent cases	5 – discussion that involves more than one control loop or multiple dependent trends. Accurately bases the elaboration or prediction on true observations of the simulation or theoretical knowledge
Dynamic effects (Vicente,1999; Kluge, 2008; Walker et al., 2010)	**Steady state**	1 – description of a system or subsystem not exhibiting transient behaviour
	Simple causality	2 – description of a simple subsystem or loop where an input causes a measurable output, single time-interval (valve opens, gas flow increases or reactor level decreases after separator level increased)
	Complex causality	5 – Explanation of a phenomenon involving multiple coupled elements, subsystems or control loops.
	Measuring decision utility	1 – description of a single, directly causal factor without substantial discussion around additional tests or diagnostics

		2 – description of a possible causal factor drawing dependent systems into the analysis, limited identification of additional steps or tests, simple recommendation that does not incorporate system-scale or multiple time-interval considerations 5- detailed and well supported root cause with one or more linked subsystems or components.
Non-transparency (Woods et al., 1990; Vicente, 1999;Funke, 2010; Kluge, 2014)	Incomplete schema Incomplete retrieval mechanism Incomplete interface Incorrect relationships implied	1 – description of a single measurement, variable or trend that appeared to be missing 2 – discussion of why a single variable, measurement or trend would improve understanding of either the process or control architecture 5- analysis of how an interface element caused a misleading or incorrect association to be made, or description of a missing element that was necessary to confirm root cause analysis, evidence of system or multi-component thinking. Recommendation considers more than the single affected unit and/or a longer time interval than the upset
MPC and RTO implementation	Knowledge acquisition of methods, algorithms, conceptual contents of MPC and RTO Integration of insights into the mental model	2 – description of objective function (production cost), trend 5 – extension of objective function into any controller or loop behaviour
Alarm flooding	Prioritizing, simplifying or decoding unexpected alarm sequences	2 – logical discussion that shows evidence of prioritizing or sorting multiple alarms or rapidly changing measurement values.

Rubric constructed with extensive reliance on: A. Kluge, S. Nazir, and D. Manca, “Advanced Applications in Process Control and Training Needs of Field and Control Room Operators,” *IIE Trans Occup*, vol. 2, pp. 121–136, 2014, doi: 10.1080/21577323.2014.920437.

Appendix 3

Manufacturer's marketing materials

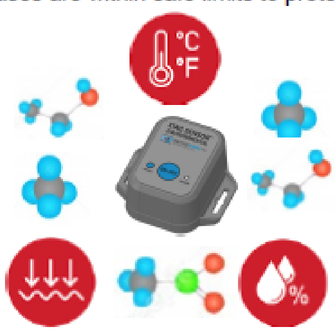
DEVICEWORX.COM TECHNOLOGIES INC.

Revision 1.0

Protect People, Crops, Animals & Equipment with the xTAG Environmental & xTAG Gas Sensors

Deviceworx xTAG Environmental Sensors accurately measure air temperature, pressure, and humidity. xTAG Gas Sensors measure the concentration of Volatile Organic Compound (VOC) and Volatile Sulphur Compound (VSC) gases in air. There are a vast array of measurable VOC/VSC gases including sewage gas, propane, natural gas (methane), construction off-gases, fire off-gases … The xTAG Gas sensor supports Artificial Intelligence (AI) and can be "trained" to measure any of these gases.

Use these xTAG Sensors to monitor the air within spaces to ensure temperature, humidity and dangerous gases are within safe limits to protect people.



Monitor the exposure of crops to damaging high temperatures and humidity within storage and shipping spaces. Recover crop damages by proving where and when they were exposed to harsh environments.

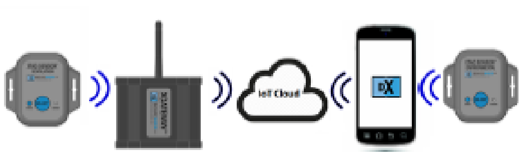
Protect livestock and pets by monitoring their exposure to damaging levels of temperature, humidity and dangerous gases.

Ensure equipment is healthy by collecting operational data such as bearing casing temperatures, filter upstream pressure (increases when clogged), exhaust gas composition etc.

Monitor fire types and burn rates. Specific fire fuels (e.g. grasses vs trees) exhaust specific VOCs. Measuring specific VOCs in smoke using the xTAG Gas Sensor can indicate what is burning as well as how fast.

Operators can review environmental and gas concentration data within Internet of Things (IoT) cloud-based reports. Cloud functions can also trigger alarm messages to Smartphones via email or text alerts whenever data crosses a safety threshold. Alarming can prove crucial in vacating a dangerous space saving the lives of people or animals within that space. Alarming and monitoring of equipment ensures that equipment faults are dealt with in a timely manner to avoid unplanned production stops that cost operations dearly. Sensor-logged crop batch environmental and location data is uploaded to the cloud for later review. Exposure of crops to poor environments can be addressed by process improvements or claims against a storer or shipper. Service providers can deploy xTAG Sensors and charge operators to view reports and receive alerts under a Software as a Service (SaaS) business model to dramatically increase revenue.

All xTAG Sensors can make wireless (long range Bluetooth) cloud connections using Deviceworx xGATEWAY IoT Gateways.



In some cases, simple, opportunistic acquisition of xTAG data to the cloud may be all that is required. In these situations, operators can use the xTAG Explorer app as an IoT Gateway. This app can collect sensor data over Bluetooth for cloud storage, reporting and alarming (using Wi-Fi or cell data).

Contact our sales team (sales@deviceworx.com) to learn more about how we can help address all of your environmental and gas monitoring challenges.

Deviceworx Technologies Inc.(www.deviceworx.com)

Page 1 of 1

1-888-409-2688

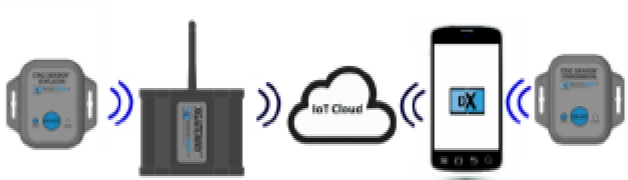
Manufacturer's sensor application materials

Improving Air Ventilation Systems To Reduce Covid-19 Transmission Using xTAG BLE Ventilation Sensors

Installing ventilation systems can have a significant effect on curbing the transmission of Covid-19 and keeping any environment healthy. Ventilation systems are only effective when working correctly. Monitoring their effectiveness requires measuring Carbon Dioxide (CO_2) within processed air at targeted locations and **comparing indoor CO2 to baseline outdoor CO2**. The difference between indoor and outdoor CO2 levels indicates how well a space is ventilated. Deviceworx new waterproof, extended temperature xTAG Ventilation Sensor design supports measuring CO2 indoors **and outdoors** to prove ventilation effectiveness. When the difference between indoor and outdoor CO2 is high, ventilation is insufficient and must be increased. When this difference is very low (often when spaces are empty) ventilation systems can be turned down or off dramatically increasing filter longevity and reducing energy use. xTAG Ventilation Sensors can substantially increase the safety of public and private spaces while ensuring that ventilation systems do not waste expensive energy.

Operators can review indoor, outdoor and differential CO2 levels within Internet of Things (IoT) cloud-based reports. Cloud functions can also trigger alarm messages to Smartphones via email or text alerts when indoor/outdoor differential CO2 levels get high and ventilation systems require attention. Fixes may include simply plugging them back into a wall socket, replacing their filter or simply increasing their output. Air ventilation service providers can charge operators to view reports and receive alerts under a Software as a Service (SaaS) business model and dramatically increase revenue beyond air ventilation hardware markups. SaaS revenue increases are the benefit to service providers while more effective ventilation and purification are the benefit to operators.

All xTAG Sensors can make wireless (long range Bluetooth) connections to the cloud via Deviceworx xGATEWAY IoT Gateway devices. xGATEWAYs can also be used to cloud-connect many disconnected air ventilation systems that have simple serial interfaces to support reporting on ventilation system health and, in some cases, providing for their remote control based on CO2 levels.



For simpler ventilation hardware (without a serial interface), direct on/off control that is based on indoor/outdoor differential C02 is possible using an xGATEWAY.

In some cases, simple, opportunistic acquisition of xTAG Ventilation Sensor CO2 data for storage and reporting in the cloud may be all that is required. In these situations, operators can use the xTAG Explorer app as an IoT Gateway. This app can collect sensor data over Bluetooth for subsequent cloud storage reporting and alarming (using a tablet Wi-Fi or cell data connection).

There are many approaches in monitoring air ventilation system effectiveness. Deviceworx can integrate xTAG Ventilation Sensors, xTAG Gas Sensors, xGATEWAYs, the xTAG Explorer app and 3rd party tools to meet the challenges presented in different installations. Contact our sales team (sales@deviceworx.com) to learn more about how we can help.

Selected sections of sensor manufacturer's data book

SENSIRION

SCD4x

Breaking the size barrier in CO_2 sensing

Features

- Photoacoustic NDIR sensor technology PASens®
- Smallest form factor: 10.1 x 10.1 x 6.5 mm³
- Reflow solderable for cost effective assembly
- Large output range: 0 ppm – 40'000 ppm
- Large supply voltage range: 2.4 – 5.5 V
- High accuracy: ±(40 ppm + 5 %)
- Digital I²C interface
- Integrated temperature and humidity sensor
- Low power operation down to < 0.4 mA avg. @ 5 V, 1 meas. / 5 minutes

Product Summary

The SCD4x is Sensirion's next generation miniature CO_2 sensor. This sensor builds on the photoacoustic NDIR sensing principle and Sensirion's patented PASens® and CMOSens® technology to offer high accuracy at an unmatched price and smallest form factor. SMD assembly allows cost- and space-effective integration of the sensor combined with maximal freedom of design. On-chip signal compensation is realized with the build-in SHT4x humidity and temperature sensor.

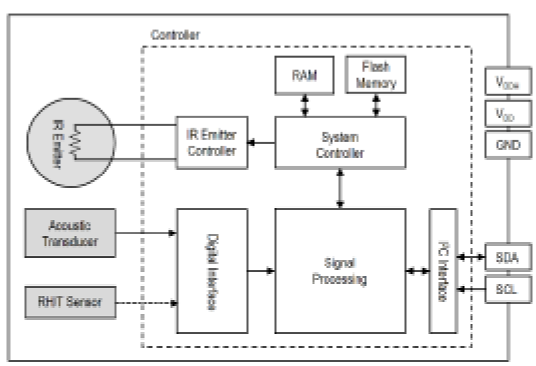
CO_2 is a key indicator for indoor air quality as high levels compromise humans' cognitive performance and well-being. The SCD4x enables smart ventilation systems to regulate ventilation in the most energy-efficient and human-friendly way. Moreover, indoor air quality monitors and other connected devices based on the SCD4x can help maintaining low CO_2 concentration for a healthy, productive environment.

Device Overview

Products	Details
SCD40-D-R2	Base accuracy, specified range 400 – 2'000 ppm
SCD41-D-R2	High accuracy, specified range 400 – 5'000 ppm, single shot mode supported
SCD42-D-R2	Separate datasheet, see https://sensirion.com/products/catalog/SCD42/

Full product list on page 22

Functional Block Diagram



1 Sensor Performance

1.1 CO_2 Sensing Performance

Default conditions of 25 °C, 50 % RH, ambient pressure 1013 mbar, default periodic measurement and 3.3 V supply voltage apply to values in the table below, unless otherwise stated.

Parameter	Conditions	Value
CO_2 output range[1]	-	0 – 40'000 ppm
SCD40 CO_2 measurement accuracy[2]	400 ppm – 2'000 ppm	± (50 ppm + 5% of reading)
SCD41 CO_2 measurement accuracy[2]	400 ppm – 5'000 ppm	± (40 ppm + 5% of reading)
Repeatability	Typical	± 10 ppm
Response time[3]	$\tau_{63\%}$, typical	60 s
Additional accuracy drift per year with automatic self-calibration algorithm enabled[4]	Typical	± (5 ppm + 0.5 % of reading)

Table 1: SCD40 and SCD41 CO_2 sensor specifications

1.2 Humidity Sensing Performance[5]

Parameter	Conditions	Value
Humidity measurement range	-	0 %RH – 100 %RH
Accuracy (typ.)	15 °C – 35 °C, 20 %RH – 65 %RH	± 6 % RH
	-10 °C – 60 °C, 0 %RH – 100 %RH	± 9 % RH
Repeatability	Typical	± 0.4 %RH
Response time[3]	$\tau_{63\%}$, typical	90 s
Accuracy drift	-	< 0.25 %RH / year

Table 2: SCD4x humidity sensor specifications

1.3 Temperature Sensing Performance[5]

Parameter	Conditions	Value
Temperature measurement range	-	- 10°C – 60°C
Accuracy (typ.)	15 °C – 35 °C	± 0.8 °C
	-10 °C – 60 °C	± 1.5 °C
Repeatability	-	± 0.1°C
Response time[3]	$\tau_{63\%}$, typical	120 s
Accuracy drift	-	< 0.03 °C / year

Table 3: SCD4x temperature sensor specifications

[1] Exposure to CO_2 concentrations smaller than 400 ppm can affect the accuracy of the sensor if the automatic self-calibration (ASC) is on.
[2] Deviation to a high-precision reference. Accuracy is fulfilled by > 90% of the sensors after calibration. Rough handling and shipping reduce the accuracy of the sensor. Sensor assembly temporarily reduces sensor accuracy. Accuracy is fully restored with FRC or ASC recalibration features > 5 days after sensor assembly. Accuracy is based on tests with gas mixtures having a tolerance of ± 1.5%.
[3] Time for achieving 63% of a respective step function when operating the SCD41 Evaluation Kit with default measurement mode. Response time depends on design-in, signal update rate and environment of the sensor in the final application.
[4] For proper function of ASC field-calibration algorithm SCD4x must be exposed to air with CO_2 concentration 400 ppm regularly. Maximum accuracy drift per year estimated from stress tests is ± (5 ppm + 2 % of reading). Higher drift values may occur if the sensor is not handled according to its handling instructions.
[5] Design-in of the SCD4x in final application, self-heating of the sensor and the environment impacts the accuracy of the RH/T sensor. To realize indicated specifications, the temperature-offset of the SCD4x inside the customer device must be set correctly (see chapter 3.6). Best RH/T accuracy is realized when operating the SCD4x in low power periodic measurement mode.

2 Specifications

2.1 Electrical Specifications

Parameter	Symbol	Conditions	Min.	Typical	Max.	Units
Supply voltage DC[6]	V_{DD}		2.4	3.3 or 5.0	5.5	V
Voltage ripple peak to peak	V_{RPP}				30	mV
Peak supply current[7]	I_{peak}	V_{DD} = 3.3 V		175	205	mA
		V_{DD} = 5 V		115	137	mA
Average supply current for periodic measurement	I_{DD}	V_{DD} = 3.3 V		15	18	mA
		V_{DD} = 5 V		11	13	mA
Average supply current for low power periodic measurement	I_{DD}	V_{DD} = 3.3 V		3.2	3.5	mA
		V_{DD} = 5 V		2.8	3	mA
Average supply current for periodic single shot measurement, 1 measurement / 5 minutes (SCD41 only)[8]	I_{DD}	V_{DD} = 3.3 V		0.45	0.5	mA
		V_{DD} = 5 V		0.36	0.4	mA
Input high level voltage	V_{IH}		0.65 x V_{DD}		1 x V_{DD}	-
Input low level voltage	V_{IL}				0.3 x V_{DD}	-
Output low level voltage	V_{OL}	3 mA sink current			0.66	V

Table 4 SCD4x electrical specifications

2.2 Absolute Maximum Ratings

Stress levels beyond those listed in **Table 5** may cause permanent damage to the device. Exposure to minimum/maximum rating conditions for extended periods may affect sensor performance and reliability of the device.

Parameter	Conditions	Value
Temperature operating conditions		-10 – 60°C
Humidity operating conditions[9]	Non-condensing	0 – 95 %RH
MSL Level		3
DC supply voltage		- 0.3 V – 6.0 V
Max voltage on pins SDA, SCL, GND		– 0.3 V to V_{DD}+0.3 V
Input current on pins SDA, SCL, GND		- 280 mA to 100 mA
Short term storage temperature[10]		- 40°C – 70°C
Recommended storage temperature		10 °C – 50 °C
ESD HBM		2 kV
ESD CDM		500 V
Maintenance Interval	Maintenance free when ASC field-calibration algorithm[11] is used.	None
Sensor lifetime[12]	Typical operating conditions	> 10 years

Table 5: SCD4x operation conditions, lifetime and maximum ratings

[6] Supply voltage must be kept stable during sensor operation
[7] Power supply should be designed with respect to peak current.
[8] On-demand measurement with freely adjustable interval. See chapter 3.10
[9] Accuracy can be reduced at relative humidity levels lower than 10 %.
[10] Short term storage refers to temporary conditions during e.g. transport.
[11] For proper function of ASC field-calibration algorithm the SCD4x has to be exposed to clean air with 400 ppm CO_2 concentration regularly.
[12] Sensor tested over simulated lifetime of > 10 years for indoor environment mission profile

Sensor calibration certification

Calibration Certification – **CO_2 Sensors**

SENSIRION
THE SENSOR COMPANY

Calibration Certification

Name and address of the manufacturer: Sensirion AG
Laubisruetistrasse 50
CH-8712 Switzerland

Description: *CO_2 Sensors*

SCD4x

The above-mentioned products are calibrated to meet the specifications according to the corresponding Sensirion data sheet. Each device is individually tested after its calibration.

Sensirion uses transfer standards for the calibration. These transfer standards are themselves subject to a scheduled calibration procedure. The calibration of the reference itself used for the calibration of the transfer standards is performed in accordance with ISO9001:2015.

The accreditation bodies of our references are full members of the International Laboratory Accreditation Cooperation (www.ilac.org). Calibration certificates issued by facilities accredited by a signatory to the ILAC Mutual Recognition Arrangement (MRA) are accepted by all signatories to the ILAC MRA.

Staefa, 16 May 2022

Stephan Weber,
Director,
Head of Quality Management, Sensirion AG

Volker Born
Manager,
Head of Quality Engineering, Sensirion AG