

Understanding and Mitigating the Effects of Operating Room Distractions

by

Suzan Ayas

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science

Mechanical and Industrial Engineering
University of Toronto

© Copyright by Suzan Ayas 2019

Understanding and Mitigating the Effects of Operating Room Distractions

Suzan Ayas

Master of Applied Science

Mechanical and Industrial Engineering
University of Toronto

2019

Abstract

Operating rooms (ORs) are complex dynamic environments with frequent distractions. Literature has limited evidence on OR distractions leading to adverse surgical outcomes; in fact, studies increasingly agree that distractions under certain circumstances can be beneficial. A notable limitation of existing studies is the drawbacks of data collection methods. This study is the first to use a naturalistic dataset, the OR Black Box, to investigate the relation of OR distractions and intraoperative technical events. Increased machine alarm rates, decreased technical skill levels, and increased body mass index levels were found to be associated with increased likelihood of observing more than one severe technical event during high criticality/medium-high difficulty procedural steps of Roux-en-Y gastric bypass operations. Mitigation strategies to reduce the negative effects of OR distractions were identified through a literature review and discussions with OR members. A taxonomy was created through systems perspective to support the selection of appropriate mitigation strategies.

Acknowledgments

I would like to thank my supervisor, Professor Birsen Donmez, for guiding me with her knowledge and experience with endless patience. I am deeply grateful for your support and supervision. I was able to learn many things from you in these two years. And I hope to be a better student during my PhD.

Secondly, I would like to thank my committee members, Dr. Teodor Grantcharov and Professor Mark Chignell, for their time, support, and guidance.

I would like to thank Dr. Grantcharov again, but this time for providing this amazing research opportunity. What you have created has inspired me to continue doing research.

Members of SST and ICSS, thank you all for the friendly environment you have created. Karthik, Alex, and my desk neighbor Amalia, thank you for your support and friendship from day one. Amar, thank you for patiently sending me the files I need. Thank you, Lauren, for teaching me so much in this short period of time; I look forward to learning more from you.

Friends back home, I miss you all a lot. As I always say on the phone, I am waiting for all of you to visit me in Toronto. Friends here at my new home, it is a great pleasure to explore the city and to share experiences with you. Ghazaleh and Mandana, I wouldn't be able to survive the winters without your friendship. Friends from HFASt lab, thank you all your support, feedback, and friendship (and coffee – thank you Mehdi).

Nazli, thank you for being my twin. This has been a huge journey for both of us, and I don't think I could be here without your friendship. I look forward for our future friendship.

Most importantly, I want to thank my family, my mum, my dad, and my brother, for their endless support from far away. Despite the distances, I can always feel your love and support with me. Without you, I couldn't have come anywhere near where I am right now. Thank you.

Table of Contents

Acknowledgments.....	ii
Table of Contents	iv
List of Tables	vi
List of Figures	viii
List of Appendices	ix
Chapter 1 Introduction	1
1.1 Background.....	2
1.1.1 Cost of Preventable Surgical Errors	2
1.1.2 Definitions of Distraction, Interruption, and Technical Event.....	5
1.1.3 Distraction and Attention Theories	5
1.1.4 Distraction Studies in the OR.....	8
1.1.5 Relationship between Distractions and Surgical Errors	11
1.1.6 Benefits of Distractions.....	15
1.2 Objectives	18
Chapter 2 Quantification of OR Distractions and their Effects using Naturalistic Data	20
2.1 Background.....	22
2.1.1 Description of Dataset used in Analysis	24
2.1.2 Model Construction.....	28
2.1.3 Data Processing for Analysis	29
2.2 Findings	32
2.2.1 Comparisons across Procedure Categories	37
2.2.2 Logistic Regression Analyses	38
2.3 Discussion	42
2.3.1 Limitations	46
2.3.2 Suggestions to Improve the Distraction Coding Framework	48
Chapter 3 A Taxonomy of Distraction Mitigation Strategies for Operating Rooms	51
3.1 Introduction.....	51
3.2 A Taxonomy of OR Distraction Mitigation Strategies	52
3.2.1 Hard Technology, Feedforward Control.....	54
3.2.2 Hard Technology, Concurrent Control	55

3.2.3	Hard Technology, Feedback Control	56
3.2.4	Soft Technology, Feedforward Control	57
3.2.5	Soft Technology, Concurrent Control.....	58
3.2.6	Soft Technology, Feedback Control	59
3.3	Discussion.....	59
Chapter 4 Conclusions and Future Work.....		61
Limitations and Future Work.....		63
References.....		65
Appendices.....		75
Copyright Acknowledgements.....		86

List of Tables

Table 1. Descriptive statistics are shown in the table for each machine alarm types... ..	35
Table 2. Descriptive statistics for surgery and across procedural step categories.	36
Table 3. Descriptions and measurements of variables used in statistical analyses	39
Table 4. Descriptive statistics of distractions, surgical team factors, and patient factors with respect to the severe technical event categories used in the dependent variable.	40
Table 5. Logistic regression results for predicting severe technical events through distractions and surgical team factors.....	41
Table 6. Logistic regression results for predicting severe technical events with patient factors..	42
Table C-1. Descriptive statistics on severe technical events	77
Table G-1. Logistic regression results for predicting severe technical events through people entering/exiting	81
Table G-2. Logistic regression results for predicting severe technical events through machine alarms	81
Table G-3. Logistic regression results for predicting severe technical events through external communications	81
Table G-4. Logistic regression results for predicting severe technical events through changeovers	82
Table G-5. Logistic regression results for predicting severe technical events through OSATS average	82
Table G-6. Logistic regression results for predicting severe technical events through BMI.....	82
Table G-7. Logistic regression results for predicting severe technical events through previous abdominal surgery information.....	82

Table H-1. Variance inflation factor (VIF) of distractions (people entering/exiting, machine alarms, and external communications) and surgical team factors (changeovers and OSATS average)..... 83

Table I-1. Variance inflation factors (VIF) of patient factors (BMI levels and previous abdominal surgery) 84

Table J-1. Descriptive statistics for machine alarms broken down with respect to their durations and the procedural step category..... 85

List of Figures

Figure 1. Dimensions of information processing according to multiple resource theory (recreated from Wickens, 2008b)..... 7

Figure 2. Illustration of Yerkes-Dodson law (recreated from Yerkes & Dodson, 1908). 16

Figure 3. Data exclusion process represented by a flow chart..... 25

Figure 4. Illustration of three system factors (procedural, surgical team, patient) involved in the relationship between distractions and severe technical events 29

Figure 5. An example operation timeline showing the procedural step sequences (LLM, HMH, LML). OSATS are scored over 20-minute blocks. ti's indicate periods during which a 20-minute block and an HMH procedural step (shown in blue) overlaps..... 31

Figure 6. Boxplots showing the distributions and average values of severe technical events, distractions across procedural step categories. 37

Figure 7. Taxonomy of OR distraction mitigation strategies..... 54

Figure D-1. Number of observed door activities (door opening/closing) with respect to their duration. 78

Figure E-1. Number of machine alarms observed that took between 1 second and 15 seconds. . 79

Figure E-2. Number of machine alarms observed that took between 3 seconds and 30 seconds. 79

Figure F-1. Normal Q-Q plots for distractions. 80

List of Appendices

Appendix A: Generalized Error Rating Tool	75
Appendix B: Global 5-point scale rating for Objective Structured Assessment of Technical Skills.....	76
Appendix C: Distribution of Severe Technical Events.....	77
Appendix D: Door Activity: Duration Analysis	78
Appendix E: Machine Alarms: Duration Analysis	79
Appendix F: Normal Q-Q Plots for Distractions	80
Appendix G: Univariate Analysis: Results	81
Appendix H: Multicollinearity Results for Distractions and Surgical Team Factors	83
Appendix I: Multicollinearity Results for Patient Factors	84
Appendix J: Descriptive Statistics of Door Opening/Closing and Machine Alarms with Respect to Procedural Step Categories	85

Chapter 1

Introduction

In the complex and dynamic setting of operating rooms (ORs), distractions occur frequently: phones ring; alarms beep; doors open and close; side conversations take place. These frequent distractions can interrupt clinicians' work, which may put patient safety at risk by leading to medical errors (Kohn, Corrigan, & Donaldson, 2000). In the last two decades, in order to improve patient safety and minimize medical errors, research in healthcare had increased its focus on the prevalence and effects of distractions. Even though research has advanced over the years, the relationship between distractions and intraoperative technical events (or more simply technical events) lacks sufficient evidence. Instead, reporting the frequency of distractions has been a common approach (Coiera, 2012). Moreover, the limited number of studies that tried to understand the relation between distractions and technical events has data collection limitations such as observer bias in observational studies, unrealistic conditions and unrepresentative participant selection in controlled experiments, and limited sample size for both methods. Despite the limited evidence, recommendations taken from other high-risk industries such as aviation have been proposed to eliminate distractions within ORs (e.g., Crew Resource Management (CRM) skills; McConaughy, 2008, checklists; Hales & Pronovost, 2006, and sterile cockpit rule; Wright, 2016). However, eliminating distractions from healthcare settings, including ORs, may actually deteriorate clinical performance as some distractions may be necessary and beneficial for better patient care (Rivera-Rodriguez & Karsh, 2010; Sasangohar, Donmez, Trbovich, & Easty, 2012). For example, distractions, such as case-irrelevant conversations, can potentially lower mental fatigue and stress in ORs (Weigl, Antoniadis, Chiapponi, Bruns, & Sevdalis, 2015) and can help surgeons bring their attention back to the surgical task (Moulton, Regehr, Lingard, Meritt, & MacRae, 2010). In order not to remove distractions when they might be beneficial to surgical performance, mitigation techniques should be developed through a systems approach. A taxonomy of OR distraction mitigation strategies can help guide distraction management within a systems approach to understand the conditions that OR distractions can be harmful or beneficial.

This thesis investigates the effects of distractions experienced in the ORs. First, the existing literature on OR distractions and surgical error relationship is reviewed, and the

limitations of the current literature is highlighted. Second, to understand the OR distraction and technical event relationship while addressing some of the existing data collection limitations, analysis conducted on a naturalistic OR dataset, OR Black Box, is presented. Third, a taxonomy of OR distraction mitigation strategies is proposed along with the support of literature and discussions with OR team members in our partner hospital where the OR Black Box data is housed. Lastly, limitations of this research and suggestions for future work are discussed.

1.1 Background

1.1.1 Cost of Preventable Surgical Errors

Medical errors, defined as “the failure of a planned action to be completed as intended (i.e., error of execution) or the use of a wrong plan to achieve an aim (i.e., error of planning)” (Kohn et al., 2000; p. 28), are common and significantly costly. A study that used medical insurance claim data estimated the measurable cost of medical errors (a preventable adverse outcome that results from improper medical management) to be US\$17.1 billion in 2008 (Van et al., 2011), of which \$5 billion was estimated to be surgery related. A technical report by the Canadian Patient Safety Institute (CPSI) estimated the cost of preventable adverse events in the Canadian acute care system to be CAD\$397 million for 2009-2010 (Etchells et al., 2012); CAD\$24 million of this estimated cost was attributed to surgical safety infections (SSI) only in one year. Another Canadian study estimated that in the fiscal year of 2000 in Canada, 7.5% of patients experienced at least one adverse event, and 36.9% of these adverse events were preventable, and the majority (51.4%) were surgery related (Baker et al., 2004). The authors noted that in one year, these rates would correspond to 70,000 patients who experience potentially preventable adverse events with consequences such as longer hospital stays, disability, or death.

These estimated costs of medical errors may not truly reflect the actual costs since the studies cited above were limited by the availability of adverse event reports and the amount of details reported; further, the overall cost estimates did not account for indirect costs such as negative physical and psychological effects on the patients, their families and the clinicians. Nevertheless, these reports show that medical errors can cost societies large amount of financial resources and most importantly, the well-being of their members. On the other hand, medical

errors are preventable, and their associated costs are avoidable through an evidence-based understanding of the underlying problems.

The solution to preventing errors is not by blaming individuals but by building defense barriers within the system against human error (Reason, 2000). High-risk industries, such as aviation and nuclear power, manage human error by acknowledging that humans are fallible and will make errors. However, in healthcare, clinicians in general are expected to make no errors, and can be blamed even when errors occur due to tasks exceeding human capabilities, for example, by overreliance on human memory (Leape, 1994). To solve this problem, either the tasks should be redesigned and/or there should be other mechanisms that act as defense barriers against medical errors such as standardization of storage location of equipment and supplies (Leape, 1994). To inform such changes, the conditions that can lead to medical errors should be investigated. A substantial attempt in understanding medical errors was reported by the Institute of Medicine in 2000, in “To Err is Human” (Kohn et al., 2000). Human and system factors contributing to medical errors were identified, and the healthcare system shortcomings were presented with improvement recommendations. One contributing factor to medical errors mentioned in this report was distractions, and distractions were recommended to be reduced. To understand the factors that may risk surgical safety, a later review article listed system factors that could contribute to adverse surgical outcomes (Vincent, Moorthy, Sarker, Chang, & Darzi, 2004): patient factors, surgical team (e.g., experience with working together), processes and procedures, key operative events (e.g., blood loss, minor and major complications, error compensation and recovery), flow of information following patient (e.g., handovers), technical skills, team performance and leadership, decision making and situation awareness, operative environment (e.g., availability and adequacy of equipment, availability of notes and records, noise and lighting, distractions), and finally, interruptions (e.g., phone calls).

As distractions and interruptions were identified to be problematic, many proceeding studies investigated the prevalence of distractions and more specifically interruptions in healthcare settings and methods to reduce them (Grundgeiger & Sanderson, 2009; Mentis, Chellali, Manser, Cao, & Schwaitzberg, 2016; Raban & Westbrook, 2014; Rivera-Rodriguez & Karsh, 2010). Rivera-Rodriguez and Karsh (2010) conducted a systematic review on interruptions (i.e., distractions that result in a break in primary tasks) in healthcare settings, including intensive care units (ICUs), emergency departments (EDs), pharmacies, and ORs. The

authors concluded that interruptions are common regardless of the setting; however, not all interruptions can be or must be eliminated. They added that certain interruptions may be necessary for patient safety. Another review of interruption studies performed for various healthcare settings (EDs, ORs, ICUs, and general wards) concluded that there was insufficient evidence on the causal link between interruptions and medical errors, and the effects of interruptions on the prospective memory (Grundgeiger & Sanderson, 2009) which is the “memory for intentions” (Reason & Hobbs, 2003, p. 44). Another systematic review, this time focusing on mitigating the effects of interruptions on medication administration errors, found weak evidence on the effectiveness of the relevant mitigation strategies that have been explored in the literature (Raban & Westbrook, 2014); the authors also reported that not all interruptions were harmful and some were necessary for patient safety.

Mentis et al. (2016) conducted a systematic review specifically on OR distraction studies and investigated the most frequent and most severe distractions; movements around monitors and movements of OR doors, case-irrelevant communication, phone calls/pagers, radio, procedure-related, and equipment-related distractions were identified as the six most common sources of distraction (Mentis et al., 2016) and the rates of OR distractions were reported to range from 6 (Sevdalis et al., 2014) to 33 distractions per hour (Persoon, Broos, Witjes, Hendriks, & Scherpbier, 2011). However, the authors also acknowledged two simulation studies (Conrad et al., 2012; Siu, Suh, Mukherjee, Oleynikov, & Stergiou, 2010) that found improved time to task completion, economy of motion, and accuracy with the presence of background music.

Overall, distractions are common in the OR, however, some of them are inherent to the OR tasks, and thus are necessary and potentially useful to improve surgical performance. For example, phone calls may be necessary during operations to consult other surgeons for improved surgical outcomes, and during routine tasks, OR team members may engage in conversations or listen to music to keep alert and focused. However, further research is needed to identify which distractions can be eliminated from the OR entirely, which distractions may have to be eliminated under certain conditions, which distractions may have to be delayed for more opportune times, and which distractions can be introduced to enhance alertness. To this end, further evidence is needed on the relations between distractions, context, and technical events.

1.1.2 Definitions of Distraction, Interruption, and Technical Event

A major challenge in studying distractions in healthcare is that the definitions of distractions and interruptions vary across studies and this variation hinders the knowledge accumulation and comparison across studies (Coiera, 2012; Grundgeiger et al., 2016; Hopkinson & Jennings, 2013; McCurdie, Sanderson, & Aitken, 2016). This thesis will adopt the relatively widely used definitions of distractions and interruptions by Healey, Sevdalis and Vincent (2006). A distraction is defined as an external source that leads to an observable break in attention, and an interruption is defined as a special case of distraction with an observable break in the performance of the primary task. It is important to recognize that the definition of distraction is limited and only identifies a source to be a distraction when there is an observable break in attention – an artifact of observational study methods. Even if a break in attention is not observable, a source can be distracting, i.e., it can interfere with the person's attention to a level that impairs their performance. Thus, this thesis expands the definition of distraction and includes sources that can interfere with attention.

As stated earlier, one of the objectives of this thesis is to understand the relation between OR distractions and intraoperative technical events, while considering context. Intraoperative technical events, or also referred to in this thesis as technical events, are defined as 'any deviation from usual medical care that causes an injury to the patient or poses a risk of harm' (World Health Organisation, 2005, p. 9). Bleedings, thermal injuries, and mechanical injuries are some examples of technical events. These events can be due to error modes such as too much/little force or inadequate visualization (Bonrath, Zevin, Dedy, & Grantcharov, 2013), which may result due to distractions interfering with surgeon's attention. The following section introduces relevant theories of attention that can explain why certain types of distractions may lead to errors whereas others may not.

1.1.3 Distraction and Attention Theories

According to the capacity model of attention (Kahneman, 1973), humans have limited attentional resources for mental activities. Each mental activity competes for attentional resources and when an activity's attentional demand is not met, the individual performs poorly on that activity or fails completely. One example from aviation is the crash of Continental Airlines flight #1943 in 1996 (Dismukes, Berman, & Loukopoulos, 2007). It was reported that

the captain engaged in non-essential conversation while preparing for landing and omitted a critical item in the landing preparation checklist. The flight crew realized that there was an issue right before landing but could not diagnose it amongst loud warning alarms. During these high workload and stressful moments, another checklist was completely neglected. As a result, the plane landed with physical damage and minor passenger injuries. This example illustrates that distractions can contribute to the omission of known tasks (e.g., forgetting to execute items on checklists known as lapses; Reason, 1990) specifically when working memory is overloaded due to task demands or its capacity reduced due to stress (Wickens, Hollands, Banbury, & Parasuraman, 2000). Similarly, attending to a distraction in the OR can tax an individual's cognitive resources, and if there is no spare capacity for the processing of distractions or the person is unable to ignore them, the surgical task performance may be negatively affected.

Certain surgical tasks may require the majority of an individual's cognitive resources, leaving little reserve for other mental activities. In order to successfully complete such surgical tasks, irrelevant environmental cues may need to be filtered out. Filtering out such cues and allocating cognitive resources on a single task is achieved by focused attention. Some environmental cues may break focused attention and divert it away from the task. This attentional diversion is governed by factors that affect selective attention, i.e., selecting what to attend to, such as cue saliency, effort, expectancy, and value (known as the SEEV model for visual attention allocation; Wickens, 2008a). Saliency refers to the amount of attention a cue draws. For example, a tray dropping in the OR can immediately and automatically divert OR team members' attention because loud noise is salient. When it is less effortful to attend to a cue, the person is more likely to attend to it. For example, surgeons may be more likely to attend to information sources that are in front of them as opposed to behind them. Expectancy and value are top-down factors that guide selective attention. If a surgeon is expecting to find a particular type of information on a display, they may guide their attention towards that display; and a surgeon may choose to attend to a machine alarm instead of a side conversation because the information conveyed by the alarm may have higher value for patient safety.

If there are sufficient cognitive resources, multiple tasks can be simultaneously performed. For this however, attention needs be allocated to multiple mental activities, requiring divided attention. Task performance during multitasking can be explained by the multiple resource theory (Wickens, 2008b). The multiple resource theory states that if two tasks are

competing for the same type of resources, the interference between them would be higher compared to two tasks that do not compete for same type of resources. The theory lists four information processing dimensions to define resources (Figure 1): stages of processing, codes of processing, modalities, and visual channels. It should be noted that even if two tasks do not share any common resources on these dimensions, there can still be interference between the two tasks. In the OR context, distractions can compete with surgical tasks for the same limited cognitive resources and can potentially degrade the surgical task performance. Conversely, surgical tasks and distractions can be time shared if they do not require similar cognitive resources. Therefore, distractions must be evaluated while taking the demands of both surgical tasks and distractions into account.

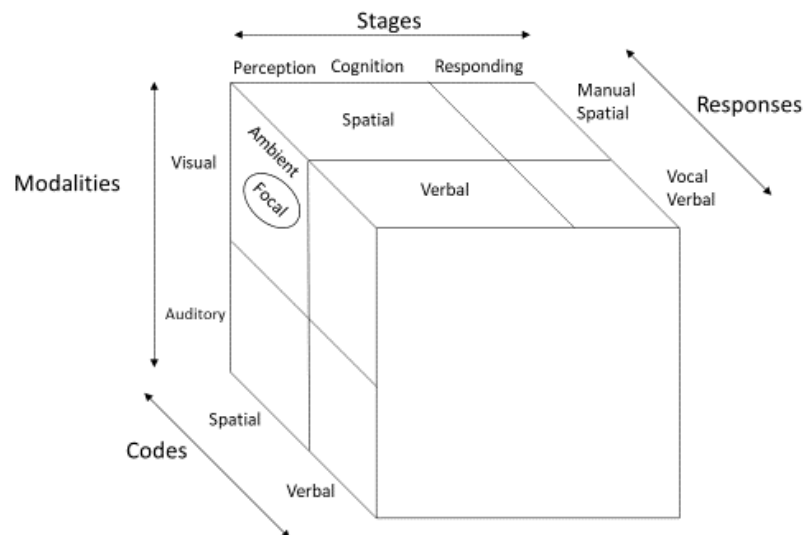


Figure 1. Dimensions of information processing according to multiple resource theory (recreated from Wickens, 2008b). Stages of processing shows different resources used for perception and cognition, and responding (i.e., selecting and executing actions). Codes of processing shows the different resources used during perception and cognition (spatial and verbal) and during responding (manual-spatial and vocal-verbal). Modalities differentiate the auditory and visual resources used for perception. Visual channels dimension separates focal and ambient processing in visual perception.

1.1.4 Distraction Studies in the OR

In the last two decades, various studies investigated the effects of distractions on surgical outcomes. A comprehensive human factors study looked at 243 neonatal arterial switch operations performed at 16 hospitals across the United Kingdom to understand the patient outcomes of adverse surgical events (de Leval, Carthey, Wright, Farewell, & Reason, 2000). The researchers grouped observed surgical events into two groups: major events that are “failures that were likely to have serious consequences for the safety of the patient” and minor events that are “failures that disrupted the surgical flow of the procedure but which, in isolation, were not expected to have serious consequences for the safety of the patient” (de Leval et al., 2000, p. 662). In their findings, both the number of major events observed per surgery if not rectified and the number of minor events observed per surgery if not rectified were found to be related to death and/or near misses. Minor events included (1) events that this thesis considers as interruptions based on the definition adopted, (2) events labelled distracters that were not explicitly defined by the authors, and (3) events that are not related to distractions or interruptions such as poor fasting policy. The frequency of all minor events was summed without separating interruptions and distracters from other minor event types, thus, the negative findings obtained for minor events in this study cannot be directly attributed to distractions or interruptions.

To investigate OR distractions using a systematic approach, Healey, Sevdalis, and Vincent (2006) created an observational tool to capture OR distractions (and interruptions) and measure their intraoperative interference levels on OR team members. The distractions that this tool captures include ten items: (1) phone, (2) bleeper (pager), (3) radio, (4) case irrelevant communication, (5) communication difficulties, (6) external staff, (7) equipment – not at hand or failing, (8) work environment – workspace and human–interface problems, (9) procedural – events intrinsic to the case work, and (10) movement in front of or behind video monitors for laparoscopic cases only. The distraction interference rating scales consist of 9 levels: scores 1-3 correspond to potential distractions that are either ignored or a circulating nurse attending the distractions, scores 4-6 indicate an observably distracted individual, scores 7-9 indicate two or more members distracted and the surgical workflow interrupted. Although interference ratings measure the observable effects of distractions, these ratings do not capture technical events but rather assess the number of team members who are visibly affected from distractions.

Using this tool, Healey et al. (2006) studied a single OR with two observers attending 29 laparoscopic and 21 open surgeries, and reported 13.6 distractions per case and 0.29 distractions per minute. The most frequent distractions observed were visual distractions behind the laparoscopic monitors (5 times per case, with interference level of 5.52), followed by beepers (3.31 times per case, with interference level of 8.56), and then surgeon initiated case irrelevant conversations (1.88 times per case, with interference level of 8.18), while the least frequent distraction was related to engaging with or reacting to radio once every case (with interference level of 3.67). The highest interference was from external staff; external staff disrupted the surgical flow 1.44 times per case with 11.2 interference level. However, as stated earlier, these interference ratings did not capture actual technical events, such as bleeding during surgery. Further, the interference ratings did not distinguish distractions that occur during different surgical steps. One individual being distracted during a critical phase of the surgery may be more detrimental to patient safety than multiple team members being distracted during a routine phase. Therefore, not taking into account surgical task demands more explicitly neglects the interactions between task demands and distractions. Some other limitations for this observational study, as the authors also noted, were difficulties of logging and categorizing frequent distractions that occurred in short time intervals. Despite these limitations, this study introduced a systematic approach to collect distraction-related data within the OR. Additionally, the findings showed that distractions can be frequent within the OR, and certain type of distractions can interfere with surgical flow more than others.

The tool developed by Healey et al. (2006) was used in many other observational studies (e.g., Antoniadis, Passauer-Baierl, Baschnegger, & Weigl, 2014; Healey, Primus, & Koutantji, 2007; Jothiraj, Howland-Harris, Evley, & Moppett, 2013; Persoon et al., 2011; Sevdalis et al., 2014; Weigl et al., 2015, 2018; Wheelock et al., 2015). These studies recorded a wide range of distractions in ORs ranging from 6 (Sevdalis et al., 2014) to 33 distractions per hour (Persoon et al., 2011). However, none of these observational studies was able to investigate the relation between distractions and technical events. They captured interference levels of distractions, and the highest interference levels were observed for distractions related to external staff (Healey et al., 2006), equipment (Antoniadis et al., 2014; Healey et al., 2007; Jothiraj et al., 2013; Persoon et al., 2011; Weigl et al., 2015, 2018), and communication and coordination (Sevdalis et al., 2014; Weigl et al., 2018). The variability in the findings of these studies suggest that there might

be large individual differences among the ORs studied or methodological differences in the operationalization of the data collection tool. Further, Weigl et al. (2018) showed that different stages of robotic-assisted radical prostatectomies were affected by different types of distractions: high-severity surgical flow disruptions were communication and coordination problems during pre-robot and docking phases, and equipment problems and communication during console time. Therefore, as suggested earlier, task demands need to be considered when examining the distraction-technical event relationship.

Interestingly, most frequent distractions were usually found to have the least interference on the OR teams (Mentis et al., 2016). For example, Healey et al. (2006) found that bleepers, phone calls, conversations, and visual distractions due to movements behind or front of the video monitor for laparoscopic operations were some of the most frequent ones, yet they did not have a high interference score since the circulating OR members dealt with these distractions. With only one member being visibly affected, these distraction categories received lower interference scores. However, even one team member being distracted may lead to a technical event. Further, although some distractions may not visibly disrupt the surgical flow, they may have effects on the surgical performance that are not observable. For example, noise can interfere with the verbal rehearsal of information in working memory (Wickens et al., 2000). Although noise such as machine alarms or background music may not generate an observable break in attention, it can thus impair cognition. In a controlled laboratory study, alarms were found to degrade novice surgeons' surgical performance only when the alarms were combined with a cognitive distraction, suggesting that alarms can have negative effects when reserve cognitive resources are limited (Ahmed, Ahmad, Stewart, Francis, & Bhatti, 2015). Thus, although the interference rating scale is informative, it is insufficient to study the effects of distractions on surgical safety.

One observational study investigated the effects of conversations and other auditory distractions on surgical-site infections (SSIs) by observing 167 elective open abdominal surgeries conducted in a Swiss university hospital (Tschan et al., 2015). Conversations were recorded as case-relevant and case-irrelevant among the sterile team and between the sterile team and the anesthesiologists. Case-relevant conversations were related to patient-at-hand or the procedures, while case-irrelevant conversations were any other type of conversations. Other auditory distractions recorded included noise produced by the non-sterile OR team members (e.g., when opening packages), OR traffic, and side conversations among the non-sterile OR

team members. SSI information was collected both through phone calls to the patients 30 days after the surgery, and by asking general practitioners to confirm the presence of SSIs and to classify SSIs as organ/space SSIs and incisional SSIs. Patient and procedure characteristics were taken from the patient records and included in the analyses. Twenty-four out of 167 cases had SSIs. It was found that there were 19.2 case-relevant conversations, 6.2 case-irrelevant conversations, 10.2 noises, 31.8 door openings, and 10.5 side conversations per hour. Due to the limited number of SSIs, the authors could only conduct univariate analyses using these factors and found that increased number of case-relevant conversations during the entire case was associated with fewer organ/space SSI incidents, while increased number of case-irrelevant conversations during closure was associated with increased incisional SSI incidents. A similar study examined how SSIs were affected by OR traffic (Young & O'Regan, 2010). The authors used electronic counters to count the number of OR door openings and closings for 46 cardiac cases in one hospital in the UK and collected patient demographics and operative details. Although they could not directly examine the relationship between OR traffic and SSIs due to small sample size, authors found that door opening and closing frequency increased for high complexity procedures and those procedures performed infrequently. They inferred that because door openings would disrupt the airflow, the risks for SSI would increase.

In general, observational studies fall short in investigating the effects of distractions on technical events due to data collection limitations of real-time human observations. In order to assess a technical event, medically trained observers are required; however, observational studies require great amount of time from such individuals. Because of such limitations, studies that look into the relationship between distractions and technical events have mostly been controlled laboratory experiments. In fact, only one observational study was found in the literature that tried to assess this relationship. Next two sections will present literature findings on the effects of distractions on technical events identified through controlled laboratory experiments and this observational study.

1.1.5 Relationship between Distractions and Surgical Errors

Two experimental studies showed that novice surgeons took longer to complete surrogate surgical tasks when presented with arithmetical questions as cognitive distractions (Ahmed et al., 2015; Goodell, Cao, & Schwaitzberg, 2006). Ahmed et al. (2015) found a significant increase in

error scores in a dissection task on a mastoidectomy simulator with cognitive distraction, whilst Goodell et al. (2006) did not find a significant relationship between cognitive distractions and errors for laparoscopic tasks (e.g., acquire–place, withdraw–insert, diathermy, manipulate–diathermy, stitch start, half square knot) on a virtual reality simulator. However, these findings fall short in demonstrating the performance of experienced surgeons under distractions since both studies had novice surgeons as participants. A similar laboratory study that used arithmetic questions as distractors using 31 novice and 9 experienced surgeons found that under distractions, novices had significantly impaired arithmetic task performance but not the surrogate surgical task performance (i.e., a laparoscopic peg transfer task), while experienced surgeons had no task performance deteriorations for either task (Hsu, Man, Gizicki, Feldman, & Fried, 2008). Ahmed et al. (2015) also noted that even within their participant group of novice surgeons, more experienced residents were less affected by distractions. These results suggest that increased experience through training to gain proficiency and automaticity in surgical tasks can moderate the effect of distractions. Training with distractions can even promote different resilience levels compared to training without distractions. A simulation study showed that residents who were trained in the presence of distractions performed better in a simulated virtual peg transfer task when auditory distractions (i.e., ambient noise, participant being summoned, and a pager call) were present in the environment compared to those who were trained in a quiet environment (Szafranski, Kahol, Ghaemmaghami, Smith, & Ferrara, 2009). Compared to residents that were trained in silence, residents that were trained with distractions had higher levels of gesture level proficiency (calculated by comparing nine decomposed gesture movements such as up, down, left, right and grasping to those of experts' performance), hand movement smoothness, tool movement smoothness, lower time to complete the peg transfer task, and lower cognitive errors (i.e., number of rings placed on a wrong peg) in the presence of distractions. Still, the presence of distractions decreased the surgical task performance for both groups.

Using realistic distractions such as visual distractions near the laparoscopic monitor, phone calls, irrelevant conversations, and loud noises, a laparoscopic cholecystectomy simulator study with 18 residents showed that major surgical errors significantly increased when residents were distracted (Feuerbacher, Funk, Spight, Diggs, & Hunter, 2012). These major surgical errors were computed by the simulator via assessing damage to internal organs, ducts, and arteries. In addition to surgical errors, distractions increased the number of participants who forgot to

complete a prospective memory task, which was to announce that they were three quarters done in the task for anesthesiologists. Pluyter, Buzink, Rutkowski and Jakimowicz (2010) conducted a simulator experiment of a laparoscopic cholecystectomy task (clip and cut assessed task) under distracting and non-distracting conditions with twelve medical interns who had no experience in surgical simulators and in laparoscopic procedures, to understand how distractions influence surgical performance. Distracting conditions involved social and technological distractions. For social distractions, music was playing in the background and a case-irrelevant conversation recording was overlaid on top of the music. Technological distraction was implemented by one experimenter operating the laparoscope non-optimally using a standardized protocol. Non-optimal handling of laparoscope was considered to put an extra cognitive strain on the surgeon since the surgeon would need to concentrate on the poorly controlled images; thus, it was considered to be distracting. The measures of surgical performance were task score, whether the task was completed, task error score, and total time required to complete the simulated task. Task score was found to be significantly lower and error score was significantly higher under distracting conditions compared to non-distracting conditions while task completion and total time required to complete the simulated task did not differ significantly. On the other hand, in a study of salpingectomy on a laparoscopic simulator with residents, distractions (frequent pagers and questions about the patient) were found to be non-significant for surrogate surgical task completion time and operative blood loss; however, participants were significantly more likely to successfully complete the simulator tasks under the no distraction condition (i.e., quiet condition; Murji et al., 2016).

While these studies did not vary the surgical task difficulty, Yang et al. (2017) investigated the effect of realistic OR distractions for an easy task (laparoscopic peg transfer) and a difficult task (laparoscopic precision cutting) under no distraction, and what the authors called mild and strong distractions. Mild distraction required the participant to take a phone call and answer a medically relevant question, while the strong distraction condition involved two separate phone calls with the second phone call requiring the participant to remember the information from the first phone call. Their measures for task performance were speed and accuracy of the surrogate surgical tasks (easy and difficult tasks), and quality of answers to the phone calls. The authors found that only strong distractions significantly decreased easy task

performance, while both distractions affected the performance of the difficult task. However, their participants were, again, novice surgeons that had little experience with laparoscopic tasks.

Looking at both novice and experienced surgeons, a controlled study found that noise, and noise with background music impaired auditory perception and speech processing for both novice and experienced surgeons when engaged in a laparoscopic peg transfer task on a simulator (Way et al., 2013). However, in this study, the primary concern was to assess the participants' speech processing performance which can be crucial for effective communication, but the surgical task performance was not analyzed. Since noise is inherent to the OR and music is preferred by some OR team members (Ullmann et al., 2008), understanding how noise and music can affect surgical task performance in addition to speech processing performance is important for OR distraction mitigation efforts.

Overall, the laboratory studies have generally showed negative effects of distractions on surgical performance, which may encourage eliminating distractions from the OR. However, it must be noted that these studies do not entirely reflect the true OR environment. The participants were mostly novice surgeons who had little experience in real OR environments. Moreover, operations in reality are not completed by one surgeon only, but they require extensive teamwork. Other OR team members will actively manage some of these distractions such as answering phone calls, or surgeons through experience may become resilient and automatized in surgical tasks to even benefit from some distractions. For instance, background music can be distracting for novice surgeons, but experienced surgeons may find it favorable.

Observational studies have more power to explain distractions in natural environments. However, as reviewed earlier, the existing observational studies have mostly looked into the interference levels of distractions instead of technical events. Only one observational study that examined the distraction-surgical error relationship was identified in the literature. A direct observational study of cardiovascular surgeries by Wiegmann et al. (2007) investigated the relationship between surgical errors and surgical flow disruptions, which were categorized as being related to teamwork issues, equipment and technology issues, resource-based issues, supervisory/training issues, and extraneous interruptions. Surgical error was defined as "an occasion in which a planned sequence of activities failed to achieve its intended outcome". The authors found that when surgical flow disruptions increased so did surgical errors; however,

teamwork failures were the only factor that was significantly linked to surgical errors. A limitation of this study was that only one observer collected data from the OR. Although the observer was medically trained and also received human factors training, the observer may have missed or misinterpreted events. Documenting both medical events and environmental events that occur frequently would be a highly demanding task for one observer. The physical presence of the observer could have also affected the OR team's behavior (known as the Hawthorne effect; McCambridge, Witton, & Elbourne, 2014). To overcome these issues of data collection, human observers can be replaced by audio and video recorders. Such naturalistic datasets are needed in examining distraction-technical event relationship. Naturalistic studies can provide unique insights from the OR by being a source of raw and relatively unbiased data that can be coded by multiple raters, and by minimizing the chances of missed and misinterpreted events by observers. Such studies can reveal when distractions are detrimental in surgical performance and further, when they can be beneficial.

1.1.6 Benefits of Distractions

As mentioned previously, distractions may interfere with those primary tasks that require similar kind of cognitive resources, and consequently, may degrade the primary task performance. In addition to consuming limited cognitive resources, distractions can create stress for OR team members (Arora et al., 2010) and in turn again have a negative effect on task performance. However, certain amounts of stress can be helpful. If a task requires focused attention, stress may improve task performance by filtering other information sources that are not required by the task (Wickens et al., 2000). According to the Yerkes-Dodson law, task performance can improve with increased arousal, with too much arousal decreasing performance, indicating an optimum level of arousal (Yerkes & Dodson, 1908). Moreover, a simple task has an optimum arousal level that is greater than that of a complex task illustrated as the peaks of the two inverted U's in Figure 2. Thus, the effects or effectiveness of distractions can change depending on task demands.

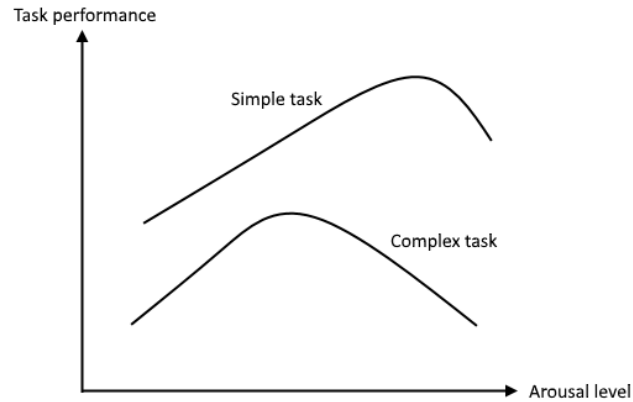


Figure 2. Illustration of Yerkes-Dodson law. Inverted U shapes show optimum levels of arousal for best task performance. This optimum level of arousal is higher for simple tasks compared to complex tasks (recreated from Yerkes & Dodson, 1908).

Simulation studies showed improvement in surgical task performance (time to complete task, and accuracy) with the presence of classical music for both residents and experienced surgeons (Kyrillos & Caissie, 2017) or high-rhythmic music for residents (Siu et al., 2010), with these sources potentially decreasing/increasing arousal. One study found no significant effects of background music for novice surgeons in simulated surgical tasks (Conrad et al., 2012). On the other hand, another controlled laboratory study showed that surgeon-selected music improved arithmetic task performance compared to no music for those surgeons who preferred to listen to music while operating (Allen & Blascovich, 1994). OR staff also expressed that background music made them feel calmer and that it improved communication (Ullmann et al., 2008). These studies show that at the least, music may be beneficial to OR performance. During high criticality steps when the individual is overloaded and over-aroused due to task demands, distractions such as calming music may bring the arousal levels down to optimum levels. High-tempo music on the other hand may increase arousal during routine task performance. Interviews suggest that during routine cases, surgeons would at times not pay enough attention to the environmental cues and would unintentionally “plough through” tasks (Moulton et al., 2010). Interviewed surgeons commented that these drifting moments could sometimes result in adverse events and near misses. Further, although sometimes the surgeons may plough through, distractions such as alarms may help surgeons receive necessary cues to slow down and be more alert. Distractions may also decrease boredom during low workload and routine tasks, bring arousal levels to more favorable levels, and hence, improve surgical performance. Therefore,

investigating distractions with considering surgical task demands can help us better understand when and how distractions can be harmful or beneficial. Potential benefits of OR distractions are slowly becoming recognized in the literature, but relevant studies are limited.

In an observational study of 56 general and orthopedic/trauma surgeries, an association between distractions and decreased mental fatigue and stress for surgeons was found (Weigl et al., 2015). A similar result was reported in a prospective cross-sectional observational study of 90 open surgeries that investigated the relationship between distractions, teamwork, and workload and stress levels of OR team members (Wheelock et al., 2015). This latter study reported that case-irrelevant conversations were observed when workload was low. Authors commented that conversations can potentially be one way of relieving stress. Although the authors of both studies suggested that communication may be a way of reducing workload, mental fatigue, and stress, the findings could also be explained the other way around due to correlational nature of their findings: OR teams may have engaged in more conversations because the workload was lower, and they had reserve cognitive capacity to do so. In order to understand the direction of this relationship, distractions can be further investigated by utilizing other workload metrics such as changes in pupil diameter or galvanic skin response. Time series datasets of distractions and such physiological measurements of the OR team members can be utilized to investigate if number of distractions increase during low workload moments, or if workload levels decrease after OR team engaged in certain distractions.

Aside from altering arousal and stress to more favorable levels, distractions may convey critical information for patient safety. A video aided observational study of laparoscopic anti reflux surgeries found communications to be the most frequent disruptive events (Zheng, Martinec, Cassera, & Swanström, 2008). Even though they did not investigate the effects of distractions on technical events, the authors noted that the conversations were mostly for exchanging information about the surgery; therefore, these “distractions” could have been mostly beneficial for patient safety rather than harmful. Similar results have been shown in other units in healthcare such as EDs and ICUs, with relevant studies reporting potential benefits of distractions such as improved situation awareness, communication of time-critical information (by other staff or by alarms), and decreasing boredom (Fisher, 1998; Rivera-Rodriguez & Karsh, 2010; Sasangohar et al., 2012). Although recent studies are acknowledging the positive effects in various healthcare domains, still little is known about the potential benefits of distractions in

healthcare in general, and in ORs in particular. Additionally, findings in different healthcare units may not be generalizable to OR environments. Overall, there seems to be some evidence to indicate that OR distractions can be beneficial. Investigating the effects of OR distractions during lower and higher workload surgical tasks may provide unique insights to the potential positive effects of distractions in addition to the negative ones. Therefore, analyses should look at not only the prevalence of distractions, but also the context that the distractions occur, and their effects on technical events and patient safety.

As stated earlier, despite the limited evidence, OR distractions have been mostly considered to be harmful (McCurdie et al., 2016). As a result, various mitigation strategies were suggested in the literature (see Feil, 2014 for a comprehensive review). Most mitigation strategies (e.g., “sterile cockpit rule”, checklists, crew resource management) are adopted from the aviation domain but their effects in ORs have not yet been investigated. Some mitigation strategies may help prevent the negative effects of distractions but may also inhibit their potential benefits. After implementing an interruption mitigation tool in a cardiovascular ICU, Sasangohar, Donmez, Easty and Trbovich (2015) observed a major decline in interruptions with personal content, a desired outcome, but also a decrease in work- and patient-related interruptions which might have a negative effect on patient safety. A systems approach to study different mitigation strategies can guide OR distraction mitigation efforts to foster the benefits of OR distractions while mitigating their harms.

1.2 Objectives

Previous section provided an overview of the literature related to OR distractions and their effects on surgical performance. Many studies looked at the prevalence of OR distractions and assessed their interference levels on the surgical flow using observational tools. Studies on distraction and technical event relationship were mostly controlled experiments; in fact, only one observational study was found that looked into this relationship. However, both controlled experiments and observational studies have limitations. More objective, more realistic, and less intrusive data collection methods as well as larger datasets are needed to further understand the distraction-technical event relationship. Documentation and analysis of OR events using video recordings can overcome the limitations of observational and controlled studies. Ethical concerns and non-compliance from personnel due to existing blame culture in healthcare impeded this

technology to be implemented in the ORs (Vincent et al., 2004) until very recently. Only one study in this fashion was identified that looked into technical performance and team performance using naturalistic data in the OR (Guerlain et al., 2005); however, no naturalistic study has been conducted yet to look into OR distractions.

Although the evidence on effects of distractions on surgical performance is not well established, distractions have mostly been considered harmful and mitigation efforts have been proposed. Effects of distractions can vary depending on cognitive demands of surgical tasks; distractions can even be valuable for patient safety. Mitigation efforts should adopt this perspective; in general, a systematic approach is needed.

This thesis aims to provide a start at addressing these gaps in the literature. First objective of this thesis is to explore OR distractions using a naturalistic dataset called the OR Black Box while also looking into differences in distractions based on the criticality and the difficulty of procedural steps. Second objective is to investigate the relationship between distractions identified and severe intraoperative technical events by performing statistical analysis on the OR Black Box dataset. With this objective, this thesis is the first study to look into the OR distraction-technical event relationship using a naturalistic dataset. By investigating the properties of harmful distractions, appropriate mitigation strategies can be selected and implemented accordingly among the many that have been suggested in the literature. The third objective of this thesis is to support the selection of appropriate mitigation strategies by systematically categorizing various strategies identified in the literature and through discussions with our partner hospital in a taxonomy. Lastly, the limitations of the OR Black Box dataset used in the analysis are discussed along with potential improvements and plans for future OR distraction studies investigating the distraction-technical event relationship as well as different mitigation tools.

Chapter 2

Quantification of OR Distractions and their Effects using Naturalistic Data

Literature has limited evidence on OR distractions leading to technical events. The limited evidence is due to drawbacks of data collection methods; controlled experiments are conducted under unrealistic conditions, while observational studies may have missed or misinterpreted events, and introduce observer bias. Naturalistic data collection is promising to help overcome these limitations. Thus, one of the aims of this thesis is to investigate the distractions observed in the OR and their effects on technical events, in particular, severe technical events, in a systemic approach, using a naturalistic dataset. No other research to date has done this. This chapter presents the analysis of the OR Black Box naturalistic dataset to investigate the link between distractions and severe technical events. OR Black Box collects audiovisual and physiological input from the ORs. This raw data is later labeled regarding technical data (technical events, errors, event severity ratings, procedural steps, and surgeons' technical skills) and distractions (doors opening and closing, people entering and exiting, machine alarms, external communications, staff changeovers, staff being late, surgeon switches, and loud music in the OR). Overall, 60 elective laparoscopic Roux-en-Y gastric bypass surgeries were used in the analysis. Access to patient information was limited to a little more than half of the size of the available dataset. Patient information included body mass index (BMI) levels and previous abdominal surgery information (a binary variable to indicate whether the patient had an abdominal surgery prior to the current case). These two variables were associated with operative difficulty.

Technical events refer to intraoperative technical events that can potentially result in injury and patient harm. Technical events were rated for their severity based on their clinical impact: severe technical events, if not rectified, can lead to adverse outcomes. Therefore, the analysis reported in this chapter focused on severe technical events. Distractions included information on when people entered/exited the OR including when the door opened and closed during entering/exiting (coded as people enter/exit), whenever a noise from a medical device was heard (coded as machine alarms), and whenever pager calls (received) or phone calls (made or received) were observed in the OR (coded as external communications). Distraction data included whenever OR members handed their shifts over (coded as changeovers). Since

changeovers would change the team composition and may have an impact on the flow of teamwork, they were considered as surgical team factors in the analysis. Only nurse changeovers could be used in the analysis as other types were rarely observed. Surgeons' technical skills were also considered as a surgical team factor as an indicator of team composition and workflow of the team. Procedural steps of operations were also considered in the analysis since steps of different criticality and difficulty would require different levels of attention and effort from the surgeons and can affect technical events. BMI levels and previous abdominal surgery history (a binary variable to indicate whether the patient had an abdominal surgery prior to the current case) were also included in the analysis as these patient factors can influence technical events.

Therefore, in order to investigate the relation between technical events and distractions, analyses were performed to predict severe technical events based on distractions (due to sample size limits, machine alarms, external communications, and people entering/exiting could be explored), while accounting for (1) different procedural steps of the surgery, (2) surgical team factors, including nurse changeovers (as other types were rare) and surgeon technical skills, and (3) patient factors, including BMI and previous abdominal surgery.

Exploration of the data revealed that almost all severe technical events occurred during what surgeons may consider high criticality/medium-high difficulty procedural steps (based on consultations with two surgeons who coded the data). Therefore, statistical models focused on the high criticality/medium-high difficulty procedural steps. Further, based on the distribution of the number of severe events within these procedural steps, the outcome variable was aggregated into a binary variable: ≤ 1 vs. ≥ 2 severe events. Given the binary nature of the outcome variable, logistic regression techniques were adopted.

It was found that out of the three distractions explored, only machine alarms were significant; a larger number of machine alarms was associated with an increased likelihood of severe technical events. As for the covariates, better surgeon technical skills were significantly associated with a decreased likelihood of severe technical events; no significant effect was found for nurse changeovers. Patient factors could not be included in this model as covariates, given that only about half of the cases had this information. However, to explore the relation between patient factors and severe technical events, another logistic regression model was built with just these two variables. A larger BMI level was found to be marginally significantly associated with

a higher likelihood of severe technical events, while previous abdominal surgery history did not have a significant effect.

The following sections provide detailed background information on the OR Black Box dataset and the data that could be used in our analysis, followed by data processing for analysis and the model construction. In addition to the logistic regression analysis reported above, descriptive statistics are also provided. Although the findings are limited based on the small sample size that could be used in this analysis, the analysis helped the research team learn the dataset, and come up with suggestions for an improved data collection framework for studying distractions. The limitations of the analysis reported in this chapter, as well as these suggestions are presented at the end of the chapter. Given that the OR Black Box dataset is growing continually, these lessons learned are valuable for future modeling efforts.

2.1 Background

OR Black Box (Surgical Safety Technologies Inc., Toronto, ON, Canada) is a first of its kind digital platform that collects OR activity via time-synchronized audiovisual recordings of the OR, laparoscopic videos of the operation, and both patient's and healthcare professionals' physiological measurements. OR Black Box was developed in Toronto's St. Michael's Hospital in 2013 with the overall goal of identifying factors that contribute to adverse events in the OR, mitigating these factors' harmful effects by targeted interventions and improve patient safety. After the initial year of installation serving as the pilot period, data collection techniques improved, and OR team members got accustomed to the technology, potentially reducing the Hawthorne effect (McCambridge et al., 2014). St. Michael's Hospital's Research Ethics Board approved the implementation of this technology and retrospective analyses of the collected data. And both OR team members and patients gave consent to be recorded by the OR Black Box.

OR Black Box uses multiple technological sources to capture intraoperative data, beginning with the patient being fully draped and ending with the drapes being removed. Audio from microphones placed in the OR, videos from wall-mounted panoramic cameras and from laparoscopes are recorded, time-synchronized, encrypted, and stored in a secure server at St. Michael's Hospital for 30 days. Within the 30 days, data is reviewed by two separate groups of subject matter experts. The first group is the distraction analysts who review the data for distractions. The analysts receive a two-month long training and participate in calibration

sessions periodically so that consistency is maintained across raters. Each distraction analyst annotates a surgery for distractions using a framework similar to the one created by Healey et al. (2006). Distractions labeled include doors opening and closing, people entering and exiting, machine alarms, external communications, staff changeovers, staff being late, surgeon switches, and loud music in the OR. After these distractions are coded, the case is reviewed by a member of the second group of subject matter experts, the clinical analyst. Clinical analysts are board-certified surgeons who received at least three months of training for administering standardized protocols for coding technical events; the clinical analysts who worked on the data utilized in this thesis had at least two years of experience in analyzing the collected intraoperative data. Clinical analysts review the raw surgical video/audio data and code for intraoperative factors such as procedural steps of the surgery, intraoperative technical events, error modes, event severity ratings, and error rectifications. Technical events are identified using the validated SEVerity of intraoperative Events and REctification framework (SEVERE; pre-publication) and error modes are identified using the validated Generalized Error Rating Tool (GERT; Bonrath, Zevin, Dedy, & Grantcharov, 2013). Later, clinical analysts rate surgeon's technical skills for each twenty-minute block of surgery using the Objective Structured Assessment of Technical Skills tool (OSATS; Martin et al., 1997), followed by non-technical skill assessment of surgeons and nurses using the Non-Technical Skills for Surgeons (NOTSS; Yule et al., 2008) and the Scrub Practitioners' List of Intraoperative Non-Technical Skills tools (SPLINTS; Mitchell et al., 2012), respectively. These annotations are completed within 30 days before the raw data is erased from the secure servers.

OR Black Box overcomes the data collection limitations that exist in today's OR distraction research. It provides a much more unbiased and complete source of data compared to other retrospective methods like questionnaires. Although one OR Black Box member is present in the OR to facilitate data collection from the moment the patient is fully draped until the drapes are removed, OR Black Box captures more natural behavior of the entire OR team compared to the unrealistic conditions and limited representativeness of the participants used in experimental studies. Moreover, the OR Black Box member in the OR does not collect intraoperative data unlike observational studies. An important benefit of the OR Black Box data is that the analysts can re-watch the data if they are unable to distinguish an event (technical or distraction) or if multiple events occurred in a short amount of time. This ability reduces the possibility of missed

or misinterpreted events by the observers when compared to direct observational studies. Further, the identification of events is made less effortful for the observers since annotations are automatically time-stamped.

2.1.1 Description of Dataset used in Analysis

OR Black Box dataset analyzed in this thesis consists of data extracted from audiovisual recordings of 267 elective laparoscopic surgeries collected between October 2017 and January 2019. These surgeries were performed at three Canadian hospitals by 10 to 15 operating surgeons; it was not always possible to distinguish the exact number of fellows from the video recordings. The dataset contained 33 types of laparoscopic surgeries. Three most common surgery types were laparoscopic Roux-en-Y gastric bypass (LRYGB; $n=64$), cholecystectomy ($n=38$), and unilateral inguinal hernia repair ($n=23$). LRYGB operations were chosen for analyses for homogeneity of data as they were the majority. Four LRYGB cases had to be removed from analysis due to missing distraction data ($n=1$) and mislabeled surgery type ($n=3$), which were identified to be mislabeled after discussions with clinical analysts. The remaining dataset contained 60 cases which took 83.2 (standard deviation, $SD=22.0$) minutes on average. The dataset included technical events and technical skill assessment for the operating surgeon coded by clinical analysts and distractions coded by distraction analysts. Because patient data were collected externally but not through OR Black Box, patient data were only available for surgeries conducted in St. Michael's Hospital where OR Black Box data is housed and not for the operations performed in the other two Canadian hospitals. So, patient data related analysis could be conducted only on this sub-sample. Figure 3 visualizes the case exclusion process.

Although 60 cases represent only a limited sample size and may not be enough to model the complexities of the OR environment, the analysis is conducted as a first step to explore this dataset. OR Black Box is installed in several hospitals and ORs across the world and the dataset is expanding rapidly. Further, the data annotation methods are being improved, and this analysis serves as identifying some of the limitations associated with distraction coding, which will inform future improvements in data collection.

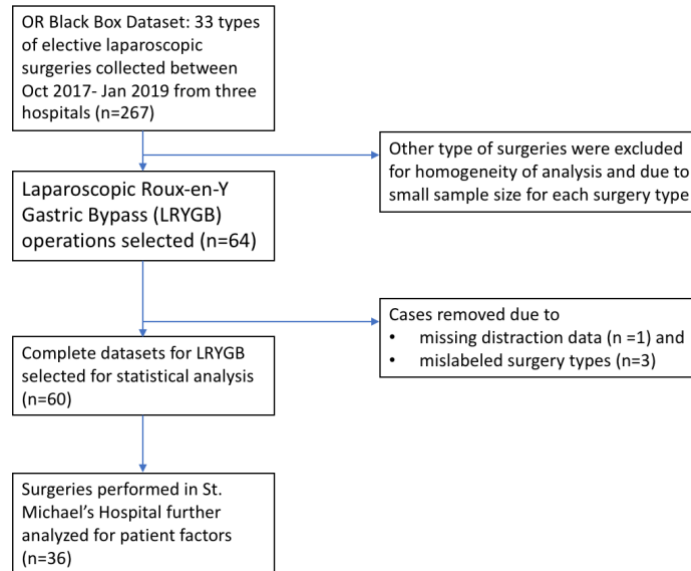


Figure 3. Data exclusion process represented by a flow chart

Technical Data

Technical events refer to intraoperative technical events defined as clinically relevant intraoperative adverse events that can potentially result in an injury and harm to the patient. Technical events were captured based on a validated framework (SEVERE, pre-publication) and included bleeding, hematoma, insufficient closure of anastomosis, ischemic injury, loss of pneumoperitoneum, mechanical injury, spillage, technical/tool failure, thermal injury, and others. If a technical event was due to the nature of the surgical task (e.g., a bleeding event during dissection), this event was associated with a “no error” label; otherwise (e.g., a bleeding event while grasping the bowel), the technical event was associated with an “error” label. Error modes were coded to indicate the reason behind a technical event that was associated with an error. Error modes included too little/much force, too little/much distance, wrong orientation, and inadequate visualization based on GERT as can be seen in Appendix A (Bonrath et al., 2013). However, this differentiation of error modes was not included in our analyses due to reliability issues regarding the coding of error modes.

Event severity refer to the clinical impact scale of an event and was determined based on a 5-point scale based on the SEVERE framework (pre-publication). Ratings 1, 2, and 3 indicate minor to moderate technical events that do not require the surgeon’s attention immediately, while ratings 4 and 5 indicate severe technical events that require immediate attention and rectification

from the surgeon for patient safety. If severe events (events that are rated 4 or 5) are not rectified, they could lead to adverse outcomes. In contrast, minor and moderate events (i.e., events rated 1,2, or 3) may not require to be rectified and are almost expected.

Procedural steps captured in the dataset for LRYGB operations are access/exposure, dissection/mobilization, reconstruction, inspection, no progress, secondary procedures, specimen resection, removal of specimen, and closure. The procedural steps in a LRYGB operation generally follow the order of access/exposure, dissection/mobilization, reconstruction (jejuno-jejunosomy), inspection, reconstruction (gastric pouch creation), inspection, reconstruction (gastrojejunosomy), inspection, and closure. Periods when the surgeon waits for instruments or stops to plan future actions are in general, labeled as no progress. And occasionally, secondary procedures such as hernia repairs or cholecystectomies are performed during LRYGB operations. Secondary procedures also contain specimen resection and removal of specimen.

Technical skills of operating surgeons are scored by clinical analysts using the OSATS tool (Martin et al., 1997). OSATS includes seven items, each rated on a 5-point scale as can be seen in Appendix B. The seven OSATS items are respect for tissue, time and motion, instrument handling, knowledge of instruments, use of assistants, flow of operations and forward planning, and knowledge of specific procedure. A score of 1 indicates poor technical skill, while 5 indicates good technical skill for a given item.

Distraction Data

The distractions annotated in the dataset used are doors opening and closing, people entering and exiting, machine alarms, external communications, staff changeovers, staff being late, surgeon switches, and loud music in the OR. Doors opening and closing, people entering and exiting, machine alarms, and external communications were coded with their start and end times, while staff changeovers, staff being late, surgeon switches, and loud music were coded as instances. Due to limited observations of staff being late (n=6) and loud music (n=7), these distractions were not considered in the analyses. As per distraction analysts' directions, surgeon switches were also excluded from analysis since surgeons, residents, and fellows were most of the times not determinable. Therefore, people entering/exiting, machine alarms, external communications, and changeovers were explored.

People entering/exiting was coded any time a person (either a scrubbed OR member or external staff) entered or exited the OR. However, the person who was entering or exiting was not identified. People entering/exiting also indicated when the door opened and when it closed while someone was entering/exiting the OR as the start and end times. So, only people entering/exiting variable was used in the current analysis.

Machine alarms were marked whenever a high-pitch sound from a medical device was heard.

External communications were coded when pager calls were received, or phone calls were received or made. Pagers, incoming phone calls, and outgoing phone calls were not distinguishable in the dataset.

Changeovers were coded when a team member ended his/her shift and handed the duty over. Changeover information was coded for nurses (n=66), surgeons (n=3), and for those who were not determinable (n=12). Surgeon changeovers were discarded in the analyses since they were rare. Not determinable changeovers, when it was not clear what OR role had a changeover, were also excluded from analyses. Although changeovers are labeled as distractions in the dataset, these also indicate team composition factors that can be controlled prior to the scheduling of a surgery, as changeovers would affect the information flow from one member to another, may affect teamwork and team performance. For these reasons, the effect of a changeover may be much wider during a surgery than an instantaneous distraction, therefore, changeovers were considered as a surgical team factor instead of a distraction in the analysis.

Patient Data

Patient risk factors (principal complaint, comorbidities, American Society of Anesthesiologists (ASA) classification status, BMI, age, and other relevant clinical information) have been identified to be one of the factors influencing surgical outcomes (Vincent et al., 2004). In our study, patients' BMI levels and previous abdominal surgery history (whether the patient had an abdominal surgery prior to the current operation) could be explored as patient factors that affected operative difficulty. Patient information was only available for those surgeries performed in St. Michael's Hospital, where the OR Black Box is housed. However, patient information for operations recorded in the other two hospitals were not available since patient

information was not collected from those two hospitals as part of the OR Black Box project. For surgeries performed in St. Michael's Hospital (n=36), BMI and previous abdominal surgery information of patients were acquired from patient admission records.

2.1.2 Model Construction

Our study aimed to investigate OR distractions and their effects on severe technical events. Due to sample size restrictions, distractions that could be explored were people entering/exiting, machine alarms, and external communications. The presence of severe technical events was chosen as the outcome variable. Severe technical events were not as common in the dataset as technical events that are of minor to moderate severity, and can be detrimental to patient safety if not rectified.

As indicated by Vincent et al.'s (2004) review article, as described in Chapter 1, there are several principle contributing factors to surgical outcomes: patient factors, surgical team (e.g., experience with working together), processes and procedures, key operative events (e.g., blood loss, minor and major complications, error compensation and recovery), flow of information following patient (e.g., handovers), technical skills, team performance and leadership, decision making and situation awareness, operative environment (e.g., availability and adequacy of equipment, availability of notes and records, noise and lighting, distractions), and interruptions (e.g., phone calls). Thus, when studying the effects of OR distractions on technical events, the effect of other factors should also be considered. The analysis reported in this chapter included three additional factors to distractions. First, nurse changeovers, which were originally coded as distractions, were considered as a surgical team factor, since changeovers would change the team composition and may affect the flow of the teamwork. Surgeons' technical skill was also considered as a surgical team factors since the coding of this variable includes interactions with other team members (e.g., use of assistants). Second, as some procedural steps may be critical or difficult and, therefore, may contribute to technical events, the criticality and difficulty of procedural steps was included as a procedural factor. Last, BMI and previous abdominal surgery information were considered as patient factors that can affect severe technical events. These three system factors (illustrated in Figure 4) were explored as covariates of the distraction-technical event relationship.

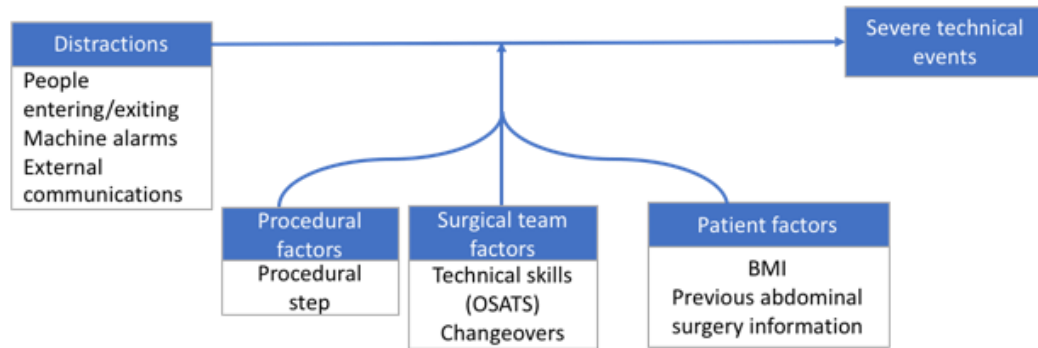


Figure 4. Illustration of three system factors (procedural, surgical team, patient) involved in the relationship between distractions and severe technical events

2.1.3 Data Processing for Analysis

Technical Events

Technical events in OR Black Box were collected for any intraoperative technical event related with injury or damage regardless of the nature of the task associated with the event. However, some technical events may be expected and common, and at times simply due to patient anatomy, such as bleeding during dissections. Given that we were interested in events that may be caused by distractions, we focused only on events where an accompanying error was also identified by the analyst, similar to the event definition put forward by the WHO (‘any deviation from usual medical care that causes an injury to the patient or poses a risk of harm’; World Health Organisation, 2005, p. 9). Further, as stated earlier, only severe technical events (ratings 4 and 5) were chosen for analysis as these events indicate a potential for significant harm to the patient.

Distractions

Machine alarms: Because machine alarms were labeled whenever a high-pitch sound from a medical device was heard by the analyst, there was an inflation in the coded machine alarms. For example, if a medical device generated a single alarm consisting of three high pitches as the alarm tone, three instances of machine alarms were coded in the OR Black Box. In order to remove this inflation in the number of alarms, two consecutive alarms that were at most 5 seconds apart were merged to be a single machine alarm.

Procedural Factors

Procedural steps: Discussions were conducted with two clinical analysts to gain a better clinical understanding of procedural steps. Procedural steps were rated as low, medium, or high based on two constructs. First was criticality, defined as the potential danger to a patient if the procedural step was done without focused attention. The second construct was difficulty, which was defined by the level of work necessary to accomplish the procedural step (Wickens et al., 2000). Two analysts separately rated each step for criticality and for difficulty. Later, the raters discussed their ratings. For those steps with different ratings, both analysts came to an agreement at the end of the discussions.

Results of discussions with the clinical analysts revealed that access/exposure and closure both had low criticality and low-medium difficulty (LLM) and mostly required physical strength. Dissection/mobilization was rated as low-medium for criticality and low for difficulty (LML). Reconstruction steps were rated as high for criticality and medium-high for difficulty. Inspection was rated as low for both criticality and difficulty; however, inspection steps required the surgeon to inspect the reconstruction steps and to assess the work completed. Therefore, these two procedural steps were joined and rated as high criticality, medium-high difficulty (HMH). The no progress and secondary procedural steps could not be assigned criticality/difficulty levels, since 1) no progress could be the surgeon either waiting for instruments, which did not require attention, or planning future actions, which required high-levels of attention, and 2) it was not possible to assess the criticality and difficulty of secondary procedures without knowing the actual secondary procedure. The overall agreement for procedure criticality between the two analysts was 92.3%, with free marginal kappa of 0.88 (95% Confidence Interval (CI): 0.66, 1.00; Randolph, 2008). The overall agreement for difficulty was 76.9%, with free marginal kappa of 0.65 (95% CI: 0.30, 1.00; Randolph, 2008). The interrater reliability scores were almost perfect for criticality and substantial for difficulty (Landis & Koch, 1977). In the end, the procedural steps were combined into five categories: LLM (access/exposure and closure), LML (dissection/mobilization), HMH (reconstruction and inspection), NP (No progress), and SP (Secondary Procedures). As will be presented later, almost all severe technical events that were associated with an error occurred during HMH. Thus, the analysis focused on this particular procedural step category.

Surgical Team Factors

Technical skills: After discussions with clinical analysts, three items of the OSATS tool (respect for tissue, knowledge of instruments, and instrument handling) were excluded from technical event analyses because the coding of these three items would depend on the occurrence of a technical event, and if used as a covariate to predict a technical event, would have resulted in a circular argument (i.e., lack of technical skill assumed to lead to events, but technical skill scored lower when an event is observed). Given that the remaining four OSATS items (time and motion, use of assistants, knowledge of specific procedure, and flow of operations and forward planning) were coded fairly independently from technical events, they were chosen to be included in our analysis.

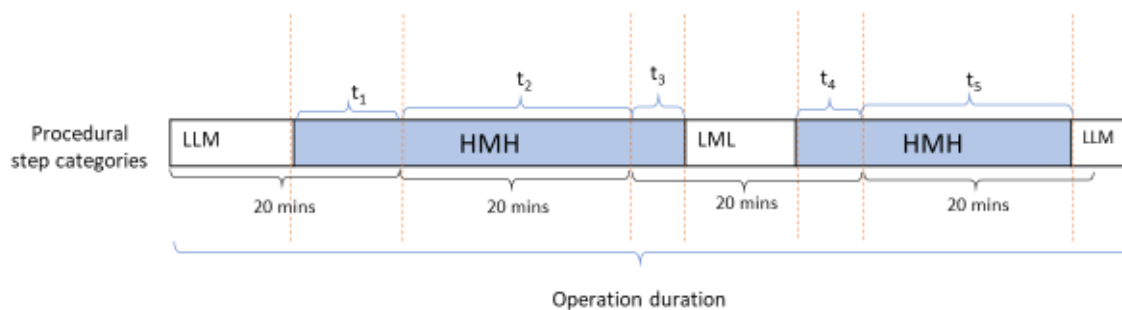


Figure 5. An example operation timeline showing the procedural step sequences (LLM, HMH, LML). OSATS are scored over 20-minute blocks. t_i 's indicate periods during which a 20-minute block and a HMH procedural step (shown in blue) overlaps.

Because clinical analysts rated OSATS in twenty-minute segments of surgery, there could be more than one OSATS score for a given HMH step (see Figure 5 for an example). Further, there were multiple HMH steps during an operation and the OSATS score for these different steps could be different and had to be aggregated (see Figure 5 with two HMH steps). In order to calculate one score for a given OSATS item during all HMH steps, a time-weighted average was calculated. The calculation is as follows:

(1) time periods when the twenty-minute blocks and HMH steps overlapped (shown as t_i 's in Figure 5) were found.

(2) The duration of these time periods (i.e., t_i 's) were used as weights.

(3) A weighted average was calculated for the four OSATS items across multiple HMH categories.

(4) The four average values obtained for four OSATS items were then summed to obtain the final OSATS variable.

Depending on the course of the surgery, t_i 's varied in magnitude. OSATS scores during t_i 's that were less than 5 minutes were excluded from the calculations as these scores may have largely represented the remainder of the coding period (i.e., >15 minutes of a different step). Possible values of the calculated OSATS variable ranged from 5 to 20 for a given operation.

2.2 Findings

All statistical analyses were conducted on RStudio version 1.1 (RStudio Team, 2016). First, exploratory analyses were conducted on the dataset to understand the distributions of severe technical events and distractions in general and across different procedural step categories. Second, logistic regression analyses were conducted in order to assess the effects of distractions on the occurrence of severe technical events. The main logistic regression model was built to examine the relationship between severe technical events and distractions (people entering/exiting, machine alarms, external communications) while controlling for procedural factors (procedural step categories) and surgical team factors (changeovers, OSATS average). Patient factors could not be included in this analysis, given the sample size limitations. To explore the relation between severe technical events and patient factors (BMI and previous abdominal surgery information), a simpler logistic regression model with just these two factors was built.

LRYGB operations ($n=60$) took 83.2 (SD = 21.97) minutes on average. During this time, on the average 1.53 severe technical events occurred per case (SD = 1.41). A potential distraction happened 47.6 times per hour (SD=20.3). This rate was more than the highest value reported in previous studies, which was 33.3 distractions per hour (Persoon et al., 2011). However, this difference was not surprising as our study did not differentiate between potential distractions,

distractions, and interruptions. Following sections present detailed descriptive statistics for the variables collected.

Severe Technical Events: Number of severe technical events ranged from none to six observations per case (for the distribution of severe technical events see Appendix C). Moreover, 45 cases had at least one severe technical event, while 15 cases had no severe technical events which created an imbalanced dataset, therefore, cases were grouped into two categories for statistical analysis: (1) Cases with at most one severe technical event (n=35), and (2) cases with at least two severe technical events (n=25).

People Entering/Exiting: Every hour, 17.8 people entered or exited the OR (SD=7.80), which also indicated the rate of door openings/closings. This finding was on the lower end of what previous observational studies had reported, which ranged from 13.4 door openings/h in plastic surgeries (Teter et al., 2017) to 40.8 door openings/h in general surgery cases (Healey et al., 2006); other studies reported 19.2/h in cardiac surgeries; Young & O'Regan, 2010; 22.8/h in a study of 24 arthroplasty, 7 arthroscopy, 6 podiatry, 5 spine, 4 general surgery, and 2 arthroscopic tendon repair; Perez et al., 2018; 40/h in pediatric orthopedic surgeries; Parikh, Grice, Schnell, & Salisbury, 2010. The variability across different studies may be due to the nature of operation performed. One observational study that collected data from a variety of surgery types (cardiac surgery, orthopedic surgeries, neurosurgery, plastic surgery, and general surgery) reported that orthopedic surgeries had 50 door openings per hour (reported as the highest rate observed), and general surgeries had 19 openings per hour (reported as the lowest rate observed; Lynch et al., 2009), similar to our finding.

During a surgery, the door stayed open for a total of 168.0 seconds on the average (SD=90.57) which corresponded to an average of 3.39% of the total case duration (SD=1.86). This result was considerably smaller from other studies. For example, Young and O'Regan, (2010) reported that the door was open approximately for 6 minutes per hour in their observational studies of cardiac operations. They also noted that the OR doors closed in 20 seconds. Lynch et al., (2009) commented that in their observational study of OR traffic, the door was open approximately 15 to 20 minutes per hour given that door took 20 seconds to close. Therefore, the differences in the durations of door staying open may be due to differences in OR door properties.

Further analysis on duration revealed two patterns in OR door openings while people entered/exited (Appendix D). The first type included those that took less than or equal to 10 seconds for the door to close from opening. 93% of all door openings took at most 10 seconds, and all cases (n=60) had these short door openings. When controlled for case duration, rate of short door openings was found to be 2.77 door openings per ten minutes (SD=1.20), totaled to 2.73% of the case duration (SD=1.28). The second type of door openings were those that took more than 10 seconds for the door to open and close. Thirty-eight out of 60 cases had at least one long door opening. For those cases that had at least one long door opening, duration of door openings ranged from 11 seconds to 4 minutes and 23 seconds (M=51.97 seconds, SD=49.43 seconds) and totaled to 0.67% of case duration on the average (SD=0.94). The rate of long door openings was found to be 0.19 per ten minutes (SD=0.21) in all 60 cases.

Short door openings were more frequent than long door openings, and the door was open for longer in total during surgery due to short door openings compared to long door openings. In order to understand their individual effects, and to mitigate their negative effects, reasons behind these two patterns of door openings need to be collected.

Machine Alarms: Machine alarms were the most frequent distractions observed with an average rate of 4.45 alarms per ten minutes (SD=2.88). Machine alarms could be heard for 65.0 seconds per case on average (SD= 50.58), which corresponded to 1.31% of the operation (SD=0.97). To the best of our knowledge, existing studies have not reported statistics on machine alarms.

Moreover, with machine alarms, three types of alarms were observed based on their duration patterns (Appendix E): individual alarms that generated a single beep and would take less than or equal to 2 seconds, short alarms that would take less than or equal to 5 seconds, and long alarms that take longer than 5 seconds. Table 1 shows the frequencies of these alarm types per case, their rates per ten minutes, and number of cases they were observed in. For those cases in which a specific alarm type was observed (i.e., excluding 0 sec duration), average durations of each type of alarm per case and percentages of case durations these alarm types were heard are reported in the last two columns of Table 1.

Table 1. Descriptive statistics are shown in the table for each machine alarm types. Numbers provided in the table are average values and in parenthesis are standard deviations.

Alarm Type	Frequency per case (SD)	Rate per ten min (SD)	No. of cases observed	For cases where particular alarm type was present	
				Observed durations per case (SD)	Percentage of case duration for observed alarms (SD)
Individual alarms ($\leq 2s$)	33.6 (20.80)	4.06 (2.62)	60	35.8s (20.83)	0.73% (0.44)
Short alarms ($>2s$ and $\leq 5s$)	1.58 (1.92)	0.20 (0.23)	42	8.93s (7.63)	0.13% (0.15)
Long alarms ($>5s$)	2.00 (2.03)	0.25 (0.26)	47	32.00s (44.35)	0.50% (0.74)

External Communications: Rate of external communications was 0.39 (SD=0.28) per ten minutes, indicating the lowest rate of distractions. In fact, not all cases had external communication: 56 cases out of 60 had at least one external communication during operation time. These observed external communications took 130.3 seconds on average (SD=101.50), which corresponded to 2.52% of the total case duration (SD=2.18). A similar result was reported in a video-aided observational study of 12 laparoscopic anti-reflux surgeries (Zheng et al., 2008). Authors reported that the OR team engaged in external communications four times per hour, which corresponded to 2.5 minutes of external communication aligning with our findings. However, they reported the incoming phone calls and pager calls. In our study pagers, incoming and outgoing calls were all combined in this variable, and pager calls, incoming and outgoing phone calls were not distinguishable.

Procedural Factors: As for procedural factors, all cases had LLM, LML, and HMH procedural step categories. However, only 13 cases had NPs, and 7 had SPs. Descriptive statistics with procedural step category breakdowns can be seen in Table 2. OSATS average, BMI, and previous surgery information were at the case level; therefore, they are constant across procedural step categories. One important finding of procedural step category breakdown was that out of 92 severe technical events observed in total, only one severe event occurred during LLM, while the remaining 91 events occurred during HMH. This finding empirically supported the validity of the categorization of these steps into high criticality and difficulty prior to the exploration of technical events within each category.

Table 2. Descriptive statistics for surgery and across procedural step categories. HMH: High criticality, medium-high difficulty category (Reconstruction, inspection). LLM: Low criticality, low-medium difficulty category (Access/ exposure, closure) LML: Low-medium criticality, low difficulty category (Dissection/ mobilization). NP: No progress. SP: Secondary procedures

	Entire case	HMH	LLM	LML	NP	SP
Duration in minutes	83.18 (21.97)	67.57 (15.73)	9.39 (4.50)	5.77 (7.24)	2.63 (1.97)	4.73 (4.35)
No. of cases	60	60	60	60	13	7
Average number of severe technical events (SD)	1.53 (1.41)	1.52 (1.41)	0.02 (0.13)	0	0	0
Average rate of people entering/exiting per ten minutes (SD)	2.97 (1.30)	2.88 (1.31)	3.44 (2.82)	3.33 (4.86)	3.85 (6.01)	3.17 (4.66)
Average rate of machine alarms per ten minutes (SD)	4.45 (2.88)	3.89 (2.93)	8.98 (3.91)	4.17 (5.24)	10.05 (11.98)	1.05 (2.02)
Percent duration of external communications (SD)	2.51% of entire case (2.18)	2.47% of HMH (2.47)	2.98% of LLM (6.24)	1.82% of LML (5.22)	0.17% of NP (0.63)	0.52% of SP (1.37)
Average rate of changeovers per hour (SD)	0.79 (0.93)	0.76 (0.92)	0.09 (0.50)	0.13 (0.51)	0	0
OSATS average (SD)	16.93 (1.56)					
BMI (n=36) (SD)	48.13 (8.23)					
Previous abdominal surgery (n=36)	No previous abdominal surgery; n=20					
	At least one previous abdominal surgery; n=16					

Surgical Team Factors: Out of 60 cases, only 33 cases had nurse changeovers. For these cases, there were 1.10 nurse changeovers on the average per case (SD=1.30). The overall rate for changeovers was 0.79 per hour (SD=0.93). The average value for OSATS average was 16.9 (SD=1.56).

Patient Factors: As mentioned, only 36 cases had patient information available. Those 36 cases had an average BMI of 48.1 kg/m² (SD =8.23). Sixteen patients had a previous

abdominal surgery, while remaining 20 patients did not have an abdominal surgery prior to the observed case.

2.2.1 Comparisons across Procedure Categories

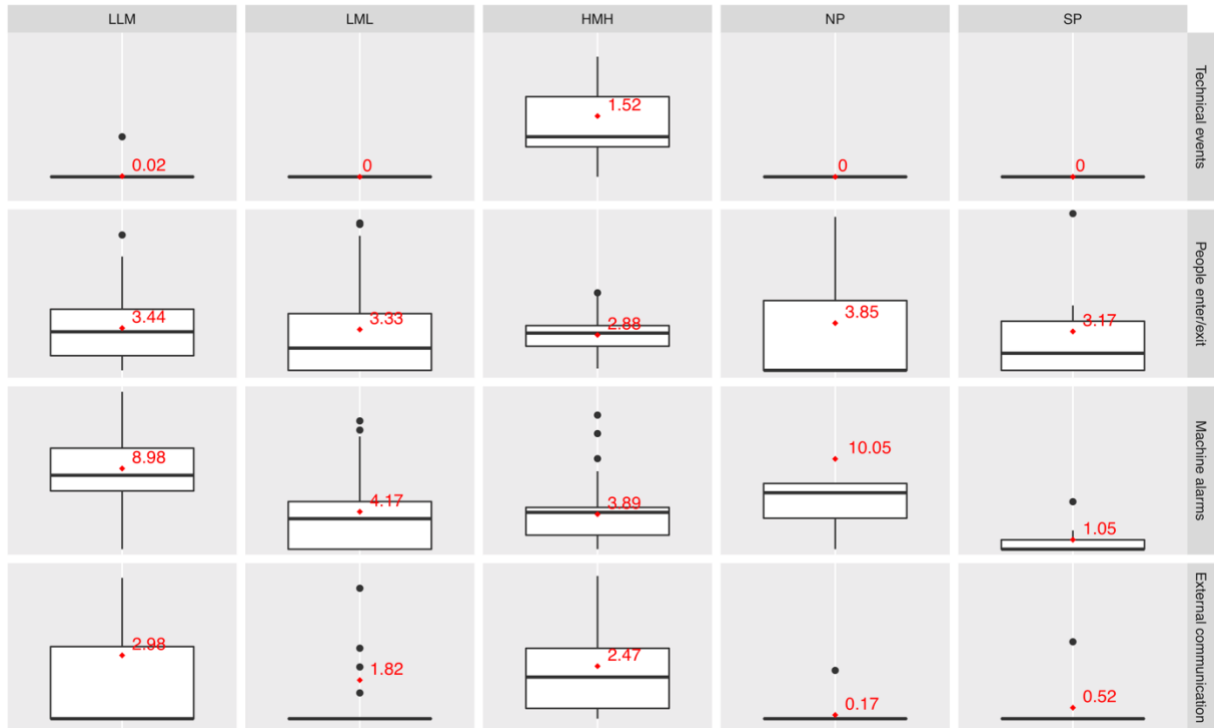


Figure 6. Boxplots showing the distributions and average values of severe technical events, distractions (rate of people entering/exiting per ten minutes, rate of machine alarms per ten minutes, and percent duration of external communications) across procedural step categories. HMH: High criticality, medium-high difficulty (reconstruction, inspection). LLM: Low criticality, low-medium difficulty (access/exposure, closure) LML: Low-medium criticality, low difficulty (dissection/mobilization). NP: No progress. SP: Secondary procedures.

Figure 6 shows the distributions of technical events and distractions across procedural step categories. As mentioned before, NP and SP categories were rarely observed in the dataset; however, each LRYGB operation involved LLM, LML, and HMH categories. Comparisons were conducted on distractions across these three common procedural step categories to understand the mean differences. However, LLM, LML, and HMH categories were not independent of each other since what happened during one category could impact what happened in the following

categories. Moreover, distributions of distractions failed the normality test and homogeneity of variances (Appendix F), violating ANOVA assumptions. Because of these reasons, Friedman's test was used to test the differences in prevalence of distractions between procedure categories. For those distractions with significant differences, post hoc tests were run. Differences between mean ranks for each pair comparisons are reported for post hoc tests. The critical level for the test statistic (i.e., critical difference) was found to be 26.22 ($\alpha=0.05$ and degrees of freedom =2).

Friedman test showed that the rate of people entering/exiting the OR did not significantly change across procedural categories, $\chi^2(2) = 4.41$, $p=0.11$. Rate of machine alarms, however, were found to be significantly different, $\chi^2(2) = 63.68$, $p<0.001$, with effect size $r=-0.25$. Post hoc tests with a Bonferroni correction showed that LLM steps had higher machine alarms compared to LML (rank difference=80) and HMH (rank difference =68.5). LML and HMH were not different for machine alarms (rank difference =11.5). For percentage of time spent on external communications during each procedural step category, significant differences were found ($\chi^2(2) =24.5$, $p<0.001$, $r=-0.25$). LLM and LML were not significantly different (rank difference =18); however, both LLM (rank difference =30) and LML were lower than HMH (rank difference =48).

Due to NP and SP being rare in the datasets, formal comparisons were not conducted on distractions experienced during these categories. During 13 NP cases, there were on average 3.85 (SD=6.01) observations of people entering and exiting per ten minutes and 10.05 (SD=11.97) machine alarms per ten minutes, and 0.17% (SD=0.63) of the time there were external communications. During 7 SP cases, there were on average 3.17 (SD=4.66) observations of people entering and exiting per ten minutes and 1.05 (SD=2.02) machine alarms per ten minutes, and 0.52% (SD=1.37) of the time there were external communications.

2.2.2 Logistic Regression Analyses

The main objective of the statistical analyses was to investigate the OR distraction-technical event relationship in a systematic manner by accounting for the identified system factors (procedural factors, surgical team factors, patient factors). As the descriptive analysis revealed, almost all severe technical events occurred during HMH procedural step category. For this reason, regression analyses focused on the OR distraction-technical event relationship during HMH category.

As noted earlier, severe technical events were chosen as the dependent variable. Since the analysis focused on the HMH procedural step category only, procedural step was not included in the regression analysis as a covariate. Surgical team factors were included as covariates. Patient factors could not be included as covariates due to the limited sample size but were explored in a separate logistic regression model. Descriptions of variables used in the models can be found in Table 3. Descriptive statistics of HMH duration, distractions, surgical team factors, and patient factors across the two dependent variable categories are presented in Table 4.

Table 3. Descriptions and measurements of variables used in statistical analyses

		Variable	Description	Measurement
Dependent variable	Severe technical events	Technical events	Severe technical events associated with errors observed during HMH category	≤1: At most one technical event was observed during HMH category
				≥2: At least two technical events were observed during HMH category
Predictor variable	Distractions	People entering/ exiting (per 10 mins)	People entering or exiting the OR during HMH category	Rate of people entering/ exiting observed in HMH category
		Machine alarms (per 10 mins)	Any machine alarms activated in the OR during HMH category	Rate of machine alarms activated in the OR during HMH category
		External communications (%time)	Any pager or phone calls received or made in the OR during HMH category	Percentage of duration of ext. communication during HMH steps
Covariates	Surgical team factors	Changeovers (per hour)	Nurse changeovers occurred in a case	Rate of nurse changeovers observed in the case
		Objective Structured Assessment of Technical Skill (OSATS)	Operating surgeons' technical skills scores during HMH category Four OSATS items were selected: Time and motion, Use of assistants, Knowledge of specific procedure, and Flow of operations and future planning	Average: Time-weighted average of OSATS item scores during HMH category, summed over four items selected
	Patient factors	BMI	Body mass index of the patient	Patient's weight divided by height squared
		Previous abdominal surgery	Abdominal surgery history of the patient	0: Patient had no abdominal surgery before current operation
1: Patient had abdominal surgery before current operation				

Table 4. Descriptive statistics of distractions, surgical team factors, and patient factors with respect to the severe technical event categories used in the dependent variable.

	Severe technical events			
	0 or 1 event (n=35)		2 or more events (n=25)	
	Mean	SD	Mean	SD
HMH duration (mins)	67.24	16.82	68.03	14.40
People entering/exiting	2.68	1.24	3.17	1.38
Machine alarms	3.12	2.40	4.97	3.29
External communications	2.24%	2.35%	2.80%	2.63%
Changeovers	0.75	0.88	0.86	1.01
OSATS	17.31	1.52	16.38	1.47
BMI (n=36)	45.74 (n=17)	7.70	50.27 (n=19)	8.30
Previous abdominal surgery (n=36)	= No previous surgery; n=11 = At least one previous surgery; n=6		= No previous surgery; n=9 = At least one previous surgery; n=10	

First, univariate analyses were conducted using each variable to predict the technical event groups. Results can be found in Appendix G. Univariate analyses showed that machine alarms (Odds ratio (OR)=1.29, 95% confidence interval (CI) = 1.05, 1.66) and OSATS average (OR=0.65, CI=0.43, 0.93) were significantly associated with observing more than one severe technical event during HMH.

Then, a multivariate logistic regression model was built with distractions as predictor variables and surgical team factors as covariates. Distractions and surgical team factors were checked for variation inflation factors (VIF), and no multicollinearity was found (see Appendix H). Results of this model showed that machine alarms and OSATS average were significant as shown in Table 5. While having increased number of machine alarms during HMH were significantly related to increased likelihood of having more than one severe technical events (OR=1.58, CI=1.18, 2.33), having increased technical skills denoted by OSATS average were found to be associated with decreased likelihood (OR=0.50, CI=0.30, 0.78). On the other hand, people entering/exiting (OR=1.04, CI=0.60, 1.78), external communications (OR=1.21, CI=0.92, 1.66), and changeovers (OR=1.10, CI=0.56, 2.16) were not significant. The goodness of fit value for the model was found to be 0.23 for Hosmer and Lemeshow's R^2 indicating a low significance for the model. When adjusted for sample size using Cox and Snell's correction method, R^2 was found as 0.27. After applying Nagelkerke's correction method for sample size, R^2 was found 0.36. All three R^2 values showed that logistic model did not fit the data too well. However,

regression model's chi square test was significant; it supported that the model was a significantly better fit than a baseline model without predictors and covariates ($\chi^2(5) = 18.7$, $p = 0.002$).

Table 5. Logistic regression results for predicting severe technical events through distractions and surgical team factors.

	Coefficient estimate (SE), p-value	OR	95%CI for OR	
			Lower	Upper
Intercept	8.76			
People entering/exiting	0.04 (0.27), $p = 0.90$	1.04	0.60	1.78
Machine alarms	0.46 (0.17) **, $p = 0.008$	1.58	1.18	2.33
External communications	0.19 (0.15), $p = 0.19$	1.21	0.92	1.66
Changeovers	0.10 (0.34), $p = 0.77$	1.10	0.56	2.16
OSATS average	-0.68 (0.24) **, $p = 0.005$	0.50	0.30	0.78

Note: $R^2 = 0.23$ (Hosmer-Lemeshow), 0.27 (Cox-Snell), 0.36 (Nagelkerke). Model $\chi^2(5) = 18.7$, $p = 0.002$. ** $p < 0.01$

Because patient factors had a limited sample size, adding them in this model would have reduced the statistical power of the test. In order to see if patient factors were associated with severe technical events, a logistic regression model was built using BMI and previous abdominal surgery information as covariates only. Table 6 shows the logistic regression results. BMI and previous abdominal surgery were checked for VIF values, and no multicollinearity was found (see Appendix I).

Table 6. Logistic regression results for predicting severe technical events with patient factors.

	Coefficient estimate (SE), p-value	Odds Ratio	95%CI for OR	
			Lower	Upper
Intercept	-4.55			
BMI	0.09 (0.05), $p = 0.07$	1.09	1.00	1.21
Previous abdominal surgery	1.02 (0.75), $p = 0.17$	2.78	0.67	13.36

Note: $R^2 = 0.10$ (Hosmer-Lemeshow), 0.13 (Cox-Snell), 0.17 (Nagelkerke). Model $\chi^2(2) = 4.85$, $p = 0.09$.

As can be seen from Table 6, BMI was marginally statistically significant in predicting the odds of having more than two severe technical events. High BMI levels (OR=1.09, CI=1.00, 1.21) increased the odds of having more than two severe technical events during HMH procedure

category; however, having a previous abdominal surgery (OR=2.78, CI=0.67, 13.36) did not. The model's R^2 using Hosmer and Lemeshow's method was 0.10. R^2 corrected for sample size by Cox and Snell's method was 0.13, and R^2 corrected for sample size by Nagelkerke's method was 0.17. All three R^2 values indicated a poor fit for the logistic model. Regression model's chi square test was marginally significant; the model was a marginally better at predicting severe technical events than chance ($\chi^2(2) = 4.85, p=0.09$). The marginal results could be due to sample size. However, these results indicated that patient factors could contribute to severe technical events and should be included in future studies of OR distraction-severe technical events relationship.

2.3 Discussion

This chapter investigated the effects of OR distractions on severe technical events using a first of its kind naturalistic dataset called OR Black Box. In order to understand the effects of OR distractions on patient safety, severe technical events were chosen, because severe technical events could have a high impact on patient safety and could lead to adverse surgical outcomes if not rectified. These events were observed in 75% of the laparoscopic Roux-en-Y operations recorded in the dataset. On average, 1.53 severe technical events per operation were observed.

OR distractions studied were selected as people entering/exiting, machine alarms, and external communications due to data limitations. Descriptive analysis showed that every hour, on the average, 18 people entered or exited the OR, 27 machine alarms went off, and two external communications took place; therefore, a distraction happened about 48 times per hour – almost one distraction every minute. These rates are slightly higher than the highest rate reported for distractions in the literature (33 distractions per hour; Persoon et al., 2011). This difference may be because our analysis focused on distractions regardless of the observable reactions from individuals that had to be identified in observational studies. These high rates show that distractions are frequent in the ORs, with some inherent to the ORs; machine alarms were observed almost every two minutes, but they cannot be removed from the ORs as they can convey critical patient information.

These frequent distractions could affect patient safety, both good and bad. They may be beneficial for the operation; for example, a nurse may leave the OR to replace malfunctioning equipment. Alternatively, these distractions may be detrimental; for example, by dividing an

individual's attention when the task at hand requires focused attention. However, ORs are complex dynamic settings with many factors contributing to severe technical events including (1) difficulty and criticality of procedural steps, (2) composition of the team and changes within the team, and (3) operative difficulty due to patient factors. In order to control for their effects, OR Black Box variables available for analysis were grouped under three system factors: (1) procedural factors (different criticality and difficulty levels of procedural steps), (2) surgical team factors (surgeon technical skills and changeovers) and (3) patient factors (BMI levels and previous abdominal surgery information).

Almost all severe technical events occurred during high criticality, medium-high difficulty procedural step category (HMH). This finding indicated that procedural factors do have an effect on severe technical events. During HMH, clinicians may have to allocate most of their cognitive resources to surgical tasks. However, distractions in the OR may compete for the resources required by the surgical tasks and may lead to surgical task degradation, leading to severe technical events. Therefore, these steps are particularly important to consider for distraction mitigation. It was found that HMH had significantly fewer external communications compared to both LLM and LML. The fewer external communications in the high focused attention phases may indicate that the OR team may be reducing their outgoing communications due to the demands of the surgery. It was however not possible to distinguish the outgoing communications from incoming communications, thus this hypothesis could not be tested with the existing data. Regardless, it may be important to filter out incoming communications that are not urgent during the high-focused phases of surgery.

An important finding was that machine alarms were significantly associated with severe technical events. These auditory distractions may put extra workload on surgeons' working memory by providing critical patient information during tasks that already exert high workload similar to the warning alarms during the Washington DC flight (Dismukes et al., 2007). It is also possible that machine alarms increase because there is a severe technical event; future analysis needs to investigate the temporal relation between alarms and technical events to test this hypothesis. Even though the number of machine alarms could increase due to the existence of a severe event, the rates of machine alarms were the lowest in the HMH phase that almost exclusively had severe events, compared to LLM and LML that had almost no severe events. However, machine alarms in the HMH phase may have a higher level of urgency even though

they may be rarer. For example, the higher rates of machine alarms observed during LLM, which consists of access/exposure and closure steps, could be due to sounds generated by equipment during preparation. The criticality of information conveyed by the alarms may help differentiate their effects on surgical performance, but such information was unspecified in the available dataset. Although no statistical comparisons were conducted, descriptive statistics of alarm types across procedural step categories are provided in the Appendix J. The urgency conveyed by the alarm through its saliency as well as duration should be investigated further. The majority of alarms were found to be instantaneous, taking less than 2 seconds. Onset and offset of salient alarms would grab attention more than continuous auditory alerts. Although alarms are essential for conveying critical patient information and thus cannot be removed from the OR, alarms can be designed in a way that their saliency adapts to surgical task demands in order to minimize their interference with task performance that requires focused attention.

People entering/exiting was not found to be significantly related with severe technical events. A limitation was that we could not distinguish the reasons for this activity using the OR Black Box data. Literature identified three most common reasons for OR traffic to be getting information, supplying equipment, and scrubbing in and out (Lynch et al., 2009; Panahi, Stroh, Casper, Parvizi, & Austin, 2012; Teter et al., 2017). Some of these reasons may be more distracting to the surgical team performance than others. For example, an external staff member entering to get information about another patient may be more distracting than a nurse bringing equipment into the OR. Besides being potentially distracting, frequent door activity and door being open for long durations can be harmful for patient safety by increasing the risks of surgical site infections through disrupting the sterile airflow of the OR (Andersson, Bergh, Karlsson, Eriksson, & Nilsson, 2012; Perez et al., 2018; Teter et al., 2017). Identifying reasons for door openings can explain the pattern of short and long door openings in future studies to find the best mitigation strategies such as limiting access to the OR for the external staff or improving OR equipment inventory management.

External communications were also not found to be associated with technical events. Differentiating pagers, incoming calls, and outgoing calls may bring additional insights on this finding. Pager calls may be more distracting and stressful for surgeons compared to an outgoing call made by a nurse since a pager call may mean an emergency for another patient. However, such inferences cannot be made from the current available data. One finding regarding external

communications was that they were significantly longer during the HMH compared to other two categories (LLM and LML). This finding may indicate that there were either more phone calls during HMH compared to pager calls, or phone calls took longer than those during LLM and LML.

In addition to distractions and procedural factors, effects of surgical team factors were also considered in the analyses. Higher surgeon technical skills were significantly related with observing less than two severe technical events during HMH. Surgeons with higher technical skills may have obtained automaticity in surgical tasks. Hence surgeons may need less cognitive resources allocated to critical and difficult surgical tasks and use cognitive reserves to accommodate distractions, lowering the chances of observing severe technical events. Although our analysis was not able to investigate the interaction effects between technical skills and distractions, training for technical skills and automaticity in surgical tasks may be a method of mitigating the negative effects of distractions in the OR as suggested and supported by other studies. Hsu et al. (2008) reported that experienced surgeons' peg-transfer task performance was not affected by arithmetic questions, suggesting that these surgeons had gained automaticity in the task. Mentis et al. (2016) also reported in their systematic review that good technical skills had mitigating effects for OR distractions. Although it must be noted that having high technical skills does not mean that surgeons will not be affected by distractions.

Changeovers, the second team factor, were not found to be significantly associated with severe technical events. Changeovers may lead to gaps in information flow (Cook, Render, & Woods, 2000), which is critical for surgical safety (Vincent et al., 2004). One important finding regarding changeovers was that they mostly occurred during HMH categories. This finding indicates an opportunity to protect HMH phases by scheduling changeovers to less demanding phases.

Problems due to patient factors have been found to be associated with high levels of stress among OR team members (Arora et al., 2010). However, due to limited sample size, patient factors could not be controlled for in our distraction analysis. In a simpler analysis, we found higher levels of BMI to be marginally significantly related with increased probability of observing more than one severe technical event. Future distraction studies should investigate patient information using larger datasets.

The key contribution of this chapter is that it reports the first naturalistic data analysis that investigated OR distraction-technical event relationship while considering other system factors. These factors were similar to the ones identified by Vincent et al. (2004) and were included in the regression analysis as far as the data allowed. Findings revealed that during HMH, higher numbers of machine alarms, lower levels of technical skills, and increased BMI levels may be related with increased chances of having more severe technical events. However, these findings are subject to many limitations. Prior to any mitigation efforts, limitations discussed in the following section need to be addressed with further studies.

2.3.1 Limitations

Distractions in this study represented sources that were potentially disruptive to attention. Data could be coded by labeling distractions as potential distractions (with no observed disruption to attention), distractions (observed disruption to attention), and interruptions (observed break in primary task). Such a framework could further inform which distractions are more disruptive to surgical flow and would enable us to more closely compare our findings to previous literature. A potential difficulty, however, for achieving this coding scheme is that attention shifts may not always be visible to the analyst from video recordings of wall-mounted cameras. Distractions that occur in the forward field-of-view of clinicians, such as the area behind laparoscopic monitors for surgeons, may attract visual attention but the shift in visual attention may not be detectable to an analyst. Eye-trackers or more cameras may be needed for this purpose. Further, the analyses only looked at associations between distractions and technical events at the procedural step level and causality cannot be inferred from these analyses. Future studies should analyze the time-series data to better understand the temporal relation between distractions and events.

There were also drawbacks due to the way distractions were coded. Because alarms were marked every time a high-pitch sound was heard instead of marking the start and end times of a device going off, consecutive alarms that were at most 5 seconds apart were artificially grouped to denote a single machine alarm. This method could have introduced a bias to our findings as it may not have captured the number of machine alarms perfectly. Therefore, coding for machine alarms should represent times a device goes on and off in addition to the tallying of the high pitch tones. Another coding related limitation was that pagers and phone calls were not

separately coded in the dataset. Pagers may indicate urgent calls for other patients and thus may take a surgeon's attention away from the current tasks, while outgoing calls made by the circulating nurse may be more easily ignored by the surgeon. Incoming or outgoing phone calls were also combined in the current data. An incoming call can be distracting because it will be received at unexpected and uncontrollable times. Therefore, analysis can benefit from differentiating pagers, incoming calls, and outgoing calls, and capturing conversation content. Similarly, it was also unknown if the communication was through OR phones or personal phones. These factors may provide important aspects when analyzing the effects of external communication on surgical performance and designing mitigation strategies. A final issue with distraction coding was the lack of information due to the difficulty of identifying certain information from video data. For example, surgeon switch information could not be used because at times the primary operator could not be reliably distinguished by coders. Similarly, it was not possible to identify the type of person entering or exiting the OR. This limitation can be overcome by differentiation of roles in the OR through the use of, for example, color-coded scrubs.

Current study looked at potential OR distractions through the surgeons' perspective. Potential distractions that were found not to affect surgeons' performance may be detrimental for nurses or anesthesiologists. Given that surgeries depend on teamwork, future investigations must consider the entire team and the effects of distractions on each individual's performance as well as the teamwork. Further, a member of OR Black Box was present in the OR during data collection, which may have degraded the naturalistic aspect of data collection. However, the member was not collecting any data themselves and was present only to ensure that data collection equipment was working properly. Finally, although the sample for current study was limited and only one type of surgery was analyzed, OR Black Box is a rapidly growing surgical safety initiative that continues to collect surgical data from partner hospitals worldwide. Future studies will have more power and their findings will be more generalizable, as there will be larger sample sizes and a wider range of operations represented from a larger number of hospitals.

Overall, the current analysis laid a foundation for future OR distraction studies that use naturalistic datasets and provided insights to improve distraction data collection methods. Suggestions for improving the OR Black Box data coding are provided in the next section.

2.3.2 Suggestions to Improve the Distraction Coding Framework

Several limitations regarding distraction data collection were identified and discussed. Improvements in distraction coding can bring more insights to future OR distraction analyses. An important limitation of the current study was that contextual details were not always captured in the dataset. Besides the type of distraction, its interference level (i.e., potential, break in attention, break in task), context (e.g., the topic of the conversation identified as case-irrelevant), source (e.g., who initiates the case-irrelevant conversation), and target (e.g., who the case-irrelevant conversation is addressed to) are required for more informative analyses. For example, case-irrelevant conversations may have varying interference levels across different target groups; direct observations of 48 general surgery procedures found that surgeons received case-irrelevant conversations with lower interference levels compared to anesthesiologists and to nurses (Sevdalis, Healey, & Vincent, 2007). Knowing the interference levels, context, sources, and targets of distractions can better inform the development of mitigation strategies.

First, the current set of distractions does not cover the most interfering distractions identified in the literature, such as equipment problems and conversations within the OR (Mentis et al., 2016). Equipment problems can be due to an equipment not working or that stops working during the operation (i.e., malfunctioning equipment) or due to a missing equipment required for the operation. Conversations in the OR can be differentiated for being case-related, work-related, case-irrelevant communication. Case-related conversations include any medical conversation related to the current case at hand (e.g., conversation about the surgical task). Work-related conversations include any medical conversation that is not related to the current case at hand but related with other clinical tasks; thus, these types of conversations may also be important and urgent. Lastly, case-irrelevant conversations include any conversation that is not related to the patient at hand or other clinical work; thus, these types of conversations may be delayed until more opportune times. An observational study of 78 endourological surgeries showed that missing or malfunctioning equipment had the highest interference ratings (mean interference rating per event was 4.97) using the interference rating tool created by Healey et al. (2006; Persoon et al., 2011). In the same study, equipment problems were followed by procedure-related conversations (i.e., case-related conversations; mean interference rating was 4.80), medically irrelevant conversations (i.e., conversations with no medical relevance; interference level was 4.79), and patient-irrelevant conversations (i.e., work-related conversations; mean

interference rating was 4.20). The interference ratings showed that these four distractions were frequently interrupting a sterile member's primary task in order to engage in the distraction. However, in our dataset, such distraction types were not coded, and their effects could not be investigated. For future analyses, equipment problems need to be captured when a problem related to a surgical instrument occurs until it is resolved; and conversations need to be captured while differentiating case-related, work-related, and case-irrelevant conversations.

Another distraction that was not included in the current dataset was background music. Background music has been studied in many controlled experiments and showed positive (Allen & Blascovich, 1994; Kyrillos & Caissie, 2017; Siu et al., 2010), negative effects (Murji et al., 2016; Way et al., 2013), and no effects (Conrad et al., 2012). The existing dataset included loud music, which was marked when an OR team member asked for the volume of the music to be lowered or the music to be switched off, instead of background music. However, background music was not coded unless a team member made this request. In order to understand the effect of music in the OR, background music as a potential distraction needs to be annotated from the time the music starts until it stops. Additionally, free text entry should be provided to distraction analysts as an option to capture the distracting events that were not considered in the framework. Because the videos are erased after 30 days; it is important to capture every OR distraction as much as possible without violating OR team member's privacy.

An important contribution to the existing framework would be capturing the reasons behind the OR traffic for the existing people entering/exiting variable. Capturing if a door was opened/closed for missing or malfunctioning equipment or conversation can supplement the understanding of the effects of door openings for targeted mitigation strategies. Similarly, urgency of machine alarms can better inform whether an alarm is a distraction that can be mitigated or whether it is needed in the OR. Documenting OR traffic reasons and alarm urgency can serve as contextual information for these distraction types.

To sum up, coding content source, target, and the interference levels of distractions, in addition to capturing their type can be more informative in understanding the effects of OR distractions. Once a more complete OR distraction range is collected, studies can obtain more insights for the OR distraction-technical event relationship. With these findings, targeted

mitigation strategies can be implemented to facilitate the positive effects of distractions while mitigating the negative effects.

Chapter 3

A Taxonomy of Distraction Mitigation Strategies for Operating Rooms

3.1 Introduction

Literature review in Chapter 1 demonstrated that distractions are not necessarily harmful in the OR; in fact, the effects are still unclear due to limitations, including those posed by data collection methods. The analyses reported in Chapter 2 addressed some of these data collection limitations by using a naturalistic dataset, the OR Black Box, and investigated the relationship between OR distractions and severe technical events. Interestingly, almost all severe technical events, which can put patient safety at risk if not rectified, occurred during high criticality/medium-high difficulty procedural step categories.

Analyses explored system factors such as procedural factors, surgical team factors, and patient factors. Results revealed that distractions (machine alarms) and surgical team factors (technical skills) were significantly associated with observing severe technical events during high criticality/medium-high procedural step categories of Roux-en-Y gastric bypass operations: more machine alarms were found to be associated with higher likelihood of having more than one severe technical event, while higher levels of technical skills were found to be related to decreased likelihood. Additionally, patient factors were found to be associated with severe technical events during high criticality/medium-high procedural step categories: Higher BMI levels were found to be marginally significantly related to increased probability of having more than one severe technical event. Although patient factors cannot be controlled, the effects of OR distractions and surgical team factors can be altered for improved surgical safety. However, these findings must be interpreted with caution due to limitations discussed in Chapter 2.

Even though the evidence is limited, and studies are inconclusive, existing studies consider OR distractions as harmful. To mitigate negative effects of distractions, many strategies have been suggested in the literature; see Feil's (2014) detailed list of mitigation strategies for a comprehensive review. Some mitigation strategies, in addition to the negative effects of distractions, may simultaneously block the potential benefits of distractions in the OR such as their efficacy in decreasing stress (Wheelock et al., 2015) and mental fatigue (Weigl et al., 2015). For example, although our results showed machine alarms to be related with technical events,

machine alarms are vital for conveying critical patient information and cannot be removed from the OR. In order not to impede the potential benefits of OR distractions, suggested mitigation strategies must consider the system as a whole. As ORs are complex dynamic systems, a categorization of mitigation strategies with a systemic approach can inform and guide OR distraction management.

In order to classify mitigation strategies within a systems perspective, this chapter proposes a taxonomy for OR distraction mitigation strategies that were identified through a review of the literature and discussions with OR team members. The generated taxonomy has three dimensions: (1) technology type (hard and soft), (2) management control type (feedforward, concurrent and feedback control) and (3) target levels (individual, team, and hospital). Additionally, the potential advantages and disadvantages of each strategy is hypothesized with the support of available literature.

3.2 A Taxonomy of OR Distraction Mitigation Strategies

The systems approach for human error management presumes that humans within a system will make errors. As Reason (2000) explains, the systems approach recognizes these errors as consequences of the system's shortcomings and emphasizes that the system should utilize defense barriers. Defense barriers possess control mechanisms, either in physical forms or in the forms of human skills and procedures, in order to prevent an error before occurring or to prevent an error from propagating into a catastrophic outcome. Further, defense barriers target multiple units of an organization: the person, the team, the task, the workplace, and the institution. Understanding the properties of defense barriers through can help the selection of targeted error management strategies.

A taxonomy that identifies similar properties of distraction mitigation strategies in the ORs can be beneficial for selecting the most appropriate mitigation strategies. This chapter proposes a taxonomy with dimensions identified using systems approach. This taxonomy has three dimensions. The type of technology is the first dimension: (1) hard and (2) soft technology. Hard technology contains tangible tools (Jin, 2005), or physical defense barriers. For example, physical signs that remind hospital staff not to initiate distractions in the ORs would be hard technology. Soft technology refers to intangible tools (Jin, 2005) that consists of trainings, procedures, and regulations to improve distraction management in the OR. For example, a

training program that trains surgeons for surgical tasks under distraction is a soft technology (Mentis et al., 2016).

Second dimension is the type of administrative control mechanisms. Reason (2016) states that there are two extremes of administrative control, feedforward and feedback control, and in between lies concurrent control. These three levels of administrative control are used in this taxonomy: Feedforward control aims to prevent errors related with distractions by predicting the future states of the surgery. For example, briefing sessions before starting a surgery could help the team members create common goals and shared expectancies for the surgery. By doing so, critical steps of the surgery can be made explicit for the team and individuals can act accordingly during those phases. Concurrent control introduces defense barriers that stay active during surgery. Rules and regulations that concern the OR (e.g., no personal texting within the OR), for example, are active throughout the surgery and help mitigate distractions in real-time. Feedback control is retrospective in the sense that it identifies incidents from past operations to provide corrective strategies to mitigate undesirable effects. A simple example would be reporting tools which would describe the events that led to undesirable outcomes to help plan for corrections for future operations.

Finally, the taxonomy specifies three target levels: (1) individuals, (2) OR team, (3) hospital management. Surgeries are performed by individual system actors such as surgeons, nurses, and anesthesiologists. Certain mitigation strategies may target individuals; for example, resumption support tools target individuals to help them return to an interrupted task successfully. However, surgeries are successful only through effective teamwork; and some strategies target the OR team, such as formalized communication protocols in the OR. Hospital can also be targeted to reduce the effects of OR distractions. An example would be introducing and executing rules and regulations within the OR that applies to the hospital.

Hard Technology (Tangible tools)	Hospital	Safety Warnings (Signs on interruption awareness)	Interruption Blocks (Limited room access during critical phases)	External Displays (Monitors outside of OR to inform external staff)	Feedback Tools (Hospital feedback to individuals about distractions experienced)
	OR team	Checklists (Checklists for required equipment)	Warning/Information Systems (Light to indicate when anesthesiologists are working on a critical task)		
	Individual		Preference Tools (List of members' preferred equipment)	Resumption Support Tools (Possibility to rewind laparoscopic images)	Feedback Tools (Anonymous feedback to hospital about distractions experienced)
Soft Technology (Intangible tools)	Individual	Training under Distractions (Training on simulators under realistic OR distractions)			
	OR team	Team Building (Formal CRM-like teamwork trainings)	Briefing (Creating surgical goals)	Communication Protocols (Read-back)	Debriefing (Reflecting on the surgery performance)
	Hospital	Mindful Management of Operations (Actively exploring possibilities of failures due to distractions)	Standardization of Team (Predetermined number of members per team)	Rules and Regulations within OR (No texting)	Systematized Performance Reporting (Culture of giving systematized feedback)
		Feedforward		Concurrent	Feedback

Figure 7. Taxonomy of OR distraction mitigation strategies with dimensions of technology type (hard/soft), management control type (feedforward, concurrent, feedback), and target (individual, OR team, hospital). Different strategies are denoted in bold with examples presented in parenthesis.

The proposed taxonomy (Figure 7) uses these three dimensions to classify the mitigation strategies. Strategies that have been identified through reviewing the literature and through discussions with OR team members fill the taxonomy cells. However, noticeably, some of these strategies overlap with multiple levels of a single dimension. This overlap indicates that the strategy carries the characteristics of multiple categories. For example, surgeons filling out preference tools before surgery can be considered as feedforward control, while scrub nurses using the same tool during surgery would be considered as concurrent control. The following sections define and discuss each mitigation strategy.

3.2.1 Hard Technology, Feedforward Control

Preference tools. Lack of regard for surgeons' preferred tools can be a source of distraction during surgery, and create surgical delays (Gillespie, Chaboyer, & Fairweather, 2012). Scrub nurses are responsible for the preparation of surgical setup and they can perform this task better if they are provided with surgeon preference cards (Mitchell et al., 2011). In an

effort to more effectively implement the use of preference cards, electronic databases are being created to store individual surgeon's equipment preferences. Providing such a system to team members can help lower their workload during equipment preparation as well as during surgery. However, these tools should allow the individuals to update their general and case-specific preferences. Additional recommendations for better integration of such e-tools with electronic health record systems can be found in Dizon, Richesson, Lytle and Hawks (2016), a root cause analysis study conducted to understand disruptions experienced with the electronic preference card system management used in orthopedic surgeries at the Duke University Health System.

Checklists. Missing tools/equipment result in significant surgical delays (Zheng et al., 2008). Staff entering and exiting the OR to retrieve missing equipment or patient information is a source of distraction (Lynch et al., 2009) in addition to being a risk factor for surgical site infections (Panahi et al., 2012). Surgical checklists are designed to aid the OR team to prepare for surgery, and have been shown to reduce adverse patient outcomes (Haynes et al., 2009) and communication failures (Lingard et al., 2008). Although checklists have been shown to be effective in general, they require focused attention; however, interruptions experienced by pilots during checklist operations is a well-known cause of aviation accidents (Dismukes et al., 2007).

Safety warnings. Individuals who are part of a complex system may not actively think about latent risk factors, such as distractions, but it is the institution's duty to remind them about risky actions (Reason & Hobbs, 2003). Similar to factory safety signs, safety signs regarding distractions can be posted near ORs. Signs that call the area to be "distraction-free" can remind individuals to refrain from walking into an operating room without checking its current usage. Through such tools, a distraction awareness culture can be built within the hospital, declining ill-timed distractions in the OR.

3.2.2 Hard Technology, Concurrent Control

Resumption support tools. Resumption of interrupted tasks can be supported by environmental cues that can remind the interrupted operator to resume a task and to resume it correctly. For example, a resumption support system that plays the last few seconds of laparoscopic images or videos before a surgical instrument was last activated could help support the surgeon after an interruption.

Warning/information systems. Certain surgical tasks are more critical than others; and not every OR team member may be involved in all critical tasks (Wadhwa et al., 2010). For example, there might be periods of surgery where the anesthesiologist has to maintain high levels of focus while the rest of the team may not have to and may therefore inadvertently create distractions to the anesthesiologist (e.g., case-irrelevant conversations). Warnings or information displays can be implemented to notify the OR team when certain team members are performing critical tasks. For example, an LED signal can be activated within the OR automatically or through the request of a team member when critical tasks are being performed. Protocols can be set about when to use such a tool, and what not to do when it is in operation.

In an observational study of ninety general surgery cases, distractions initiated by external staff were found to be the most prevalent type of distraction, with an average of 7.52 distractions per hour initiated by external staff (Wheelock et al., 2015). **Interruption blocks.** To reduce distractions caused by external staff, non-essential personnel's access to the OR can be limited during surgery. For example, doors may be unlocked through keycards. **External displays.** Alternatively or in addition, screens placed outside the OR can inform other hospital personnel about the current OR operations (Lynch et al., 2009), with the assumption that staff who are warned about the fact that there is a surgery in the room would refrain from entering the OR. The information provided can even highlight critical phases of surgery so that those distractions that are important but non-urgent may be delayed until a critical surgical phase is over.

3.2.3 Hard Technology, Feedback Control

Feedback tools. High reliability organizations support and encourage the reporting of adverse events and near-miss (Reason, 2008). Therefore, individual OR team members and OR teams should be encouraged to report issues about distractions and be provided with a system that is non-punitive. It should be noted that individuals are not always good at recognizing root causes contributing to errors (Lawton et al., 2012), therefore, may not realize that one of the root causes of an adverse surgical event may be distraction. Therefore, guidance should be provided through training or design of reporting tools for the individuals to describe the conditions surrounding an event in their entirety.

Just like individuals and OR teams can provide feedback to the hospital through reporting systems, the hospital can provide feedback to the individuals and OR teams. For example, feedback generally on technical and nontechnical skills (e.g., teamwork), and in particular on distraction-related behaviors can help mitigate distractions. OR data collection systems (e.g., audio, video, patient vitals) have been proposed and implemented to collect objective data on both technical and non-technical performance (Langerman & Grantcharov, 2017). These datasets can be used for general feedback as well as distraction-specific feedback, as well as for general research and training purposes. However, such data collection systems may not be accepted by OR teams until a “no-blame culture” (Reason, 2008) is adopted by the hospitals.

3.2.4 Soft Technology, Feedforward Control

Training under distraction. In a simulator study of laparoscopic surgery, experienced surgeons were found to be affected less by cognitive distractions compared to novice surgeons (Hsu et al, 2008). With experience, certain tasks become automatic, creating more spare cognitive resources for the operator. It appears that training that can increase surgical experience levels can help mitigate the effects of distractions. However, training under distractions might be beneficial for both experienced and novice surgeons, and make them more resilient to distractions (Mentis et al., 2016). Although training under distractions have been suggested as a way of mitigating OR distractions, there is limited evidence for their effectiveness (e.g., Szafranski, Kahol, Ghaemmaghami, Smith, & Ferrara, 2009).

Team building. Team members who are not engaged in critical tasks can step in to deal with OR distractions that may interfere with the performance of critical tasks conducted by the rest of the team (Moulton et al., 2010). Effective teamwork would prescribe that team members would recognize such situations and take action. On the other end, poor teamwork can generate additional distractions (e.g., personal conversations started by a team member when focused attention is required in the OR). In their observational study of ninety general surgery cases, Wheelock et al. observed that the second most prevalent distraction type was case-irrelevant conversations initiated by a member of the OR team, with an average of 7.6 distractions of this type per hour. Teamwork skills can be improved through training programs such as training and feedback (e.g., Wakeman & Langham, 2018).

Briefings. Preoperative briefing sessions have been shown to reduce surgical flow disruptions (Henrickson, Wadhera, ElBardissi, Wiegmann, & Sundt, 2009) and surgical delays (Nundy et al., 2008). Through briefing, surgical goals, expectancies, and critical tasks can be made explicit to the team members, and a shared understanding can be formed. Team members can understand when to refrain from initiating distractions and when to handle distractions for other team members. Barriers to the implementation of briefings were observed to relate to time limitations and also misperceptions of briefings as being “a waste of time” (Henrickson et al., 2009). In a prospective intervention study, Lingard et al. (2008) suggested that such barriers can be overcome through the design of briefing sessions with the participation of OR team members.

Proactive management of distractions. Hospitals should adopt a proactive management strategy where they should try to anticipate distraction-affected adverse events and ways to prevent them. Review of successful surgeries in addition to adverse events can inform hospitals about best practices and issues that need to be addressed.

Standardization of team. OR teams that have previously worked together have been shown to outperform new OR teams (Wiegmann et al., 2010). Standardizing OR teams as much as possible can accelerate team building, improving how teams handle distractions and reducing distractions that may arise due to poor technical and communication skills.

3.2.5 Soft Technology, Concurrent Control

Communication protocols. Communication failures involving communicating too little/too much, unclear (e.g., ambiguous non-verbal cues), or wrong time, and failing to understand or hear a message may contribute to surgical errors (Van Beuzekom, Boer, Akerboom, & Hudson, 2010). Communication protocols can be implemented within the OR to minimize such communication failures, improving the OR team’s resilience to distractions but also minimizing distractions that may arise from communication failures. An example would be read-backs to verbally acknowledge that a message or a request was received from a team member. Another example is the adoption of the sterile-cockpit rule; refraining from unnecessary communications during critical phases of surgery (Hohenhaus & Powell, 2008). However, for the sterile-cockpit rule to be effective in the OR, critical phases of surgery should be made explicit to the OR team members (Wadhera et al., 2010) before case starts or on preference cards.

Rules and regulations within the OR. In surgical simulator studies, residents' and medical students' surgical performance was found to decline when they were presented with phone and pager distractions (Murji et al., 2016; Yang et al., 2017). In order to reduce the number of distractions caused by these devices, their use in the OR may be regulated by hospitals such as leaving them at a monitored station outside the OR put up a note after case. Through regulations, hospitals may choose to make any of the other mitigation techniques presented in our taxonomy mandatory.

3.2.6 Soft Technology, Feedback Control

Debriefings. With structured debriefing sessions at the end of or shortly after a surgery, individuals can provide formal and constructive feedback to their team members to help them learn desirable distraction-related behaviors. Compared to hospital-provided feedback, debriefing sessions have the advantage of providing feedback when the memory of surgical events is fresh. Like briefing sessions, debriefings may also face barriers for implementation.

Systematized performance reporting. For feedback tools in general to be effective, a culture of “no-blame” should be fostered within the hospital environment. By creating a system of feedback, where feedback is adopted widely across individuals, OR teams, and the hospital, a positive safety culture can be developed in general, as well as it relates to distractions.

3.3 Discussion

In this chapter, we proposed a taxonomy of OR mitigation strategies and populated it with strategies suggested by earlier OR distraction research and also provided new suggestions (e.g., standardization of teams, resumptions support tools) based on the distraction literature at large as well as our informal discussions with OR team members. The dimensions of the taxonomy were adopted from the systems approach in human error management. Areas of further research and potential barriers to effectiveness were also discussed.

Distractions are inherent in the OR environment and pose threats on surgical performance. They may also have a positive impact on surgical performance but also on overall work environment. Likewise, literature is increasingly recognizing the potential benefits of distractions in healthcare such as decreasing stress (Wheelock et al., 2015), boredom, and mental fatigue (Weigl et al., 2015), improving surgical task performance with background music

(Kyrillos & Caissie, 2017; Siu et al., 2010) and providing critical information about patient safety either related to the patient-at-hand or related to overall work, i.e., to other patients (Sasangohar et al., 2012). It is important to ensure to foster positive distractions that are patient- or work-related, while mitigating negative ones. Our taxonomy identifies the circumstances these suggested mitigation strategies can be potentially effective. However, more research is needed to clarify how surgical outcomes may be affected by OR distractions and by reducing or completely removing them.

Thorough examinations are needed before implementing different mitigation strategies since these strategies, both in forms of hard and soft technology, can become sources of distraction themselves if designed, implemented, or maintained poorly. Safety culture needs to be adopted and supported, which can be achieved through the inclusion of feedforward, concurrent, and feedback control mechanisms for managing OR distraction. Success of distraction mitigation efforts requires active participation from individuals, OR teams, and hospital management. This systematic approach in categorizing mitigation strategies is aimed to assist OR distraction management and to improve patient safety together with work conditions of OR team members. However, further research is needed to assess the effectiveness of different strategies in practice.

Chapter 4

Conclusions and Future Work

Similar to other high-risk industries such as aviation, ORs are prone to human error. In order to build effective defense barriers against human error in ORs, contributing factors need to be investigated. Literature identified one contributing factor as distractions in the OR (Kohn et al., 2000); however, no clear conclusion about the effects of OR distractions can be drawn due to limitations of existing studies. Data collection methods are one of the important limitations: (1) Observational studies have not been able to directly study this relationship due to difficulties experienced by the observers, such as recording all distractions and intraoperative technical events while keeping fully focused during long surgeries and (2) controlled experiments lack the real OR conditions, as such studies using simulators usually have novice surgeons as participants, exclude teamwork, and are limited in sample size. These shortcomings require improvements in data collection methods to expand our knowledge of OR distractions.

OR distraction studies generally consider distractions to be harmful even though there is not adequate evidence on the effects of distractions. With the aim of mitigating the supposed negative effects, strategies have been taken from other high-risk industries and applied to the ORs. However, implementing these strategies can also have undesirable outcomes, such as hindering the potential benefits of OR distractions. Some distractions can be beneficial for patient safety; for example, conversations may help surgeons bring their attention back to the surgical task (Moulton et al., 2010) and may decrease stress (Weigl et al., 2015). Thus, before implementing mitigation strategies, the effects of OR distractions and associated mitigation strategies must be thoroughly studied. A taxonomy of mitigation strategies can inform OR distraction management about effectiveness, advantages, and disadvantages when selecting an appropriate mitigation strategy. However, such a taxonomy is missing in the literature.

This thesis aimed to understand the effects of OR distractions on technical events and to classify relevant mitigations strategies. Major findings of this thesis are as follows:

- Chapter 2 presented the first naturalistic study to investigate OR distractions and their effects on intraoperative technical events using the OR Black Box platform.

The OR Black Box platform collects audiovisual data from ORs through multiple inputs; recordings from microphones in the OR, cameras on the walls, the laparoscope, and data from medical devices are collected, time-synchronized and encrypted to be stored on secure servers. The collected data is then ready for analysts to code for distractions and technical events using standardized protocols. Moreover, the OR team members, although they are aware and they consent to being recorded by the system, are not pressured by being watched continuously by an observer present in the room. Such a platform can capture more naturalistic behavior and more objective and complete data from the OR.

Explorations of OR Black Box dataset of LRYGB operations revealed that severe technical events occurred almost always during procedural steps categorized by clinical analysts to be of high criticality/medium-high difficulty. Statistical analysis on the dataset revealed that during high criticality/medium-high procedure categories (1) increased number of machine alarms in the OR was associated with increased chances of observing more than one severe technical event; (2) higher levels of technical skills were associated with decreased chances of observing more than one severe technical event. High rates of machine alarms may compete for the cognitive resources needed for successful completion of the surgical tasks and may potentially degrade surgical performance. High levels of technical skills may indicate a gained automaticity in the surgical tasks, which could reduce the amount of cognitive resources needed for completing the surgical tasks successfully and can increase the surgeons' resilience against distractions.

Small sample size did not allow controlling for patient factors when analyzing the effects of distractions; however, a simpler model found increased BMI levels to be marginally significantly associated with an increased likelihood of more severe technical events during high criticality/medium-high procedure categories. Therefore, it is important to include patient factors as a covariate in future studies.

- Chapter 3 presented a taxonomy that categorized OR distraction mitigation strategies using a systems approach to guide OR distraction management.

After reviewing the literature and discussing with OR team members at our partner hospital, mitigation strategies were identified and the taxonomy was populated under three

dimensions: technology type, administrative control, and target levels. Strategies did not necessarily fall into one level in any given dimension; some strategies displayed characteristics of two levels. For example, preference tools show feedforward control when preparing equipment for the upcoming operation, and concurrent control when informing nurses on the equipment the surgeon preferred during surgery.

Major contributions of this thesis are as follows:

- Analyses described in Chapter 2 is the first naturalistic data study to investigate the effects of distractions on technical events.
- Analyses laid out a groundwork for future studies; recommendations for data collection improvements were reported to the OR Black Box team.
- The proposed mitigation taxonomy is the first to categorize the mitigation strategies through a systems approach.

Limitations and Future Work

A significant limitation in the study of understanding the effects of OR distractions on technical events was the distraction coding framework. First, the data did not indicate whether the surgeon was visibly distracted or interrupted, a missed opportunity to better understand the effects of different sources on surgeon's attention. Second, the range of distractions captured were limited. For example, distractions related to equipment problems were not captured in the current dataset. Lastly, contextual information for distractions was largely missing; for example, activities collected as door opening/closing did not explain why the OR door opened. Another limitation experienced in the statistical analyses were the limited sample size. However, as OR Black Box is a growing project, future studies will be able to utilize larger and improved datasets. By collecting more contextual information and more distraction types, the effects of distractions on OR technical events will be further explored. Interaction effects between distractions and other system factors will be studied. In addition, future analysis will compare the prevalence of distractions before and after technical events and conduct time series analysis to investigate the temporal aspects of the effects of distractions on technical events.

There were also limitations in creating the taxonomy of mitigation strategies for OR distractions. One limitation was that the literature was not reviewed systematically. And so, there may be other mitigation strategies that were not considered in the current taxonomy. Similarly, the analysis of advantages and disadvantages of different strategies were limited to the extent of the literature reviewed. Future research on the taxonomy will expand on the existing literature on mitigation strategies through a systematic literature review. In order to understand the qualitative data acquired from the literature review, the grounded theory will be used as a formal method, which is a method of generating theories by collecting and coding data to identify categories emerging from data (Strauss, 1987). Grounded theory consists of iterations of collecting data, coding data into categories based on common concepts and themes, and comparing and merging these categories. OR distraction mitigation taxonomy will be extended by applying grounded theory methodology on the data collected from a systematic review.

In summary, intraoperative data is being collected through the OR Black Box platform from multiple hospitals every day. With increased sample size and improved distraction coding framework, future studies can reveal more insights from ORs. Expanding the OR distraction mitigation taxonomy through a systematic literature review and analysis based on grounded theory can aid in selecting the most appropriate mitigation strategies. And by understanding and mitigating OR distractions, surgical safety can be improved.

References

- Ahmed, A., Ahmad, M., Stewart, C. M., Francis, H. W., & Bhatti, N. I. (2015). Effect of distractions on operative performance and ability to multitask-A case for deliberate practice. *Laryngoscope*, *125*, 837–841.
- Allen, K., & Blascovich, J. (1994). Effects of Music on Cardiovascular Reactivity Among Surgeons. *JAMA: The Journal of the American Medical Association*, *272*, 882–884.
- Andersson, A. E., Bergh, I., Karlsson, J., Eriksson, B. I., & Nilsson, K. (2012). Traffic flow in the operating room: An explorative and descriptive study on air quality during orthopedic trauma implant surgery. *American Journal of Infection Control*, *40*, 750–755.
- Antoniadis, S., Passauer-Baierl, S., Baschnegger, H., & Weigl, M. (2014). Identification and interference of intraoperative distractions and interruptions in operating rooms. *Journal of Surgical Research*, *188*, 21–29.
- Arora, S., Hull, L., Sevdalis, N., Tierney, T., Nestel, D., Woloshynowych, M., ... Kneebone, R. (2010). Factors compromising safety in surgery: stressful events in the operating room. *American Journal of Surgery*, *199*, 60–65.
- Baker, G. R., Norton, P. G., Flintoft, V., Blais, R., Brown, A., Cox, J., ... Tamblyn, R. (2004). Canadian adverse events study: The incidence of adverse events among hospital patients in Canada. *Canadian Medical Association Journal*, *170*. doi:10.1503/cmaj.1040498
- Bonrath, E. M., Zevin, B., Dedy, N. J., & Grantcharov, T. P. (2013). Error rating tool to identify and analyse technical errors and events in laparoscopic surgery. *British Journal of Surgery*, *100*, 1080–1088.
- Coiera, E. (2012). The science of interruption. *BMJ Quality & Safety*, *21*, 357 LP – 360.
- Conrad, C., Konuk, Y., Werner, P. D., Cao, C. G., Warshaw, A. L., Rattner, D. W., ... Gee, D. W. (2012). A quality improvement study on avoidable stressors and countermeasures affecting surgical motor performance and learning. *Ann Surg*, *255*. doi:10.1371/journal.pone.0178059

- Cook, R. I., Render, M., & Woods, D. D. (2000). Gaps in the continuity of care and progress on patient safety. *Bmj*, *320*, 791–794.
- de Leval, M., Carthey, J., Wright, D., Farewell, V., & Reason, J. (2000). Human factors and cardiac surgery: a multicenter study. *Journal of Thoracic and Cardiovascular Surgery*, *119*, 661–672.
- Dismukes, R., Berman, B., & Loukopoulos, L. (2007). *The Limits of Expertise*. London: Routledge. doi:<https://doi.org/10.4324/9781315238654>
- Dizon, J. M., Richesson, R. L., Lytle, K. S., & Hawks, S. J. (2016). Factors disrupting the preference card management process: A root cause analysis. *AORN Journal*, *103*, 105.e1-105.e12.
- Etchells, E., Mittmann, N., Koo, M., Baker, M., Krahn, M., Shojania, K., ... Daneman, N. (2012). *The economics of patient safety in acute care: Technical report*. Retrieved from <http://www.patientsafetyinstitute.ca>
- Feil, M. (2014). Distractions in the operating room. *Pennsylvania Patient Safety Authority*, *11*, 45–52.
- Feuerbacher, R. L., Funk, K. H., Spight, D. H., Diggs, B. S., & Hunter, J. G. (2012). Realistic distractions and interruptions that impair simulated surgical performance by novice surgeons. *Archives of Surgery*, *147*, 1026–1030.
- Fisher, C. D. (1998). Effects of external and internal interruption on boredom at work: Two studies. *Journal of Organizational Behavior*, *522*, 503–522.
- Gillespie, B. M., Chaboyer, W., & Fairweather, N. (2012). Factors that influence the expected length of operation: Results of a prospective study, 3–12.
- Goodell, K. H., Cao, C. G. L., & Schwaitzberg, S. D. (2006). Effects of Cognitive Distraction on Performance of Laparoscopic Surgical Tasks, *16*, 94–98.
- Grundgeiger, T., Dekker, S., Sanderson, P., Brecknell, B., Liu, D., & Aitken, L. M. (2016). Obstacles to research on the effects of interruptions in healthcare. *BMJ Quality and Safety*,

25, 392–395.

Grundgeiger, T., & Sanderson, P. (2009). Interruptions in healthcare: Theoretical views. *International Journal of Medical Informatics*, 78, 293–307.

Guerlain, S., Adams, R. B., Turrentine, F. B., Shin, T., Guo, H., Collins, S. R., & Calland, J. F. (2005). Assessing team performance in the operating room: Development and use of a “Black-Box” recorder and other tools for the intraoperative environment. *Journal of the American College of Surgeons*, 200, 29–37.

Hales, B. M., & Pronovost, P. J. (2006). The checklist—a tool for error management and performance improvement. *Journal of Critical Care*, 21, 231–235.

Haynes, A. B., Weiser, T. G., Berry, W. R., Lipsitz, S. R., Breizat, A.-H. S., Dellinger, P., ... Gawande, A. A. (2009). A surgical safety checklist to reduce morbidity and mortality in a global population. *The New England Journal of Medicine*, 360, 491–499.

Healey, A. N., Primus, C. P., & Koutantji, M. (2007). Quantifying distraction and interruption in urological surgery. *Quality and Safety in Health Care*, 16, 135–139.

Healey, A. N., Sevdalis, N., & Vincent, C. A. (2006). Measuring intra-operative interference from distraction and interruption observed in the operating theatre. *Ergonomics*, 49, 589–604.

Henrickson, S. E., Wadhera, R. K., ElBardissi, A. W., Wiegmann, D. A., & Sundt, T. M. (2009). Development and pilot evaluation of a preoperative briefing protocol for cardiovascular surgery. *Journal of the American College of Surgeons*, 208, 1115–1123.

Hohenhaus, S. M., & Powell, S. M. (2008). Distractions and Interruptions: Development of a Healthcare Sterile Cockpit. *Newborn and Infant Nursing Reviews*, 8, 108–110.

Hopkinson, S. G., & Jennings, B. M. (2013). Interruptions during nurses’ work: A state-of-the-science review. *Research in Nursing and Health*, 36, 38–53.

Hsu, K. E., Man, F. Y., Gizicki, R. A., Feldman, L. S., & Fried, G. M. (2008). Experienced surgeons can do more than one thing at a time: Effect of distraction on performance of a

simple laparoscopic and cognitive task by experienced and novice surgeons. *Surgical Endoscopy and Other Interventional Techniques*, 22, 196–201.

Jin, Z. (2005). *Global Technological Change: From Hard Technology to Soft Technology*. Intellect. doi:10.5840/techne201115217

Jothiraj, H., Howland-Harris, J., Evley, R., & Moppett, I. K. (2013). Distractions and the anaesthetist: A qualitative study of context and direction of distraction. *British Journal of Anaesthesia*, 111, 477–482.

Kahneman, D. (1973). *Attention and Effort*. Prentice-Hall. doi:10.2307/1421603

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (2000). *To Err Is Human*. Institute of Medicine. doi:10.17226/9728

Kyrillos, R., & Caissie, M. (2017). Effect of music on surgical skill during simulated intraocular surgery. *Canadian Journal of Ophthalmology*, 52, 538–542.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data, 33, 159–174.

Langerman, A., & Grantcharov, T. P. (2017). Are We Ready for Our Close-up? *Annals of Surgery*, 266, 934–936.

Lawton, R., McEachan, R. R. C., Giles, S. J., Sirriyeh, R., Watt, I. S., & Wright, J. (2012). Development of an evidence-based framework of factors contributing to patient safety incidents in hospital settings: A systematic review. *BMJ Quality and Safety*, 21, 369–380.

Leape, L. (1994). Error in medicine. *JAMA: The Journal of the American Medical Association*, 272, 1851–1857.

Lingard, L., Regehr, G., Orser, B., Reznick, R., Baker, R., Doran, D., ... Whyte, S. (2008). Evaluation of a preoperative checklist and team briefing among surgeons, nurses, and anesthesiologists to reduce failures in communication. *Archives of Surgery*, 143, 12–17.

Lynch, R. J., Englesbe, M. J., Sturm, L., Bitar, A., Budhiraj, K., Kolla, S., ... Campbell, D. A.

- (2009). Measurement of foot traffic in the operating room: Implications for infection control. *American Journal of Medical Quality*, *24*, 45–52.
- Martin, J. A., Regehr, G., Reznick, R., Macrae, H., Murnaghan, J., Hutchison, C., & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery*, *84*, 273–278.
- McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, *67*, 267–277.
- McConaughy, E. (2008). Crew resource management in healthcare the evolution of teamwork training and medteams®. *Journal of Perinatal and Neonatal Nursing*, *22*, 96–104.
- McCurdie, T., Sanderson, P., & Aitken, L. M. (2016). Traditions of research into interruptions in healthcare: A conceptual review. *International Journal of Nursing Studies*, *66*, 23–36.
- Mentis, H. M., Chellali, A., Manser, K., Cao, C. G. L., & Schwaizberg, S. D. (2016). A systematic review of the effect of distraction on surgeon performance: directions for operating room policy and surgical training. *Surgical Endoscopy and Other Interventional Techniques*, *30*, 1713–1724.
- Mitchell, L., Flin, R., Yule, S., Mitchell, J., Coutts, K., & Youngson, G. (2011). Thinking ahead of the surgeon. An interview study to identify scrub nurses' non-technical skills. *International Journal of Nursing Studies*, *48*, 818–828.
- Mitchell, L., Flin, R., Yule, S., Mitchell, J., Coutts, K., & Youngson, G. (2012). Evaluation of the Scrub Practitioners' List of Intraoperative Non-Technical Skills (SPLINTS) system. *International Journal of Nursing Studies*, *49*, 201–211.
- Moulton, C., Regehr, G., Lingard, L., Meritt, C., & MacRae, H. (2010). Slowing down to stay out of trouble in the operating room: Remaining attentive in automaticity. *Academic Medicine*, *85*, 1571–1577.
- Murji, A., Luketic, L., Sobel, M. L., Kulasegaram, K. M., Leyland, N., & Posner, G. (2016). Evaluating the effect of distractions in the operating room on clinical decision-making and

patient safety. *Surgical Endoscopy*, 30, 4499–4504.

Nundy, S., Mukherjee, A., Sexton, J. B., Pronovost, P. J., Knight, A., Rowen, L. C., ... Makary, M. A. (2008). Impact of preoperative briefings on operating room delays: A preliminary report. *Archives of Surgery*, 143, 1068–1072.

Panahi, P., Stroh, M., Casper, D. S., Parvizi, J., & Austin, M. S. (2012). Operating room traffic is a major concern during total joint arthroplasty. *Clinical Orthopaedics and Related Research*, 2690–2694.

Parikh, S. N., Grice, S. S., Schnell, B. M., & Salisbury, S. R. (2010). Operating room traffic: Is there any role of monitoring it? *Journal of Pediatric Orthopaedics*, 30, 617–623.

Perez, P., Holloway, J., Ehrenfeld, L., Cohen, S., Cunningham, L., Miley, G. B., & Hollenbeck, B. L. (2018). Door openings in the operating room are associated with increased environmental contamination. *American Journal of Infection Control*, 46, 954–956.

Persoon, M. C., Broos, H. J. H. P., Witjes, J. A., Hendriks, A. J. M., & Scherpbier, A. J. J. M. (2011). The effect of distractions in the operating room during endourological procedures, 437–443.

Pluyter, J. R., Buzink, S. N., Rutkowski, A. F., & Jakimowicz, J. J. (2010). Do absorption and realistic distraction influence performance of component task surgical procedure? *Surgical Endoscopy and Other Interventional Techniques*, 24, 902–907.

Raban, M. Z., & Westbrook, J. I. (2014). Are interventions to reduce interruptions and errors during medication administration effective?: a systematic review, 414–421.

Randolph, J. J. (2008). No Title.

Reason, J. (1990). *Human error*. Cambridge University Press.

Reason, J. (2000). Human error: models and management. *BMJ*, 320, 4–6.

Reason, J. (2008). *Managing the Risks of Organizational Accidents*. London: Routledge.

doi:10.4324/9781315543543

- Reason, J., & Hobbs, A. (2003). *Managing Maintenance Error: A Practical Guide*. Ashgate.
- Rivera-Rodriguez, A. J., & Karsh, B. T. (2010). Interruptions and distractions in healthcare: Review and reappraisal. *Quality and Safety in Health Care, 19*, 304–312.
- RStudio Team. (2016). RStudio: Integrated Development Environment for R. Boston, MA: RStudio, Inc. Retrieved from <http://www.rstudio.com/>
- Sasangohar, F., Donmez, B., Easty, A. C., & Trbovich, P. L. (2015). Mitigating nonurgent interruptions during high-severity intensive care unit tasks using a task-severity awareness tool: A quasi-controlled observational study. *Journal of Critical Care, 30*, 1150.e1-1150.e6.
- Sasangohar, F., Donmez, B., Trbovich, P., & Easty, A. C. (2012). Not all interruptions are created equal: Positive interruptions in healthcare. *Proceedings of the Human Factors and Ergonomics Society, 824–828*.
- Sevdalis, N., Healey, A. N., & Vincent, C. A. (2007). Distracting communications in the operating theatre. *Journal of Evaluation in Clinical Practice, 13*, 390–394.
- Sevdalis, N., Undre, S., Mcdermott, J., Giddie, J., Diner, L., & Smith, G. (2014). Impact of intraoperative distractions on patient safety: A prospective descriptive study using validated instruments, 751–758.
- Siu, K. C., Suh, I. H., Mukherjee, M., Oleynikov, D., & Stergiou, N. (2010). The effect of music on robot-assisted laparoscopic surgical performance. *Surgical Innovation, 17*, 306–311.
- Strauss, A. L. (1987). *Qualitative Analysis for Social Scientists*. Cambridge University Press.
- Szafranski, C., Kahol, K., Ghaemmaghami, V., Smith, M., & Ferrara, J. J. (2009). Distractions and surgical proficiency: an educational perspective. *American Journal of Surgery, 198*, 804–810.
- Teter, J., Guajardo, I., Al-Rammah, T., Rosson, G., Perl, T. M., & Manahan, M. (2017). Assessment of operating room airflow using air particle counts and direct observation of door openings. *American Journal of Infection Control, 45*, 477–482.

- Tschan, F., Seelandt, J. C., Keller, S., Semmer, N. K., Kurmann, A., Candinas, D., & Beldi, G. (2015). Impact of case-relevant and case-irrelevant communication within the surgical team on surgical-site infection. *The British Journal of Surgery*, *102*, 1718–1725.
- Ullmann, Y., Fodor, L., Schwarzberg, I., Carmi, N., Ullmann, A., & Ramon, Y. (2008). The sounds of music in the operating room. *Injury*. doi:10.1016/j.injury.2006.06.021
- Van Beuzekom, M., Boer, F., Akerboom, S., & Hudson, P. (2010). Patient safety: Latent risk factors. *British Journal of Anaesthesia*, *105*, 52–59.
- Van, J., Bos, D., Rustagi, K., Gray, T., Halford, M., Ziemkiewicz, E., & Shreve, J. (2011). The \$17.1 Billion Problem: The Annual Cost Of Measurable Medical Errors. *Health Affairs*, *30*, 596–603.
- Vincent, C., Moorthy, K., Sarker, S. K., Chang, A., & Darzi, A. W. (2004). Systems approaches to surgical quality and safety: From concept to measurement. *Annals of Surgery*.
- Wadhera, R. K., Parker, S. H., Burkhart, H. M., Greason, K. L., Neal, J. R., Levenick, K. M., ... Sundt, T. M. (2010). Is the “sterile cockpit” concept applicable to cardiovascular surgery critical intervals or critical events? The impact of protocol-driven communication during cardiopulmonary bypass. *Journal of Thoracic and Cardiovascular Surgery*, *139*, 312–319.
- Wakeman, D., & Langham, M. R. (2018). Creating a safer operating room: Groups, team dynamics and crew resource management principles. *Seminars in Pediatric Surgery*, *27*, 107–113.
- Way, T. J., Long, A., Weihing, J., Ritchie, R., Jones, R., Bush, M., & Shinn, J. B. (2013). Effect of noise on auditory processing in the operating room. *Journal of the American College of Surgeons*, *216*, 933–938.
- Weigl, M., Antoniadis, S., Chiapponi, C., Bruns, C., & Sevdalis, N. (2015). The impact of intra-operative interruptions on surgeons’ perceived workload: an observational study in elective general and orthopedic surgery. *Surgical Endoscopy and Other Interventional Techniques*, *29*, 145–153.
- Weigl, M., Weber, J., Hallett, E., Pfandler, M., Schlenker, B., Becker, A., & Catchpole, K.

- (2018). Associations of intraoperative flow disruptions and operating room teamwork during robotic-assisted radical prostatectomy. *Urology*, *114*, 105–113.
- Wheelock, A., Suliman, A., Wharton, R., Babu, E. D., Hull, L., Vincent, C., ... Arora, S. (2015). The impact of operating room distractions on stress, workload, and teamwork. *Annals of Surgery*, *261*, 1079–1084.
- Wickens, C. D. (2008a). Applied attention theory. *Ergonomics*.
doi:10.1080/00140130802295564
- Wickens, C. D. (2008b). Multiple resources and mental workload. *Human Factors*, *50*, 449–55.
- Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2000). *Engineering Psychology and Human Performance. Engineering psychology and human performance* (Third Edit). Prentice-Hall.
- Wiegmann, D. A., Eggman, A. A., Elbardissi, A. W., Henrickson, S., Sundt, T. M., & Crichton, M. (2010). Improving cardiac surgical care : A work systems approach. *Applied Ergonomics*, *41*, 701–712.
- Wiegmann, D. A., Elbardissi, A. W., Dearani, J. A., Daly, R. C., & Sundt III, T. M. (2007). Disruptions in surgical flow and their relationship to surgical errors: An exploratory investigation. *Surgery*, *142*, 658–665.
- World Health Organisation. (2005). WHO Draft Guidelines for Adverse Event Reporting and Learning Systems. *Geneva, Switzerland: Author. Retrieved March, 16*, 80.
- Wright, M. I. (2016). Implementing No Interruption Zones in the Perioperative Environment.
- Yang, C., Heinze, J., Helmert, J., Weitz, J., Reissfelder, C., & Mees, S. T. (2017). Impaired laparoscopic performance of novice surgeons due to phone call distraction: a single-centre, prospective study. *Surgical Endoscopy*, *31*, 5312–5317.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482.

- Young, R. S., & O'Regan, D. J. (2010). Cardiac surgical theatre traffic: time for traffic calming measures? *Interactive CardioVascular and Thoracic Surgery*, *10*, 526–529.
- Yule, S., Flin, R., Maran, N., Rowley, D., Youngson, G., & Paterson-Brown, S. (2008). Surgeons' non-technical skills in the operating room: Reliability testing of the NOTSS behavior rating system. *World Journal of Surgery*, *32*, 548–556.
- Zheng, B., Martinec, D. V., Cassera, M. A., & Swanström, L. L. (2008). A quantitative study of disruption in the operating room during laparoscopic antireflux surgery. *Surgical Endoscopy and Other Interventional Techniques*. doi:10.1007/s00464-008-0017-7

Appendices

Appendix A Generalized Error Rating Tool (GERT; Bonrath et al., 2013)

Video code			Rater code		
Surgical task group	Error mode	Time of observation	Total number	Event (description/time)	Mechanism of event
Abdominal access	Too much force/distance				
	Too little force/distance				
	Wrong orientation				
	Inadequate visualization				
Use of retractors	Too much force/distance				
	Too little force/distance				
	Wrong orientation				
	Inadequate visualization				
Use of energy devices	Too much force/distance				
	Too little force/distance				
	Wrong orientation				
	Inadequate visualization				
Grasping and dissection	Too much force/distance				
	Too little force/distance				
	Wrong orientation				
	Inadequate visualization				
Cutting, transection and stapling	Too much force/distance				
	Too little force/distance				
	Wrong orientation				
	Inadequate visualization				
Clipping	Too much force/distance				
	Too little force/distance				
	Wrong orientation				
	Inadequate visualization				
Suturing	Too much force/distance				
	Too little force/distance				
	Wrong orientation				
	Inadequate visualization				
Use of suction	Too much force/distance				
	Inadequate visualization				
Other unclassified	Description/time:				

Appendix B Global 5-point Rating Scale for Objective Structured Assessment of Technical Skills (OSATS; Martin et al., 1997)

OSATS Item	Rating Scale				
	1	2	3	4	5
Respect for tissue	Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time and motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument handling	Repeatedly makes tentative or awkward moves with instruments		Competent use of instruments although occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness
Knowledge of instruments	Frequently asked for wrong instrument or used an inappropriate instrument		Knew the names of most instruments and used appropriate instrument for the task		Obviously familiar with the instruments required and their names
Use of assistants	Consistently placed assistants poorly or failed to use assistants		Good use of assistants most of the time		Strategically used assistants to the best advantage at all times
Flow of operation and forward planning	Frequently stopped operating or needed to discuss next move		Demonstrated ability for forward planning with steady progression of operative procedure		Obviously planned course of operation with effortless flow from one move to the next
Knowledge of specific procedure	Deficient knowledge. Needed specific instruction at most operative steps		Knew all important aspects of the operation		Demonstrated familiarity with all aspects of the operation

Appendix C Distribution of Severe Technical Events

Table C-1. Descriptive statistics on severe technical events

Number of severe technical events observed per case	Number of cases	Technical event group	Number of cases within the technical event group
0	15	Cases with at most one severe technical event	35
1	20		
2	14	Cases with at least two severe technical events	25
3	5		
4	3		
5	2		
6	1		

Predicting for no severe technical event (n=15) vs. at least one severe technical event (n=45) created an imbalanced data set for logistic regressions. This could create bias for the minority group (i.e., no severe technical event). To have a more balanced dataset, cases were divided into two groups as: (1) cases with at most one severe technical event (n=35) and (2) cases with at least two severe technical events (n=25).

Appendix D Door Activity: Duration Analysis

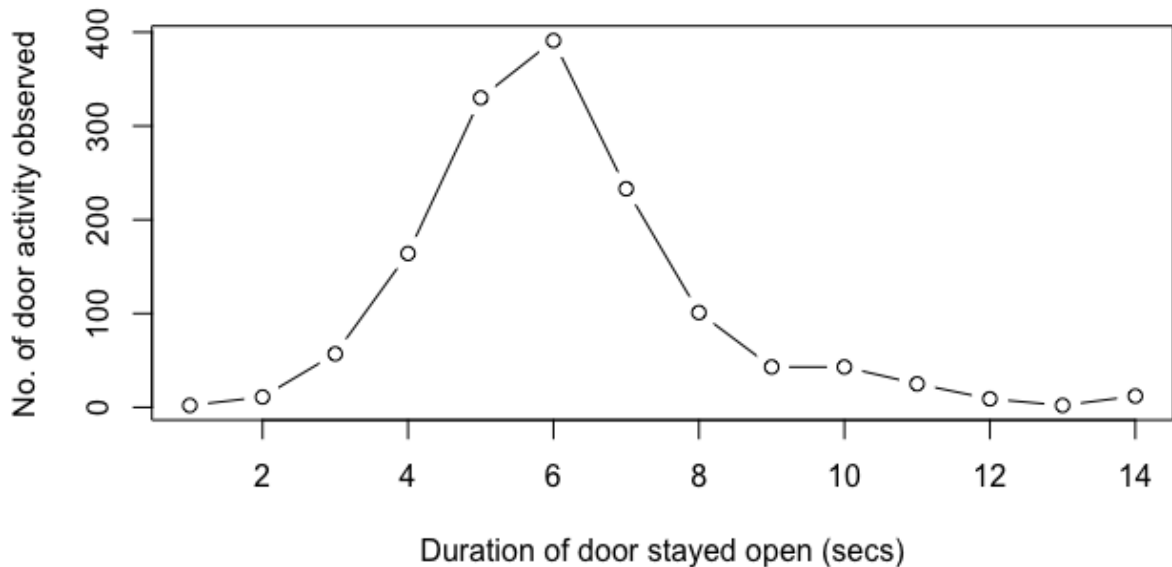


Figure D-1. Number of observed door activities (door opening/closing) with respect to their duration

Duration of door being open ranged from 1 seconds until 4 minutes and 23 seconds. Figure D-1 shows a part of the distribution of number of observed door activities (door opening/closing) with respect to the duration of door being open. Because durations longer than 10 seconds were less common than those less than or equal to 10 seconds, 10 second was chosen as the threshold for door opening categories to indicate short door openings ($\leq 10s$) and long door openings ($>10s$). This separation indicated that there were times when the door was kept open; however, the reasons for such a pattern in door opening/closing was unknown.

Appendix E Machine Alarms: Duration Analysis

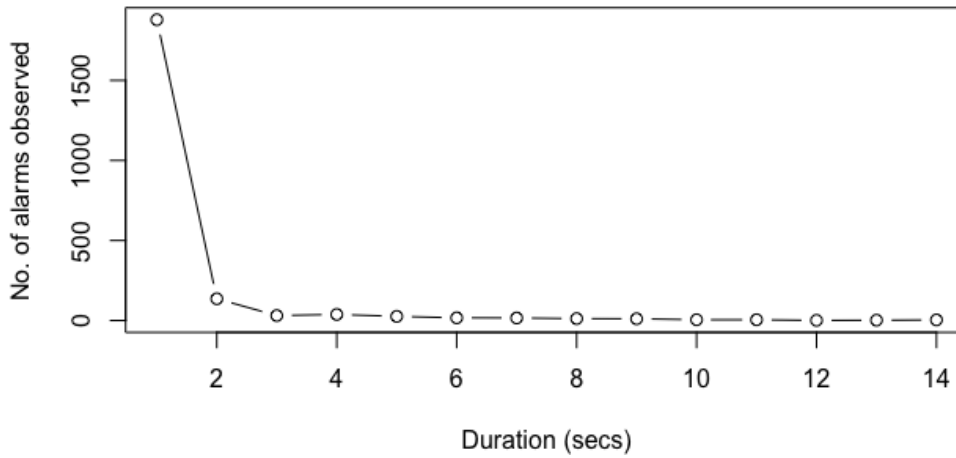


Figure E-1. Number of machine alarms observed that took between 1 second and 15 seconds

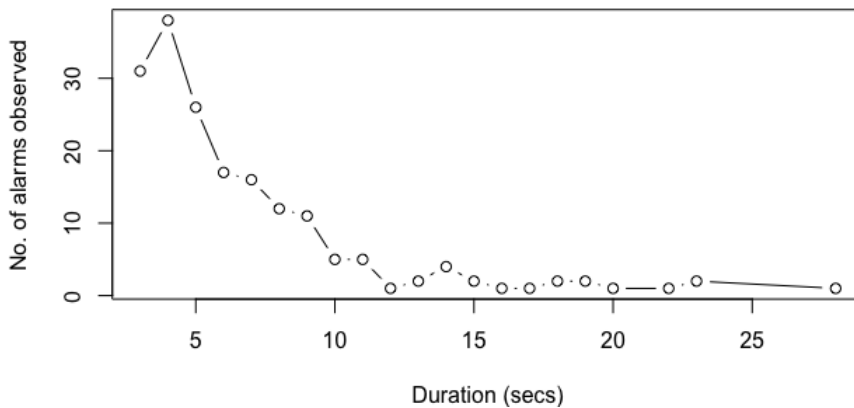


Figure E-2. Number of machine alarms observed that took between 3 seconds and 30 seconds

1375 out of 1472 (93%) machine alarms were observed to have durations of less than or equal to 2 seconds (Figure E-1). Any alarm larger than 5 seconds were observed less than 20 times in the entire dataset (Figure E-2). A stratification was observed for 2 seconds and 5 seconds as thresholds. Therefore, remaining machine alarms were separated after 5 seconds to indicate long machine alarms.

Appendix F Normal Q-Q Plots for Distractions

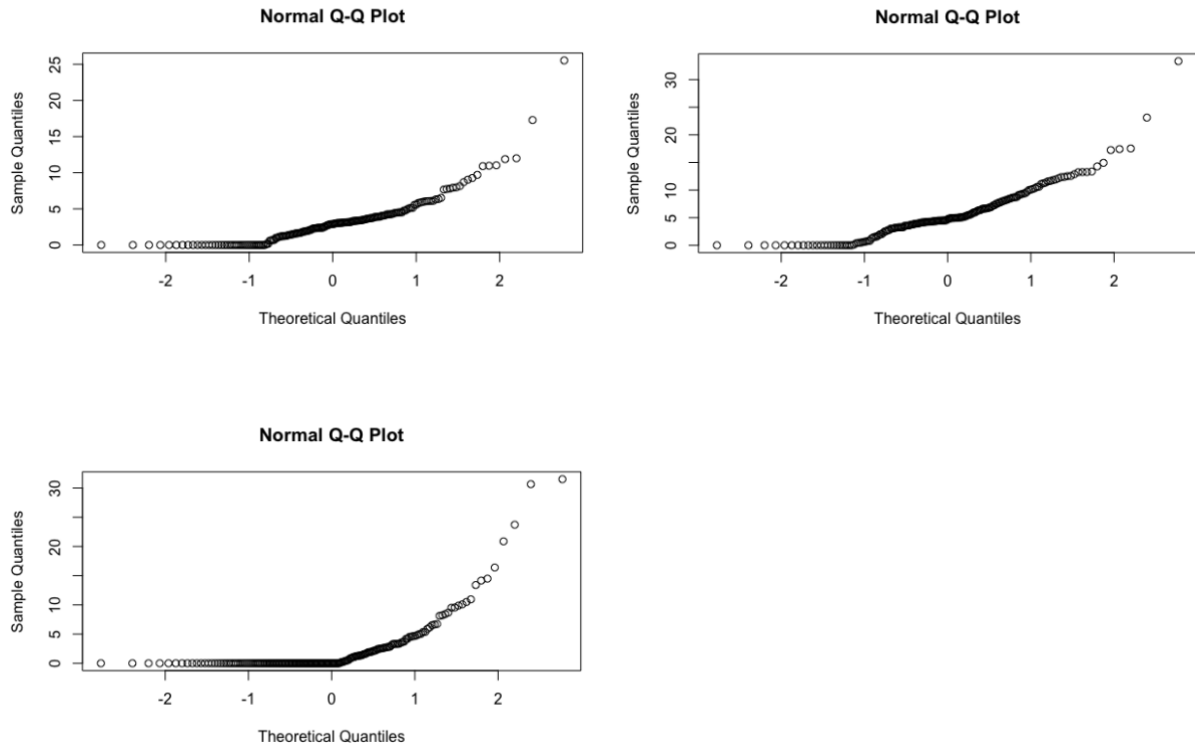


Figure F-1. Normal Q-Q plots for distractions (rate of people entering/exiting per ten minutes; top left; rate of machine alarms per ten minutes; top right; percent time of external communications; bottom left)

Distraction data failed in normality based on visual tests (Figure F-1) and Shapiro-Wilk tests. Shapiro-Wilk tests were significant for people entering/exiting ($W=0.79$, $p<0.001$), machine alarms ($W=0.88$, $p<0.001$) and external communications ($W=0.55$, $p<0.001$).

Homogeneity of variance was tested using Levene's test. People entering/exiting ($F(2, 19) = 11.2$, $p<0.001$) violated this assumption while machine alarm ($F(2, 177) = 2.38$, $p=0.10$) and external communications ($F(2, 18) = 1.21$, $p=0.30$) did not violate the homogeneity of variance test.

Appendix G Univariate Analysis: Results

Table G-1. Logistic regression results for predicting severe technical events through people entering/exiting

	Coefficient estimate (SE), p-value	Odds Ratio	95%CI for odds ratio	
			Lower	Upper
Intercept	-1.21			
People entering/exiting	0.30 (0.21), p=0.16	1.35	0.90	2.10

Note: $R^2 = 0.026$ (Hosmer-Lemeshow), 0.034 (Cox-Snell), 0.046 (Nagelkerke). Model ($\chi^2 (1) = 2.10$, $p=0.15$).

Table G-2. Logistic regression results for predicting severe technical events through machine alarms

	Coefficient estimate (SE), p-value	Odds Ratio	95%CI for odds ratio	
			Lower	Upper
Intercept	-1.33			
Machine alarms	0.25 (0.11) *, p=0.03	1.29	1.05	1.66

Note: $R^2 = 0.077$ (Hosmer-Lemeshow), 0.099 (Cox-Snell), 0.133 (Nagelkerke). Model ($\chi^2 (1) = 6.26$, $p < 0.05$). * $p < 0.05$

Table G-3. Logistic regression results for predicting severe technical events through external communications

	Coefficient estimate (SE), p-value	Odds Ratio	95%CI for odds ratio	
			Lower	Upper
Intercept	-0.57			
External communications	0.09 (0.10), p=0.39	1.10	0.89	1.37

Note: $R^2 = 0.09$ (Hosmer-Lemeshow), 0.013 (Cox-Snell), 0.017 (Nagelkerke). Model ($\chi^2 (1) = 0.76$, $p=0.38$).

Table G-4. Logistic regression results for predicting severe technical events through changeovers

	Coefficient estimate (SE), p-value	Odds Ratio	95% CI for odds ratio	
			Lower	Upper
Intercept	-0.44			
Changeovers	0.13 (0.28), p=0.64	1.14	0.65	2.01

Note: $R^2 = 0.003$ (Hosmer-Lemeshow), 0.004 (Cox-Snell), 0.005 (Negelkerke). Model (χ^2 (1) = 0.22, p=0.64).

Table G-5. Logistic regression results for predicting severe technical events through OSATS average

	Coefficient estimate (SE), p-value	Odds Ratio	95% CI for odds ratio	
			Lower	Upper
Intercept	6.85			
OSATS average	-0.43 (0.19) *, p=0.03	0.65	0.43	0.93

Note: $R^2 = 0.068$ (Hosmer-Lemeshow), 0.088 (Cox-Snell), 0.119 (Negelkerke). Model (χ^2 (1) = 5.55, p=0.018. * p<0.05

Table G-6. Logistic regression results for predicting severe technical events through BMI

	Coefficient estimate (SE), p-value	Odds Ratio	95% CI for odds ratio	
			Lower	Upper
Intercept	-3.41			
BMI	0.07 (0.04), p=0.10	1.08	0.99	1.18

Note: $R^2 = 0.058$ (Hosmer-Lemeshow), 0.077 (Cox-Snell), 0.103 (Negelkerke). Model (χ^2 (1) = 2.89, p=0.09).

Table G-7. Logistic regression results for predicting severe technical events through previous abdominal surgery information

	Coefficient estimate (SE), p-value	Odds Ratio	95% CI for odds ratio	
			Lower	Upper
Intercept	-0.20			
Previous abdominal surgery	0.71 (0.68), p=0.30	2.04	0.54	8.15

Note: $R^2 = 0.022$ (Hosmer-Lemeshow), 0.03 (Cox-Snell), 0.04 (Negelkerke). Model (χ^2 (1) = 1.1, p=0.29).

Appendix H Multicollinearity Results for Distractions and Surgical Team Factors

All variance inflation factors values were less than 10; no multicollinearity was found.

Table H-1. Variance inflation factor (VIF) of distractions (people entering/exiting, machine alarms, and external communications) and surgical team factors (changeovers and OSATS average)

	People entering/exiting	Machine alarms	External communications	Changeovers	OSATS average
VIF	1.28	1.45	1.36	1.09	1.44

Appendix I Multicollinearity Results for Patient Factors

All variance inflation factors values were less than 10; no multicollinearity was found.

Table I-1. Variance inflation factors (VIF) of patient factors (BMI levels and previous abdominal surgery)

	BMI	Previous abdominal surgery
VIF	1.09	1.09

Appendix J Descriptive statistics of machine alarms with respect to procedural step categories

Although mean comparison tests were not conducted, the results (Table J-1) indicated certain differences between different procedural step categories for machine alarms; LLM category appeared to have higher rates and higher percent time compared to HMH and LML categories. However, statistical tests are needed to conclude valid comparison results.

Table J-1 Descriptive statistics for machine alarms broken down with respect to their durations and the procedural step category.

				For cases where particular alarm type was present					
Machine alarm types	n			Average rate per ten minutes (SD)			Average percent time (SD)		
	HMH	LLM	LML	HMH	LLM	LML	HMH	LLM	LML
Individual alarms (<=2s)	55	59	40	3.68 (2.73)	7.40 (3.40)	3.63 (4.96)	0.69 (0.44)	1.60 (0.91)	0.94 (0.88)
Short alarms (>2s and <=5s)	21	32	4	0.10 (0.19)	0.84 (0.93)	0.17 (0.72)	0.19 (0.15)	1.02 (0.47)	1.88 (1.48)
Long alarms (>5s)	27	34	9	0.15 (0.23)	0.93 (1.12)	0.50 (1.99)	0.64 (0.77)	2.90 (3.04)	9.62 (17.20)

HMH: High criticality, medium-high difficulty (reconstruction, inspection). LLM: Low criticality, low-medium difficulty (access/exposure, closure) LML: Low-medium criticality, low difficulty (dissection/mobilization). NP: No progress. SP: Secondary procedures.

Copyright Acknowledgements