THE INFLUENCE OF LEADERSHIP AND EXTREME CONTEXTS ON PHYSICAL AND PSYCHOLOGICAL OUTCOMES

by

Amy Leigh Akers

A thesis submitted to the Graduate Program in Management - School of Business

in conformity with the requirements for the

Degree of Doctor of Philosophy

Queen’s University

Kingston, Ontario, Canada

September, 2015

Copyright © Amy Leigh Akers, 2015
Abstract

The need to consider context in organizational and leadership research has been highlighted by many scholars (e.g., Johns, 2006; Rousseau & Fried, 2001), and in today’s increasingly volatile world, the need to understand leadership in extreme environments has never been greater. This dissertation investigates the impacts of leadership in extreme contexts (i.e. where there exists “risks of severe physical, psychological or material consequences [...] to organizational members or their constituents” Hannah, Uhl-Bien, Avolio, & Cavarretta, 2009, p. 897), using two studies in the extreme context of healthcare.

Study 1 used growth-curve models to examine the changes in psychological outcomes (i.e. empathy, self-efficacy, daytime sleepiness) of medical and nursing students during their first eight-months of practical experience. Predictors of changes were also studied, including leadership behaviours (i.e. transformational leadership, laissez-faire leadership, abusive supervision), and context-relevant experiences (e.g., patient morbidity and mortality, beneficiary contact, boredom). Findings show that both leadership and context-relevant experiences influence the development of empathy, self-efficacy, and daytime sleepiness.

Study 2 investigated the impact of leadership (i.e. transformational leadership, laissez-faire leadership, over-controlling leadership) in the operating room on proximal and distal surgical outcomes (i.e. proximal: errors, blood loss; distal: complications during recovery at the hospital, complications post-discharge). Leadership effects were hypothesized to occur through the mediators of psychological safety and boredom, but no support was found for any mediating effects. However, certain direct relationships between leadership and surgical outcomes were moderated by surgical complexity. Leadership predicted proximal and distal surgical outcomes, but the strength of the relationship depends on the complexity of the surgery.
Results of both studies are presented, and the findings, avenues of future research, and practical implications are discussed. The dissertation concludes with a general discussion of the studies and areas of future research.
Acknowledgements

I would have never completed this dissertation without the guidance, love, and support from all of my family and friends. I cannot put into words how grateful I am to all of you, but I will try:

First, to my supervisor, Dr. Julian Barling, thank you so much for everything you did for me, from helping me identify my passion for organizational behaviour during my masters, to sitting with me for the final hours leading up to my defense. You truly went above and beyond and helped me accomplish feats I never thought possible. I am lucky to have been able to work with you!

Second, thanks to Dr. Darren Beiko, without whom none of this research would have been possible. Thank you for your enthusiasm for the research projects, and for making them happen!

Third, to my committee members, Dr. William Cooper and Dr. Christopher Miners. Your support and feedback during the preparation of my dissertation helped me strengthen my work, and make it something of which I am truly proud.

Next, to all of my friends. To my non-academic friends, thank you for being patient and supportive, lending your ears when I needed them, and giving me fun distractions at other times. To those who I shared the PhD experience with, you have no idea how critical you were to my success. Thank you for all of the knowledge you shared (both work-related and practical), and advice you gave me. To the Research Group, thank you for providing such amazing feedback for all of my research projects and dissertation work. A special thank you to Erica Carleton, who shares my love and appreciation for advanced statistical techniques, to Jennifer Robertson, Angela Dionisi and Alyson Byrne, who were my wise senior colleagues and ready at the drop of
a hat to help me out with anything, and to Wren Montgomery, who was right beside me at every milestone during the program.

Thank you to my family. To my mom, dad, and brother Anders, thank you for your endless love, encouragement, support, and laughter. To my extended family – the Akers, Rodericks, Littles, Smiths, and Jane – thank you for always being there, rooting me on. And to my grandmother, Dr. Ruth Gonnsen, who sadly passed away earlier this summer, thank you for being my inspiration and role model for the constant pursuit of knowledge.

Finally, to my husband Jason. You have been through the ups and downs of a PhD program, without getting the degree at the end. Thank you for always being there for me, thank you for your patience, and thank you for your unconditional love.
Statement of Originality

I hereby certify that all of the work described within this thesis is the original work of the author. Any published (or unpublished) ideas and/or techniques from the work of others are fully acknowledged in accordance with the standard referencing practices.

Amy Leigh Akers

September, 2015
# Table of Contents

Abstract ................................................................................................................................... ii  
Acknowledgements ................................................................................................................ iv  
Statement of Originality ........................................................................................................ vi  
List of Tables .......................................................................................................................... viii  
List of Figures ......................................................................................................................... x  
Chapter 1: General Introduction ............................................................................................ 1  
  1.1 References ....................................................................................................................... 16  
Chapter 2: A Longitudinal Study of Psychological Change in Extreme Contexts ................. 22  
  2.1 Introduction ....................................................................................................................... 23  
  2.2 Method ............................................................................................................................. 57  
  2.3 Data Analysis .................................................................................................................. 62  
  2.4 Results ............................................................................................................................. 65  
  2.5 Discussion ......................................................................................................................... 72  
  2.6 References ......................................................................................................................... 97  
Chapter 3: Scalpel! Leadership Behaviours and their Consequences .................................. 151  
  3.1 Introduction ....................................................................................................................... 152  
  3.2 Method ................................................................................................................................ 191  
  3.3 Data Analysis .................................................................................................................. 200  
  3.4 Results ............................................................................................................................. 204  
  3.5 Discussion ......................................................................................................................... 211  
  3.6 References ......................................................................................................................... 231  
Chapter 4: General Discussion ............................................................................................... 270  
  4.1 References ......................................................................................................................... 281  
Appendix A – Study 1: Full questionnaire ............................................................................... 285  
Appendix B – Study 2: Training Materials ............................................................................... 289  
Appendix C – Study 2: Observation Sheet ............................................................................... 311  
Appendix D – Study 2: Team member survey ......................................................................... 315  
Appendix E – Study 2: Postoperative outcomes to be collected ........................................... 316
List of Tables

Table 2-1: Scale reliabilities and descriptive statistics from pilot sample ........................................127
Table 2-2: Descriptive statistics and scale reliabilities .................................................................128
Table 2-3: Correlation table .........................................................................................................129
Table 2-4: Growth-curve models of empathy with leadership ......................................................129
Table 2-5: Growth-curve models of empathy with experiences ....................................................131
Table 2-6: Growth-curve models of self-efficacy with leadership .................................................132
Table 2-7: Growth-curve models of self-efficacy with experiences ..............................................133
Table 2-8: Growth-curve models of daytime sleepiness with leadership .......................................134
Table 2-9: Growth-curve models of daytime sleepiness with experiences .....................................135
Table 2-10: Sensitivity of covariance matrix on results of empathy and leadership model........136
Table 2-11: Sensitivity of covariance matrix on results of empathy and experiences model .........137
Table 2-12: Sensitivity of covariance matrix on results of self-efficacy and leadership model .138
Table 2-13: Sensitivity of covariance matrix on results of self-efficacy and experiences model .................................................................................................................................139
Table 2-14: Sensitivity of covariance matrix on results of daytime sleepiness and leadership model ........................................................................................................................................140
Table 2-15: Sensitivity of covariance matrix on results of daytime sleepiness and experiences model .............................................................................................................................141
Table 2-16: Post-hoc analysis examining leadership and daytime sleepiness ..............................142
Table 3-1: Descriptive statistics, reliability, and inter-rater agreement .......................................249
Table 3-2: Correlation table .........................................................................................................250
Table 3-3: Model results of transformational leadership .............................................................250
Table 3-4: Model results of laissez-faire leadership ....................................................................252
Table 3-5: Model results of over-controlling leadership ..............................................................253
Table 3-6: Model results of leadership predicting boredom .......................................................254
Table 3-7: Simple slopes analysis of transformational leadership and follow-up complications255
Table 3-8: Simple slopes analysis of laissez-faire leadership and blood loss per hour vs. expected ......256
Table 3-9: Simple slopes analysis of laissez-faire leadership and complications upon discharge
..................................................................................................................................................257
Table 3-10: Simple slopes analysis of over-controlling leadership and blood loss per hour vs. expected..................................................................................................................................................258
Table 3-11: Simple slopes analysis of over-controlling leadership and complications at follow-up ..................................................................................................................................................259
Table 3-12: Simple slopes analysis of over-controlling leadership and psychological safety ....260
Table 3-13: Simple slopes analysis of over-controlling leadership and complications at discharge (post-hoc analysis)..................................................................................................................................................261
List of Figures

Figure 2-1: Mean empathy across study periods ................................................................. 143
Figure 2-2: Mean self-efficacy across study periods .......................................................... 144
Figure 2-3: Mean daytime sleepiness across study periods ................................................. 145
Figure 2-4: Mean daytime sleepiness across study periods by gender ............................... 146
Figure 2-5: Results of empathy growth-curve analysis ...................................................... 147
Figure 2-6: Results of daytime sleepiness growth-curve analysis by gender ......................... 148
Figure 2-7: Results of growth-curve analysis of self-efficacy in leadership model .............. 149
Figure 2-8: Results of growth-curve analysis of self-efficacy in experiences model .......... 150
Figure 3-1: Illustration of full model .................................................................................. 262
Figure 3-2: Transformational leadership and complexity interaction on follow-up complications .................................................................................................................................. 263
Figure 3-3: Laissez-faire leadership and complexity interaction on blood loss per hour vs. expected ...................................................................................................................... 264
Figure 3-4: Laissez-faire leadership and complexity interaction on complications at discharge 265
Figure 3-5: Over-controlling leadership and complexity interaction on blood loss per hour vs. expected ...................................................................................................................... 266
Figure 3-6: Over-controlling leadership and complexity interaction on follow-up complications .................................................................................................................................. 267
Figure 3-7: Over-controlling leadership and complexity interaction on psychological safety 268
Figure 3-8: Over-controlling leadership and complexity interaction on complications at discharge .................................................................................................................................. 269
Chapter 1

General Introduction

To say that leadership is important is an understatement. With decades of research, and journals devoted specifically to the topic, leadership is consistently shown to influence individuals (e.g., Ilies, Nahrgang, & Morgeson, 2007), teams (e.g., Burke et al., 2006), and organizations (e.g., Koene, Vogelaar, & Soeters, 2002). At the same time, researchers (e.g., Johns, 2006; Rousseau & Fried, 2001) have called for greater consideration of context in the study of organizations – leadership research included. As Johns (2006) noted, considering context has the potential to make our research and writing better, and make our work more interesting to ourselves (as academics and researchers) and other consumers (e.g., managers and practitioners). Being able to contextualize our research allows managers to better understand and apply research findings.

Noting the need to incorporate context into research, leadership researchers have identified a hole in the leadership literature (Hannah, Uhl-Bien, Avolio, & Cavarretta, 2009): Do extreme contexts influence the expression and impact of leaders’ behaviours? In this respect, “extreme contexts” has been cited as one of the least researched areas of leadership research (Hannah et al., 2009), and the research that does exist is largely anecdotal, simulated, and retrospective.

Hannah et al. (2009) define an extreme context as one in which “risks of severe physical, psychological or material consequences (e.g., physical harm, devastation or destruction) to organizational members or their constituents exist” (p. 897). It would be hard to argue that leadership would not matter in or be shaped by such contexts – and in fact some contend that leadership becomes even more important in extreme contexts (Baran & Scott, 2010).
Because acknowledgement about the importance of leadership in extreme contexts remains unsupported, research needs to be undertaken to understand precisely how extreme contexts impact leader behaviours, and how leader behaviours impact extreme contexts and their outcomes. Extreme contexts are an environment where leadership perhaps matters the most, and gaining an understanding of how leadership functions and what makes for effective leadership is something that is essential for behavioural scientists to investigate (Campbell, Hannah, & Matthews, 2010).

Do the same leadership behaviours that help a CEO steer a corporation to financial success also influence a pilot to land a plane safely? Do these leadership behaviours aid in the capture of serial killer? Do surgeons’ leadership behaviours influence patient outcomes? Assuming that the same leadership behaviours used in traditional contexts will have the same effects in extreme contexts is premature at best, and perhaps dangerous at worst. As Osborn, Hunt and Jauch (2002) note, “a change in context changes leaders, leadership and leadership effectiveness” (p. 802).

Finally, as the world becomes more volatile, the need to study leadership in extreme contexts is more important. Even normal organizations have an increasing probability of being faced with extreme circumstances (e.g., school shootings, terrorist attacks, natural disasters), and the effective way to lead under those circumstances remains unknown. Beginning to understand leadership in extreme contexts has perhaps never been more imperative.

My dissertation focuses on the extreme context of the medical field, looking at the experiences of medical trainees (i.e. medical and nursing students) in early clinical exposure in the first study (Chapter 2), and examining the effects of surgeons’ leadership behaviours in the operating room on surgical outcomes in the second study (Chapter 3). The remainder of this
chapter includes a detailed review of the extreme context framework developed by Hannah et al. (2009) and how it applies to the medical context, with my own extension and a discussion of the specific leadership theories I have applied, followed by an overview of the two studies.

**Extreme Contexts**

Recognizing the need to examine leadership in different contexts, specifically extreme contexts, Hannah et al. (2009) proposed a framework for studying leadership in extreme contexts. The framework illustrates the factors inherent in extreme contexts, and identifies variables that might attenuate or intensify extreme contexts. Within the framework, Hannah et al. (2009) made clear that extreme contexts are not homogenous, and identify variations across extreme contexts.

Below I review Hannah et al.’s framework, and illustrate how the framework applies to the medical context considered in each study (i.e. the medical training environment, and the operating room). I then expand upon their framework by incorporating the aspect of boredom at work.

**Hannah et al.’s Extreme Context Framework**

Hannah et al. (2009) defined an extreme context as one in which members of an organization or its constituents are exposed to a risk of severe physical, psychological, or material consequences. It is important to note that, while there are similarities, an extreme context is not the same as a crisis. A crisis is defined as a low probability, suddenly occurring event that threatens high priority goals and offers little or no response time, and is often of ambiguous cause and effect (Hermann, 1969). In extreme contexts, the threats are much more specific and severe (as outlined in the definition), organizations are often able to anticipate,
prepare for, and react to extreme events, and are frequently able to identify the causes and effects of extreme events (Hannah et al., 2009).

Hannah et al. identified five factors comprising an extreme context, namely the temporal state of the organization, the magnitude of potential consequences, the probability of those consequences occurring, the proximity of organizational members, and the form of threat.

The temporal state identifies three stages an extreme organization could be experiencing: preparation, response, or recovery (or transitioning between them). Time spent in each stage could differ, and effective leadership behaviours might differ across stages. For example, leadership behaviours in the preparation stage might be focused on maintaining readiness and overcoming complacency of team members. Leadership in the response stage (similar to crisis mode) could emphasize acting quickly and providing direction and sense making (e.g., Weick, 1993). Leadership in the recovery stage might be focused on learning from events that occurred, providing emotional support, and facilitating the move back to the preparation stage. In the medical context, effort is placed in the preparation stage on prevention of an escalation to a response stage, which in this case could be a patient complication. More time could also be spent in the recovery stage, as employees monitor and focus on healing patients. In the operating room, the majority of surgeries also begin in a preparation stage (save for some emergency situations when a patient enters the operating room in a critical state). One final note for the medical context: Employees may be experiencing all of the temporal states almost simultaneously, as different patients may be in different states at the same time, which is different than the temporal state dynamics experienced by, for example, fire fighters who are focused on one fire at a time.
Three factors vary across extreme contexts, and help explain why effective leadership in a military context may differ from a surgical context. These three factors are form of threat, magnitude of consequences, and probability of the consequence occurring. Hannah et al. (2009) separated threats into physical, psychological, and material. In the medical setting, the forms of threat are quite obvious. Most threats would be physical: death or injury of the patient, injury and exposure to disease for team members, and physical exhaustion. Psychological threats could include psychological damage to a patient in the event that a surgery does not go as planned, and psychological damage of team members being involved in or witnessing traumatic events. Material threats would be typified by the threat of legal action by the patient or the patient’s family following adverse outcomes.

The magnitude of consequences serves as an indicator of the potential extreme nature of the situation. In the case of surgery and the practice of medicine, a patient’s death would be more severe than the loss of a limb, and that would be more severe than a scar.

But magnitude alone does not make a situation extreme – there must be an adequate probability of these consequences occurring as well. Together, the magnitude and probability could affect how team members behave, what leadership behaviours might be appropriate, and how employees could react. In the training of students, each patient presents a unique set of risks and conditions, which could influence how they are treated by medical caregivers, and how those caregivers react and behave. The magnitude and probability of any consequences will also vary across surgeries: a tonsillectomy with a risk of bleeding and infection (U.S. National Library of Medicine, 2012a) cannot be compared with a heart bypass surgery, with a risk of major organ failure and stroke (U.S. National Library of Medicine, 2012b). Surgeons are able to identify very specific potential negative outcomes and probabilities to each of their patients prior
to a surgery. These negative outcomes and probabilities are likely high on the minds of a surgical team prior to entering a surgery, and can certainly be expected to impact how they will behave during that surgery.

Finally, the proximity of team members to the threat and to each other is Hannah et al.’s final factor. Proximity is split into four forms: physical proximity to the event, physical distance between team members, psychological proximity to the event, and psychosocial distance between team members. Physical proximity refers to how physically close team members are to a potential extreme event. Physical distance considers by how much leaders and team members are physically separated from each other, as this has been shown to influence the impact, dynamics and quality of leadership interactions (e.g., Bass, 1998). Psychological proximity refers to the proximity of team members to the extreme event “in the sense that those members affiliated with leaders and followers are potentially in harm's way” (Hannah et al., 2009, p. 906). Psychosocial distance refers to the social distance between leaders and their team members, and Hannah et al. (2009) expect this will be a factor in determining the levels of trust and cohesion in a team. In the training of students, the physical proximity and distance is generally confined to the department or medical centre. Psychosocial distance may be greater, as trainees are newcomers to the field, and thus hold less status and could be more strongly influenced by role models. Similarly, medical trainees may be more susceptible to psychological ramifications, as the psychological distance may be lower in their early clinical experiences. In terms of surgery, team members are generally physically and psychologically close to each other and to the threat. Physical distance is also more constant, but psychosocial distance will vary: Surgical teams often have very tightly defined roles, and members often wear surgical masks, but there is a
chance that some team members have worked together before, and simple behaviours such as team introductions that reduce distance also vary.

In all extreme contexts, Hannah et al. proposed several factors that would buffer any negative effects of the extreme context that might help members and leaders perform more effectively, as well as factors that might exacerbate performance problems in an extreme context – the authors call these “attenuators” and “intensifiers” respectively. Attenuators include psychological resources, social resources, and organizational resources. Psychological resources such as optimism, positive emotions, self-efficacy, and resilience improve adaptability when the body is under stress and may serve to improve performance. Social resources such as physical response networks, as well as cohesion, morale, and supportive social networks improve response, functioning, and performance in extreme situations. Finally, organizational resources, both tangible (e.g., technology, financial resources) and intangible (e.g., adaptability, perceived support) serve to improve performance in extreme contexts.

On the other hand, intensifiers such as time (urgency, frequency, duration) and complexity act to increase pressures and decrease performance in extreme contexts. Urgency reduces the amount of time to think through decisions and act upon impending threats. More frequent extreme events allow an organization to better adapt and learn, as do longer duration extreme events (though too much frequency could potentially harm adaptation and learning, as employees have little time to recover and learn from events). Each surgery will vary in the different time pressures present. However, the medical field is accustomed to a higher frequency of extreme events, so learning in this setting may be higher (e.g., Vashdi, Bamberger, & Erez, 2013). Finally, complexity is defined as the level of dynamism of an environment due to interrelated variables acting in unexpected ways. Even seemingly small events in some
environments interact in such a way to cause highly unpredictable outcomes. Working with the human body is a constant, but important intensifier in the medical context. As is the case in aviation (Amalberti & Wioland, 1997), or nuclear power plants (Perrow, 1984), it is rarely ever only one mishap or error that escalates into a crisis, but a series of errors that interact to worsen the situation (Perrow, 1984). This is expected to be the case in the medical field (Helmreich, 2000).

**Boredom at Work**

To further develop Hannah et al.’s framework, I add boredom at work. While boredom seems to be almost antithetical to an extreme context, existing literature suggests it may be an important factor to consider in such contexts. Hannah et al. (2009) touched upon boredom when they discuss complacency during the preparation stage (e.g., “that will never actually happen”) as an obstacle that leaders face in extreme contexts. While not identical to complacency, boredom might help explain behaviors within the extreme context framework. I also believe that boredom could influence not only the preparation phase, but be a factor at any stage of the cycle. Although largely ignored by organizational researchers (Fisher, 1998), boredom has been studied in more traditional contexts, and the results raise concerns not only for its negative impacts in normal settings, but the potential harm boredom could cause in extreme contexts. By studying boredom in extreme contexts, new meaning could potentially be given to the oft-used term, *dying of boredom*.

Fisher (1991) defined boredom as “a transient affective state in which the individual feels a pervasive lack of interest in the current activity” (p. 3). Fisher also noted that it often takes considerable effort to refocus and maintain attention on the task at hand, and Smith and Ellsworth (1985) found that those suffering from boredom tended to divert attention away from
the cause of boredom rather than try to increase attention paid to it. Researchers have uncovered a number of causes of boredom in the workplace that Fisher categorized into task, work, environment, and personal causes.

Tasks that are low-skill, repetitive, and involve monitoring a situation while looking for infrequent events (e.g., lifeguarding, quality control) elicit boredom (Fisher, 1987). While surgery is certainly not a low-skill profession, many high context environments do involve waiting (or preparing) for infrequent events (e.g., waiting for test results, routine rounds). Other task-related causes of boredom include having little or nothing to do, especially following a period of high workload. This contrast effect could be especially salient in an extreme context, where swings in workload and excitement are likely to be more extreme. Generally speaking, boredom arises when individuals’ subjective appraisal of the current task does not meet their expectations of stimulation. One would expect that individuals working in extreme contexts would be anticipating high levels of stimulation and excitement – but instead are met with routine. The contrast is such that boredom in extreme contexts is likely to be salient.

Work environment factors related to boredom include having “boring” coworkers, policies and programs that inhibit individual control, and other coworkers expressing boredom toward an activity (Fisher, 1991; Griffin, 1983). The latter factor can have large effects as coworkers and superiors have a great influence on how employees perceive a job – in this sense boredom can be contagious (Barsade, 2002) and spread throughout a team, and the leader might exert significant influence over the boredom.

Two individual-level factors influence boredom. As individuals’ performance capacity increases – finding the same task easier or less challenging – their susceptibility to boredom may increase (Drory, 1982). This misfit between skills and task has received empirical support, but
only for simple tasks (Fisher, 1998). While there is no evidence of this relationship in more complex tasks, it would be expected that medical professionals (e.g., nurses, surgeons) have high performance capacity, and as task difficulty decreases, boredom could increase. Personality variables have also been associated with boredom: Extraverts become bored more easily (Guest, Williams, & Dewe, 1978), as do sensation-seeking individuals (Best & Kilpatrick, 1977; Zuckerman, 1979). Sensation-seeking individuals are also more likely to pursue careers in extreme contexts (Fisher, 1991). Together, these results suggest that some individuals are likely more susceptible to boredom: sensation-seekers with high performance capacities who experience extreme high and low workload situations. Thus it is reasonable to expect that boredom in extreme contexts not only exists, but might even be more prevalent than in traditional contexts. In recent research by Smith-Jentsch et al. (2012) in which 12 NASA astronauts were interviewed, boredom was a common complaint of astronauts on long-haul missions.

Finally, there would be no need to consider boredom if its effects were inconsequential, but this is not the case. Research in traditional settings suggests that boredom creates inattention and withdrawal from tasks (Spector et al., 2006), which also increases the chances of errors (Cox, 1980; Drory, 1982; O'Hanlon, 1981). As such, boredom has been associated with safety-related outcomes such as workplace injuries amongst young workers (Frone, 1998), and minor accidents (e.g. damage to loading docks) caused by truck drivers (Drory, 1982). In addition to physical consequences, boredom is also associated with absenteeism (Drory, 1982), and related to affective responses such as agitation, restlessness, and emotional distress (Robinson, 1975). These findings suggest that the consequences of boredom, when translated to extreme contexts, could negatively impact psychological states, and create safety and performance problems.
Testing “Traditional” Leadership Theories in Extreme Contexts

Throughout their framework, Hannah et al. (2009) provided behavioural prescriptions for what they expected to be effective leadership under different circumstances, without specifically referring to the traditional leadership theories that organizational researchers are familiar with (e.g. transformational leadership, abusive supervision). Similarly, other researchers have examined specific extreme contexts and created corresponding taxonomies of leadership: For example, Parker, Yule, Flin, and McKinley (2011) developed a taxonomy of leadership for the operating room, highlighting what they believed were the key characteristics and behaviours of an effective surgeon, and Wong, Bliese, and McGurk (2003) reviewed the literature on military leadership, finding a similar array of characteristics of effective leadership.

These taxonomies are helpful in distinguishing some of the unique contextual capabilities and characteristics that may be needed in extreme contexts (e.g., surgical competency), but they also discount the decades of research on leadership from traditional contexts. So instead of developing new theories of leadership, my dissertation applies existing theories that I suggest translate to extreme contexts. Furthermore, the taxonomies described above neglected to consider what constitutes ineffective leadership in their efforts to define what is effective. Thus I also investigate the potential consequences of ineffective leadership, again using existing leadership behaviours that have proven ineffective in traditional contexts. As this is one of the first empirical examinations of leadership in extreme contexts, I have attempted to gain a well-rounded view of the effects of leadership in each study, by studying one form of what is considered positive leadership (i.e., transformational leadership), one form to represent the
absence of leadership (i.e., laissez-faire leadership), and one form of what is considered negative leadership (i.e., abusive supervision in Study 1, over-controlling leadership in Study 2) ¹.

Transformational leadership was selected as the positive leadership behaviour due in part to the expansive existing literature that demonstrates its variety of positive effects in traditional organizations (e.g., Kelloway, Turner, Barling, & Loughlin, 2012; Wang, Courtright, & Colbert, 2011), that help explain how the behaviours and effects might extend to extreme contexts. Though it should be noted that other positive leadership theories, such as leader-member exchange (e.g., Graen & Uhl-Bien, 1995), are often strongly correlated (Dancey & Reidy, 2004) with transformational leadership (e.g., leader-member exchange: \( r = .73 \), Dulebohn, Bommer, Liden, Brouer, & Ferris, 2012; \( r = .71 \), Wang, Law, Hackett, Wang, & Chen, 2005), suggesting that the behaviours in each of these theories are highly similar. Laissez-faire leadership was selected as the form of absence of leadership, as it is the most widely-studied leadership theory that effectively captures the behaviours of nonleadership. Finally, abusive supervision was chosen to represent negative leadership in Study 1 because it has been suggested that nurses and medical students experience abusive-like behaviours on the job (e.g., Cook, Arora, Rasinki, Curlin, & Yoon, 2014; Ferns & Meerabeau, 2007), and findings from traditional organizations suggest detrimental effects of working for an abusive supervisor (e.g., Aryee, Chen, Sun & Debrah, 2007, Tepper, 2000). Abusive supervision is also the most commonly studied type of “destructive leadership”, and produces similar effects to other negative types of leadership (e.g., Schyns & Schilling, 2013). Over-controlling leadership was selected as the negative form of leadership for Study 2, as since it was an observational study, I expected (through anecdotal

¹ While transactional leadership behaviours are also commonly discussed in the spectrum of leadership behaviours (e.g., Avolio & Bass, 2001), it was excluded from this research because of (a) its often strong correlations with other behaviours being studied (e.g., Judge & Piccolo, 2004), and (b) it is conceptualized as management, opposed to leadership.
reports, and findings showing the prevalence of micromanagement; Box, 2012) that it would be a more frequent behaviour in the operating room than abusive supervision. While it is a newer construct, early findings suggest its negative effects (e.g., Dupre & Barling, 2006).

A broader examination of the definitions of each theory and their corresponding effects on the outcomes of interest appear in Chapters 2 and 3.

**A Test “Within” Versus a Test “of” Extreme Contexts**

At this point, I emphasize that my research is not a test of Hannah et al.’s extreme context framework. Instead, I will use Hannah et al.’s framework as a basis for understanding how different leadership behaviours and contextual variables impact outcomes within the extreme context of the surgical operating room and broader medical training environment.

**Overview of Manuscripts**

**Manuscript 1**

Manuscript 1 examines how the psychological variables of empathy, self-efficacy, and daytime sleepiness change throughout the first sustained clinical experience of medical and nursing students. Empathy was expected to decrease overall, while self-efficacy and daytime sleepiness were expected to increase, during this time period. I also propose a number of variables to explain the changes, beyond the impact of time. These variables include leadership behaviours of supervisors (i.e. transformational leadership, laissez-faire leadership, abusive supervision), patient morbidity and mortality, beneficiary contact, and boredom. I predict that transformational leadership will increase empathy and self-efficacy, and decrease feelings of sleepiness, and that laissez-faire leadership and abusive supervision will decrease empathy and self-efficacy, and increase feelings of sleepiness. The remaining variables represent experiences specific to the medical context that students could have. Specifically, I predict that contact with
patient morbidity and mortality will decrease empathy and self-efficacy, as well as increase daytime sleepiness. Beneficiary contact is predicted to increase empathy and self-efficacy, and decrease feelings of sleepiness. Boredom is predicted to decrease empathy and self-efficacy. These hypotheses are tested on a longitudinal sample of medical and nursing students over an eight-month period, using growth-curve models.

**Manuscript 2**

Manuscript 2 examines the outcomes of surgeons’ leadership behaviours (i.e. transformational leadership, laissez-faire leadership, over-controlling leadership) in the operating room on both proximal (i.e. errors, blood loss) and distal surgical outcomes (i.e. complications during recovery at the hospital, complications post-discharge measured at least a month following the surgery). I predict that leadership exerts its influence through two mediators: psychological safety, and boredom. I also predict that the complexity of the surgery will serve as a moderator, amplifying or weakening relationships in the model. I test my hypotheses using a large sample of surgeries with data from multiple sources, including observational measures of leadership, survey measures of meditational and moderator variables, observational measures of errors, and objective/archival measures of blood loss and complications.

**Summary**

In summary, my dissertation explores the effects of leadership and other contextual elements on physical and psychological consequences in the extreme context of healthcare. The need to understand leadership effects in extreme contexts grows increasingly important, and my dissertation begins to bridge the gap between the broad literature on leadership in traditional contexts and what we need to know about the transferability of that research into more unstable environments. This dissertation also provides new practical knowledge that practitioners and
policy-makers will be able to use to further improve our healthcare system, and more broadly, anyone working in or affected by other extreme contexts.
1.1 References


Chapter 2

A Longitudinal Study of Psychological Change in Extreme Contexts

Abstract

To date, most research in extreme contexts has focused on physical consequences (e.g., accidents, errors, injury). This study attempts instead to understand some of the psychological consequences faced by those working in such contexts. Specifically, this study focuses on the medical context, and the development of empathy, self-efficacy, and daytime sleepiness over time. Following a sample of medical and nursing students over eight months during their first sustained clinical experience, I predicted that their experiences with transformational leadership, laissez-faire leadership, and abusive supervision, as well as morbidity and mortality of patients, beneficiary contact, and boredom would explain shifts in students’ empathy, self-efficacy beliefs, and daytime sleepiness, over and above the effects of time. Results showed that across five six-week time periods and 166 person-time data points, (a) laissez-faire leadership negatively predicted empathy, (b) transformational leadership and depth of beneficiary experiences positively predicted self-efficacy beliefs, and laissez-faire leadership and morbidity and mortality experiences negatively predicted self-efficacy beliefs, and (c) laissez-faire leadership, and morbidity and mortality experiences, positively predicted daytime sleepiness, and depth of beneficiary contact negatively predicted daytime sleepiness, all in the predicted direction. Transformational leadership and frequency of beneficiary contact also predicted daytime sleepiness, but positively, which was in the opposite direction than predicted. Results are discussed, along with potential avenues of future research and possible practical implications.
2.1 Introduction

An extreme context is defined as one in which “risks of severe physical, psychological or material consequences (e.g., physical harm, devastation or destruction) to organizational members or their constituents exist” (Hannah, Uhl-Bien, Avolio, & Cavarretta, 2009, p. 897). Most research in extreme contexts, such as law enforcement, firefighting, combat, and surgery have invariably focused on physical consequences (e.g., death or injury of employees or constituents), as these are understandably the most salient and important consequences facing individuals and constituents in such contexts. Hospitals and physicians are also evaluated on such metrics as cure rates, mortality rates, and errors (Garrett, 2007) – all physical consequences.

Of equal importance are the psychological consequences faced by employees and constituents in extreme contexts; however, to date, these consequences have not received sufficient attention. This study aims to remedy this by examining the effects of leadership and contextual experiences on psychological consequences of medical and nursing students. These employees face an environment each day that is wrought with uncertainty, and frequently see (and could even cause) suffering, death, and destruction.

To assume that these contexts only have important physical consequences for employees and constituents is not only mistaken, it could be dangerous: Psychological consequences of working within an extreme context could impact the health and well-being of employees, which in turn could affect employees’ effects on physical consequences (e.g., errors, patient safety). Yet other than perhaps research on burnout (e.g., Adriaenssens, De Gucht, & Maes, 2015; Embriaco, Papazian, Kentish-Barnes, Pochard, & Azoulay, 2007), little is understood about the psychological consequences associated with working in an extreme context, and how working in an extreme context affects the development of these psychological experiences.
In this study, I will examine the psychological impact of early professional experiences within extreme contexts, namely psychological consequences faced by student physicians and nurses while in a vulnerable stage of their training, experiencing their first sustained patient interactions while being socialized into the greater medical environment. More specifically, I investigate (a) the development of empathy, self-efficacy, and daytime sleepiness over an eight-month period, and (b) how leadership (i.e. transformational leadership, laissez-faire leadership, abusive supervision) and contextual experiences (e.g., morbidity and mortality of patients, beneficiary contact, boredom) influence this development.

This sample provides an excellent opportunity to examine the development of psychological consequences during newcomer socialization. Newcomer socialization is the process of acquiring “the knowledge, skills, behaviours, and attitudes required for effective participation in an organization” (Allen, McManus, & Russell, 1999, pg. 456). Researchers have suggested that during this process, employees are more susceptible to influence, as they are actively engaged in navigating and understanding their environment through interactions with peers and supervisors (Jablin, 2001; Scott & Myers, 2005; Van Maanen, 1978). This susceptibility to influence may be further exacerbated by a tendency for newcomers to view norms as rigid (opposed to flexible), resulting in an overwhelming need to conform (Liu, Wang, Bamberger, Shi, & Bacharach, 2015). Thus, I would expect that the effects of leadership and contextual experiences have an important effect on the development of empathy, self-efficacy, and daytime sleepiness during the first year of clinical experience. Studying newcomers also makes this a strong test of the theoretical model that I develop and explain below.

Newcomer socialization occurs through formal and informal organizational influences (e.g., training, mentoring; e.g., Bauer, Bodner, Erdogan, Truxillo, & Tucker, 2007), and through
the proactive behaviours of newcomers (e.g., information seeking, observation, seeking feedback) (e.g., Ashford & Black, 1996; Miller, 1996; Morrison, 1993). Through these processes, successful newcomer socialization is positively associated with job performance, occupational commitment, and intentions to stay (e.g., Bauer et al., 2007). However, ineffective socialization processes can affect the development of potentially dysfunctional outcomes such as heavy drinking (Liu et al., 2015), and emotional detachment (e.g., Scott & Myers, 2005). This is important, because newcomer socialization can have lasting impacts on individuals throughout their careers, as the knowledge gained becomes the basis for everyday behaviours (e.g., Ashforth & Saks, 1996; Schein, 1968). Understanding the effects of the socialization of nursing and medical students is not only important with respect to immediate outcomes — the focus of my study — but for the students’ long-term career success.

Below, I review the research on each of the psychological variables (i.e. empathy, self-efficacy, daytime sleepiness) in both traditional and extreme contexts, followed by a discussion of their potential antecedents. I then outline the methodology used to test the hypothesized relationships, and present and discuss the results.

**The Nature and Development of Empathy**

Empathy has a rich history of research, but also varying definitions, which can be characterized as either affective or cognitive (Davis, 2006). The former suggests that empathy is an individual’s response to the perceived emotional experience of another (Mehribian & Epstein, 1972). The latter conceptualizes empathy as an individual’s ability to accurately assess another’s emotional state and to be aware of his or her effect on others (Becker & Sands, 1988). While distinct and sometimes treated as a multidimensional construct (Cox et al., 2012), these two characterizations are often combined and a unidimensional approach to empathy taken
For the purposes of this research, I define empathy as including both the cognitive ability to accurately perceive emotions and feelings of another, and the accompanying emotional response to those emotions and feelings.

The need for both cognitive and affective empathy in a complete conceptualization of empathy can be traced to one of the earlier researchers of empathy (Davis, 1980), who theorized that empathy is composed of four sub-factors: perspective-taking, fantasy, empathic concern, and personal distress. Perspective-taking and fantasy are sub-factors of cognitive empathy. Perspective-taking reflects the ability to anticipate the behaviour and reactions of others – to understand another individual’s psychological point-of-view. Fantasy reflects the tendency to imagine the thoughts and feelings of fictitious characters (e.g., in a book or movie). Empathic concern and personal distress are sub-factors of affective empathy. Empathic concern is the emotional, sympathetic response to others in distress. Finally, personal distress encompasses internal feelings and anxiety in tense interpersonal situations. Researchers do not agree whether personal distress (Minio-Paluello, Lombardo, Chakrabarti, Wheelwright, & Baron-Cohen, 2009; Zhou, Valiente, & Eisenberg, 2003), or fantasy (Baron-Cohen & Wheelwright, 2004) are components of empathy, suggesting that they may be related to, but not part of empathy. Thus, there is consensus that empathy is reflected in perspective-taking (cognitive empathy) and empathic concern (affective empathy), and this is the approach that I follow in my research.

Empathy is a sought-after ability both in and out of the workplace. While I use a unidimensional approach to empathy, combining both cognitive and affective empathy, I would be remiss if I did not briefly discuss the positive effects of empathy found by researchers who did not follow this approach. In general, individuals high in affective empathy are more likely to
exhibit greater self-esteem, social functioning, communication skills, feelings of compassion, a tendency for forgiveness, and less anxiety (Davis, 1983; 2000; Regan & Totten, 1975). Those high in cognitive empathy are also better able to handle interpersonal conflict, inhibit aggression (Richardson, Hammock, Smith, Gardner, & Signo, 1994), elicit helping and other prosocial behaviours (Davis, 2006), and be more cooperative (Becker & Sands, 1988; Davis, 1983; Eisenberg & Miller, 1987). In the workplace, affective empathy has been related to more effective leadership (Kellett, Humphrey, & Sleeth, 2002; 2006), and managers high in affective empathy elicit positive affect and well-being in their employees (Scott, Colquitt, Paddock, & Judge, 2010). Cognitive empathy is related to positive diversity attitudes (Madera, Neal, & Dawson, 2011), and lower relational conflict (LeBlanc, Gilin, Calnan, & Solarz, 2012).

However, there are a few negative consequences of high empathy that are worth noting. In general, empathy has been associated with higher levels of guilt (Stangert, Kavussanu, & Ring, 2012), depressive symptoms in individuals with low psychosocial resources (Schieman & Turner, 2001), and altruism toward a single individual to the detriment of others (e.g., Batson, 1998; Batson, Klein, Highberger, & Shaw, 1995). Further, in an experimental manipulation, Loggia, Mogil, and Bushnill (2008) showed that participants in the high empathy group experienced altered physical pain perception: Those high in empathy felt more intense and unpleasant pain in the experiment than those with low empathy. In the workplace, empathy may harm employees’ success, as “eavesdropping” on negative emotions is associated with lower workplace ratings by colleagues and supervisors of performance and organizational fit (Elfenbien & Ambady, 2002). Also applicable to the medical workplace, researchers found that empathy of ambulance paramedics was related to higher susceptibility during traumatic events, and corresponding symptoms of trauma such as sleeplessness, anger, and frustration (Regehr,
Goldberg, & Hughes, 2002). So while overall empathy is a desirable characteristic, leaders should be aware of some of the negative consequences faced by those with high empathy.

As I noted earlier, both dimensions are needed for a full understanding of empathy. In research examining the experience of empathy, Kerem, Fishman and Josselson (2001) concluded that both aspects of empathy play a role in individuals’ experience of empathy. These authors suggested that empathy rarely exists without some level of cognitive empathy, and that affective empathy serves to produce an even more meaningful experience.

**Physician and Nurse Empathy**

The practice of medicine requires frequent interactions with a variety of different patients, each presenting unique problems, personalities, and histories. Communication is needed to diagnose, provide treatment for, and give ongoing support and care for these individuals. Thus, empathy is seen as a desirable trait in physicians (Haslam, 2007) and nurses (Williams & Stickley, 2010) for the proper care of this population. There is a large body of research on physician-specific empathy and medical-related outcomes, and empathy is a component of most medical training – the Royal College of Physicians and Surgeons of Canada includes empathy as a skill that needs to be cultivated during medical training, and further developed throughout a physician’s career (The CanMEDS Framework, 2013).

Physician-specific empathy is defined in much the same way as general empathy with one exception: Physician empathy is most often targeted toward patients rather than a generalized “other”. Thus, Colliver, Conlee, Velhurst and Dorsey (2010) defined physician empathy as the physician’s “cognitive and vicarious understanding of the patient as a person” (pg. 588), and Hojat et al. (2002) defined physician empathy as “an ability to understand the
patient’s inner experiences and perspective and a capability to communicate this understanding” (pg. 1564).

The medical focus on empathy is well justified. Physician and nurse empathy are positively related to many medical outcomes, including faster and more effective patient recovery and health (e.g., Burns & Nolen-Hoeksema, 1992; Hojat, et al., 2011; Kaplan, Greenfield, & Ware, 1989; Stewart, 1995). This is largely due to the positive influence of empathy on the physician-patient relationship: When physicians are more empathetic, patients are more likely to (a) communicate and be more forthcoming with symptoms and concerns (Halpern, 2001; Kim, Kaplowitz, & Johnston, 2004; Larson & Yao, 2005), and (b) place greater trust in physicians and follow their instructions (Kim, et al., 2004). In fact, patients rate empathy as the second-most important quality in a physician (after competence; Colliver, Willis, Robbs, Cohen, & Swartz, 1998). Beyond interactions with patients, empathy is associated with actual physician performance and errors (Hojat, et al., 2002; West, et al., 2006), and nurse and physician well-being (Bourgault et al., 2014; Larson & Yao, 2005). With such positive consequences, interest in the development of empathy in medical professionals has gained attention, but findings have shown that empathy tends to decline throughout medical training, as discussed below.

**Change in Empathy**

“Students often begin their training with considerable empathy and altruism; however, they are taught in medical school to focus on more objective aspects of patient care; to cure and rule out disease. The emotional distance between doctor and patient becomes worse during residency. The isolation, long hours of service, chronic lack of sleep, fear of failure and constant exposure to tragedy serve to extinguish any empathy and altruism that may be left.” (Seaberg, Godwin, & Perry, 2000, p. 1433)
Despite the positive consequences of physician and nurse empathy, many researchers have found a significant decline in levels of empathy during medical training, both in medical school (e.g., Hojat et al., 2004; Neumann, et al., 2011), residency (e.g., Colliver et al., 2010; Roh, Hahm, Lee, & Suh, 2010), and undergraduate nursing (e.g., Ward, Cody, Schaal, & Hojat, 2012). The steepest declines often emerge during students’ first sustained experiences with patients (e.g., Austin, Evans, Magnus, & O’Hanlon, 2007; Hojat, et al., 2009; Ward et al., 2012). Some studies show no significant declines in empathy (e.g., Rosenthal et al., 2011; Williams et al., 2014), though Pedersen (2010) argued that even the stunting of empathy (i.e. non-growth) is problematic: medical educators should target increases in empathy during medical training.

The precise cause of any decline has not been investigated, but researchers have proposed a number of reasons: For example, empathy decline might be a result of increased cynicism of the value of the medical profession, or hardening to protect oneself (Colliver et al., 2010). Others argue that medical training is often embedded in a culture of detachment, and the belief that emotions cloud objectivity (Coulehan & Williams, 2001; Hojat, et al., 2002; Larson & Yao, 2005). Many students are taught with a dominant focus on the biomedical scientific approach of medicine, with a lack of consideration of empathy (Pedersen, 2010; Spiro, McCrea Curnen, Peschel, & St. James, 1993). Given the increased susceptibility to influence during newcomer socialization described earlier, the early messages that empathy is not important may predict a decline in empathy early in training. Finally, others believe that a decrease in empathy could be attributed to a heavy workload, fatigue, and burnout (Bellini, Baime, & Shea, 2002; Larson & Yao, 2005).

Given these different findings and suggestions, I predict that:

**Hypothesis 1:** Empathy of medical trainees will decrease over the first eight months of clinical experience.
The Nature and Development of Self-efficacy

Self-efficacy reflects “judgments of how well one can execute courses of action required to deal with prospective situations” (Bandura, 1982, p. 122). Believing that you can successfully execute a given task influences the behaviours you enact and your persistence of action (Bandura, 1986). Self-efficacy is a key predictor of intrinsic motivation and performance (e.g., Bandura, 1993; Bandura & Schunk, 1981; Stajkovic & Luthans, 1998), and mediates the effects of knowledge on action (Bandura, 1982).

Since the inception of the theory, self-efficacy has garnered much attention by organizational researchers who have identified its numerous positive organizational outcomes. The most common outcome associated with self-efficacy is individual performance (e.g., Judge & Bono, 2001; Sadri & Robertson, 1993; Stajkovic & Luthans, 1998). Gist (1987) even found self-efficacy to be a better predictor of performance than past performance. Goal-setting partially mediated the relationship found by Gist, as those with higher self-efficacy were more likely to set more difficult goals for themselves. Other outcomes associated with self-efficacy include work withdrawal/engagement, organizational commitment, and job satisfaction (Judge & Bono, 2001; Walumbwa, Wang, Lawlwer, & Shi, 2004).

Bandura identified four sources through which self-efficacy develops. The sources include (in decreasing order of importance) performance accomplishments, vicarious experiences, verbal persuasion, and physiological/emotional states. First, performance accomplishments (especially repeated successes) are thought to be the most influential predictor (Bandura, 1977), as they are the most experiential. Accomplishments increase mastery expectations, and failures decrease self-efficacy beliefs. Bandura (1977) explained that accomplishments and failures have an especially strong impact on self-efficacy early on in an
individual’s exposure to a new task, making self-efficacy within medical and nursing students facing their first sustained clinical experience especially interesting. Second, vicarious experiences (Bandura, 1977) are effective when individuals see others successfully complete a task. This effect is stronger when those completing the task are perceived as similar to the individuals whose self-efficacy is in question. Again, this is important to medical trainees who are continually exposed to a range of peers and role models, both effective and ineffective. Third, verbal persuasion is effective in fostering self-efficacy, especially when the source is credible, as would be the case with respected teachers. Finally, physiological and emotional states such as fear and anxiety may impact individuals’ perceived ability to complete a task, thus decreasing their self-efficacy judgments.

Empirical support has emerged for these four sources as antecedents of self-efficacy, for example, the relationship between past performance and self-efficacy (Bandura, 1982; Chen & Bliese, 2002; Tierney & Farmer, 2002), and the effect of psychological strain (an emotional state) on self-efficacy (Chen & Bliese, 2002). More specifically to my study, the antecedents of self-efficacy have also been examined indirectly in research focusing on the impact of leadership on follower self-efficacy. Leaders are in a position to foster a trusting environment and provide the feedback and encouragement needed to improve followers’ self-efficacy. Supportive leadership, transformational leadership, and positive leadership climates are positively linked to self-efficacy at both the individual and group level (Chen & Bliese, 2002; Choi, Price, & Vinokur, 2003; Jung & Sosik, 2002). Leadership also facilitates the development of self-efficacy by providing greater role clarity (Chen & Bliese, 2002), setting challenging goals for employees, and providing job enrichment (Parker, 1998; Tierney & Farmer, 2002).

**Physician and Nurse Self-efficacy**
Unlike empathy, self-efficacy has not attracted much attention in medical research or extreme contexts more generally, as either a predictor or outcome variable. As a predictor variable, however, I expect that the motivation and persistence created by self-efficacy may be related to relevant outcomes in the medical context, making it relevant for students and full-time physicians and nurses. Indeed, Artino Jr. et al. (2012) noted that high self-efficacy is particularly important to success in medicine, where there are many challenges and obstacles to successful performance. For example, self-efficacy is related to medical student clinical performance (e.g., Artino, La Rochelle, & Durning, 2010; Artino, Hemmer, & Durning, 2011; Mavis, 2001), and to the confidence and ability of nursing students to undertake advanced nursing roles (Hayes, 1998). Thus an understanding of how self-efficacy develops over time is needed, especially at vulnerable times such as during newcomer socialization.

Given the importance of early performance accomplishments and failures on the development of self-efficacy, the extent of between-individual variability makes predictions of mean change difficult. As discussed above, Bandura’s theory suggests that both accomplishments and failures affect self-efficacy, especially early on in experiences with a new task. This suggests that students enjoying early success experiences would be expected to increase their self-efficacy, and those experiencing early failures would decrease their self-efficacy. Findings summarized by Baumeister, Bratslavsky, Finkenauer, and Vohs (2001) as the “bad is stronger than good” principle suggest that failures may even have a stronger effect on self-efficacy than successes. However, as Artino et al., (2012) point out, the environment faced by these students is full of challenges to successful performance. While this might suggest that more students may experience decreased self-efficacy, I propose that the opposite will be true. Because of the new environment and challenges to success, performance accomplishments may
be more celebrated, consequently having a stronger effect on self-efficacy. Failure, while disappointing, may have a weaker effect on self-efficacy. Accomplishments and failures considered together, a mean upward trend of self-efficacy would be expected. Finally, research suggests that successful newcomer socialization is associated with increases in self-efficacy, as newcomer socialization helps reduce uncertainty (e.g., Bauer et al., 2007). In summary, a mean increase in self-efficacy over time is expected.

While the development of self-efficacy has not been examined in medical and nursing students, limited evidence supports the notion that self-efficacy may increase over time. In a cross-sectional study, Artino et al. (2012) investigated the levels of self-efficacy across all years of medical school. Significant differences emerged, as senior students tended to have higher levels of self-efficacy, suggesting an upward trend in self-efficacy across the years of medical training (though this was a cross-sectional study and did not examine intra-individual change).

While different factors influence the development of self-efficacy, overall I predict that:

**Hypothesis 2:** Self-efficacy will increase over the first eight months of clinical experience.

The Nature and Development of Daytime Sleepiness

Sleep is a personal resource and source of energy that spans physical, cognitive, and emotional realms: Both sleep quality and sleep quantity impact physical functioning, mental cognition, and individuals’ emotional states. Sleep can be defined as “a state of immobility with greatly reduced responsiveness, which can be distinguished from coma or anesthesia by its rapid reversibility” (Siegel, 2005, p. 1264). The propensity to sleep increases the longer one has been awake, and decreases the longer one has been asleep (Borbely, 2009; Siegel, 2005).

Time spent sleeping, or hours of sleep lost, is a common focus of sleep research (e.g., (Barnes, Schaubroeck, Huth, & Ghumman, 2011; Vodorholzer, Al-Shajlawi, Weske, Feige, &
However, when discussing sleep, the concern is not only with the quantity of sleep, but also the individual’s subjective sleep quality. The number of hours individuals sleep is not necessarily related to the quality of their sleep; sleep quality is influenced by a number of other factors (Johns, 1991). For this research, I will focus on self-perceived sleep quality as the outcome of interest, specifically subjective daytime sleepiness. Daytime sleepiness is a common measure used to diagnose underlying sleep problems (Johns, 1991), and is correlated with sleep quantity and quality (e.g., Johns, 1991; Pilcher, Ginter, & Sadowsky, 1997; Roehrs, Zorick, Wittig, Conway & Roth, 1989). Sleepiness reflects the “the transitional state between wakefulness and sleep” (Shen, Barbera, & Shapiro, 2006, p. 64). Daytime sleepiness measures these feelings, and is associated with a greater tendency to fall asleep, lower cognitive and motor performance, and changes in mood (Shen et al., 2006).

Daytime sleepiness has been gaining more attention recently in terms of organizational research. Most workplace research has focused on the impact of sleepiness on occupational safety (e.g., Mullins, Cortina, Drake & Dalal, 2014): Daytime sleepiness is positively related to workplace accidents and injuries (Lindberg, Carter, Gislason, & Janson, 2001; Melamed & Oksenberg, 2002; Uehli et al., 2014), and negatively related to safety behaviour (DeArmond & Chen, 2009). Daytime sleepiness is also related to decreased motivation and coherence, increased irritability (Dalbokova, Tzenova, & Ognjanova, 1995), difficulties concentrating, inefficient decision-making and problem-solving (Alapin et al., 2000), and absenteeism (Hackett & Bycio, 1996).

**Physician & Nurse Daytime Sleepiness**

While it has been an issue for decades, sleep is an ever-increasing concern for medical professionals and trainees. Researchers have investigated the prevalence of sleep problems in
residents, medical students, and nursing students, often finding above-average levels of sleepiness (Howard, Gaba, Rosekind, & Zarcone, 2002) and below-average hours of sleep per night (Ford & Wentz, 1984). These conditions are also associated with a greater incidence of depression, anger (Ford & Wentz, 1984), mood disturbances (Harrison & Horne, 2000), and physical health (Hughes & Rogers, 2004) in medical professionals.

Daytime sleepiness in medical professionals is most notably associated with errors and accidents (e.g., Chen, Vorona, Chiu, & Ware, 2008; Gold, et al., 1992; Suzuki, Ohida, Kaneita, Yokoyama, & Uchiyama, 2005): errors in surgery and in drug administration, and occupational accidents and injuries including incorrect operation of medical equipment, and needlestick injuries. Daytime sleepiness is also associated with poor patient exam performance by medical students (Rodrigues, Viegas, Abreu e Silva, & Tavares, 2002), likelihood of falling asleep on the job (Gold, et al., 1992), and impaired motor performance (Saxena & George, 2005).

The causes of daytime sleepiness in the medical context are often cited as high workloads, work hours, and stress, along with shift work and on-call duties (Landrigan, et al., 2004; Gold, et al., 1992). People working over 80 hours per week will not have the opportunity to sleep as long. Similarly, individuals constantly switching between daytime and nighttime shifts would experience their sleep as disturbed.

Thus, I predict that:

**Hypothesis 3:** Daytime sleepiness will increase over the first eight months of clinical experience.

**Predicting Psychological Change**

In the previous sections, I predicted overall trends of empathy, self-efficacy, and daytime sleepiness throughout medical training. In this sense, time is the predictor variable. However, I
am more curious about why each of these variables change, and why some individuals may see positive or negative, or larger or smaller changes than others. I will examine possible reasons for these changes, including (1) leadership behaviours of supervisors (including transformational leadership, laissez-faire leadership, and abusive supervision), and (2) experiences that are relevant to the medical context (including morbidity and mortality experiences, beneficiary contact, and experienced boredom). Individuals will have different experiences with each of these variables that could impact changes in empathy, self-efficacy and daytime sleepiness, potentially amplifying a change, weakening it, or even changing its direction. Each will be discussed in turn, along with hypothesized effects on empathy, self-efficacy, and daytime sleepiness.

**Leadership**

In this study, I am interested in transformational leadership, laissez-faire leadership, and abusive supervision, and their effects on empathy, self-efficacy, and daytime sleepiness. Transformational leadership is one of the most prominent positive forms of leadership behaviours studied. Initially conceived by Burns (1978) and later refined by Bass (1985, 1990), transformational leadership is comprised of four behaviours: inspirational motivation, idealized influence, individualized consideration, and intellectual stimulation. Inspirational motivation involves motivating employees to achieve more than they thought possible, setting challenging (but realistic) goals and standards, and demonstrating belief and confidence in employees. Idealized influence reflects a leader’s focus on what is best for the employees and the organization, and includes having a long-term vision, creating a collective sense of mission, and showing integrity, and involves providing a role model consistent with the leader’s vision. Within individualized consideration, leaders give special attention to the individual needs of
employees, both for success and development. The leader thus acts as mentor, showing care, compassion, and empathy, and forming a trusting relationship (Podsakoff, MacKenzie, Moorman & Fetter, 1990). Finally, intellectual stimulation involves encouraging employees to think for themselves, and to challenge their own assumptions, and ask for suggestions, input, and feedback.

Laissez-faire – literally “leave to do” in French – leadership is a style of leadership behaviour that is just that: leave employees to do their work on their own. In laissez-faire leadership, the leader is passive, avoidant and unwilling to make decisions or take responsibility, and does not use the formal authority given to them by the organization (Barling, Christie, & Hoption, 2010; Hinkin & Schriesheim, 2008). Furthermore, Hinkin and Schriesheim (2008) note that laissez-faire leadership’s lack of response to both good and poor performance (reward omission and punishment omission, respectively) is a particularly important component of laissez-faire.

Finally, abusive supervision is defined by Tepper (2000) as the “… sustained display of hostile verbal and nonverbal behaviors, excluding physical contact” (p. 178). Abusive supervision reflects the perception of leaders’ behaviours, which are thought to be willful and intentional (Tepper, 2007). Abusive supervision is sustained over time, as opposed to supervisors who may be having a particularly bad day and take it out on employees once in a while (Tepper, 2007).

In traditional organizations, these three leadership behaviours have been linked to general well-being (e.g., Kuoppala, Lamminpaa, Liira, & Vaino, 2008), perceived meaningfulness of work (e.g., Arnold, Turner, Barling, Kelloway, & McKee, 2007), psychological empowerment
(e.g., Avolio, Zhu, Koh, & Bhatia, 2004), psychological distress, and sleep loss (e.g., Rafferty, Restubog & Jimmieson, 2010), to name a few.

While the majority of research in extreme contexts has focused on the physical effects of leadership (e.g. Bass, Avolio, Jung, & Berson, 2003; Beaton, Johnson, Infield, Ollis, & Bond, 2001; Dvir, Eden, Avolio, & Shamir, 2002), a few studies have also examined the psychological effects of leadership behaviours in the extreme context of the military. On the positive side, Jones et al. (2012) showed that perceived positive leadership was associated with lower levels of mental disorders and PTSD outcomes in a sample of UK Armed Forces personnel serving in Afghanistan. Dvir et al. (2002) demonstrated a positive relationship between transformational leadership of platoon leaders during basic training and follower development, including self-efficacy, engagement and morality. In a theoretical article, Bartone (2006) proposes that transformational leadership could also play a role in the development of resilience and hardiness of military members in the face of stress. Beyond (and arguably even in) the military, there is little empirical research examining the influence of leadership on psychological outcomes in extreme contexts.

Given the positive effects that leadership has shown in other contexts and for other psychological outcomes, I detail in the discussion below how leadership might play a role in the development of the outcomes of interest in this study. As discussed earlier, this sample represents newcomers to the fields of medicine and nursing. As such, students are expected to be especially susceptible to environmental cues (e.g., Morrison, 1993), and organizational influences (e.g., Katz, 1990). The behaviours of supervisors are a source of environmental cues through their interactions with newcomers, which are expected to have an especially important impact on newcomer socialization.
Empathy. I predict that the leadership behaviours of supervisors will have a significant impact on changes in empathy of medical and nursing students. While the relationship between leadership behaviours and follower empathy has not been widely studied, I suggest that leadership will influence follower empathy.

I believe that the behaviours inherent in transformational leadership could improve the empathy of followers. First, in emphasizing that leaders choose to do the right thing, idealized influence may refocus students on the wider impact of their behaviours on patients, and encourage them take a broader, more idealistic approach to treatment, with a long-term outlook. Shamir, House, and Arthur (1993) argued that value congruence between leader and follower is one of the mechanisms through which transformational leadership exerts influence. This long-term, value-driven view may help counteract any decreases in empathy thought to arise from short-term feelings of not having an impact. Through the same mechanisms, the use of idealized influence may also help students maintain initial career ideals, abandonment of which is a potential cause of empathy decline (Neumann et al., 2011). Second, through intellectual stimulation, transformational leadership teaches followers to challenge their assumptions and think about problems in different ways. Doing so may result in enhanced perspective-taking, as students think about patients’ problems in different ways. Finally, transformational leaders are more likely to be empathetic themselves (e.g., Butler & Chinowsky, 2006), which is a core aspect of individualized consideration. Displays of individualized consideration by leaders toward students and patients are environmental signals that newcomers may interpret as expected and necessary behaviours for their own success (e.g., Ashford & Black, 1996). Thus simply having a role model who behaves empathetically, students may themselves see a boost in their own empathy (Brazeau, Schroeder, Rovi, & Boyd, 2010).
Thus I predict:

**Hypothesis 4a:** Transformational leadership behaviours will predict positive changes in empathy across the first eight months of clinical experience.

I predict that laissez-faire leadership will have a detrimental impact on the development of empathy for two reasons. First, laissez-faire leadership conveys the message of a culture that does not value caring. As theories of newcomer socialization suggest, newcomers actively seek information about their new environment, and are expected to identify and internalize such information as a norm of behaviour (e.g., Morrison, 1993). When supervisors neglect the development of medical trainees, students may mimic the absent behaviors of their supervisors (Brazeau et al., 2010), reducing their own empathetic behaviours. Second, laissez-faire leadership is related to destructive employee behaviours such as conflict with co-workers, distress, and workplace bullying (Skogstad, Einarsen, Torsheim, Schanke Aasland, & Hetland, 2007). These relationships suggest that laissez-faire leadership creates a climate in which negative behaviours are acceptable since there are no consequences, and so students expend less effort on behaving in an empathetic manner. Thus I predict that:

**Hypothesis 4b:** Laissez-faire leadership will predict negative changes in empathy across the first eight months of clinical experience.

Characterized by the demeaning and insulting of followers, abusive supervision is antithetical to empathy, and coping with an abusive supervisor would be a substantial drain on followers’ psychological resources (Rafferty, et al., 2010). Those subjected to abusive supervision are likely to have little energy to behave empathetically, and are also likely to take cues from the supervisor that they are in a culture that does not value caring, further impeding the development or maintenance of empathy (Larson & Yao, 2005). In support of this, Neumann et al. (2011) suggested that medical student abuse by supervisors and mentors was a key
antecedent of medical students’ distress and empathy decline. So while researchers have not examined the link between abusive supervision and empathy explicitly, I suggest that abusive supervision will negatively affect the development of empathy.

Thus I predict that:

**Hypothesis 4c:** Abusive supervision will predict negative changes in empathy across the first eight months of clinical experience.

**Self-efficacy.** The relationships between leadership behaviours and follower self-efficacy have been widely studied in traditional contexts (e.g., Chen & Bliese, 2002; Tierney & Farmer, 2002). I expect that these relationships may develop in the medical context, through any of the four antecedents outline by Bandura: Performance accomplishments, vicarious experiences, verbal persuasion, and physiological and emotional states (Gong, Huang, & Farh, 2009). Leadership is related indirectly to performance, and a common mediator of that relationship is self- and group-efficacy (Walumbwa, Avolio, & Zhu, 2008). Leaders influence the vicarious experiences of followers, and also use verbal persuasion with followers. Finally, leaders can impact the physiological and emotional states of followers through the work environments that they create (for example, by creating a trusting environment and optimizing job characteristics; Chen & Bliese, 2002).

Transformational leaders are expected to foster followers’ self-efficacy by improving their affective states, providing necessary resources, and developing a trusting relationship. First, transformational leadership changes the way followers view their own capabilities by improving their affective state (Tsui, Chen, & Cheng, 2009), and ensuring their needs are met through individualized consideration. A positive affective state is a pre-condition for high self-efficacy (Bandura, 1977). Bandura (1977) suggested that individuals experiencing positive psychological states have higher self-efficacy because they are more confident in their
capabilities. Second, transformational leadership provides necessary resources, and sets high expectations and challenging goals, which in turn increase self-efficacy (Durham, Knight, & Locke, 1997; Dvir et al., 2002; Shea & Howell, 1999). Finally, transformational leadership fosters high levels of trust, and sets a positive example for followers, both of which make a transformational leader a credible source for persuasion, encouraging and empowering followers to persist at their tasks (Dvir et al., 2002). Not surprisingly, transformational leadership is associated with follower self-efficacy in traditional organizations (e.g., Den Hartog & Belschak, 2012; Dvir et al., 2002; Walumbwa et al., 2008).

I expect that the relationships between transformational leadership and follower self-efficacy will translate to extreme contexts, as in addition to mechanisms described above, transformational leadership will help reduce ambiguity and foster confidence in the face of more extreme consequences. While research on the influence of leadership on self-efficacy in more extreme contexts is sparse, evidence exists for the impact of transformational leadership on self-efficacy in firefighting (Pillai & Williams, 2004), and the military (Dvir et al., 2002). Furthermore, leadership intervention studies have shown that transformational leadership training leads to increases in firefighters’ perceptions of their abilities to achieve future success (Beaton et al., 2001), as well as self-efficacy in the military (Dvir et al., 2002). It should be noted that in both of these cases the research was done in instances where extreme events are perhaps less frequent than those faced by medical trainees (i.e. firefighting and military training, versus routine medical care). Given support for the role of transformational leadership in self-efficacy development, I predict:

**Hypothesis 5a:** Transformational leadership will predict positive changes in self-efficacy across the first eight months of clinical experience.
In contrast to transformational leadership, I expect that laissez-faire leadership, characterized by a lack of positive or negative feedback, will negatively influence the development of students’ self-efficacy through the increased ambiguity and stress students would encounter.

First, consistent feedback, whether positive or negative, is necessary for employees to understand how their performance is developing. This feedback is also necessary to give employees goals and actions to improve their future performance. Beyond feedback, other forms of verbal persuasion (e.g., encouragement) enhance self-efficacy beliefs. Laissez-faire leadership provides little guidance, resources, and clarity (Skogstad et al., 2007), impeding the ability of employees to successfully do their jobs, which forms a primary basis for their self-efficacy beliefs.

Second, laissez-faire leadership is also a significant source of stress for employees (e.g., Skogstad et al., 2007). As Bandura (1977) described in his original theory, negative emotional arousal could cause individuals to question their competence; the negative emotions resulting from laissez-faire leadership could predict lower self-efficacy. Chen and Bliese (2002) showed that “psychological strain” was negatively related to self-efficacy for employees in the lower levels of the organization, which suggests that the stress created by laissez-faire leadership could result in lower self-efficacy for trainees. Thus:

**Hypothesis 5b**: Laissez-faire leadership will predict negative changes in self-efficacy across the first eight months of clinical experience.

While behaviours that reflect abusive supervision are much different than those of laissez-faire leadership, similar mechanisms (i.e. performance feedback, stress) are expected to influence self-efficacy.
First, while an abusive supervisor might be expected to give (negative) feedback, it is unlikely to be constructive or provide employees with information that would improve future performance. The more sustained this negative feedback, the more employees will be left questioning their capability.

Second, abusive supervision is related to employee stress (e.g., Tepper, 2007), and other negative forms of psychological arousal, such as anxiety, emotional exhaustion, and depression (Tepper, 2000), which, as described above, could leave individuals questioning their performance abilities. The relationship between abusive supervision and negative psychological states also extends to the medical context. In an examination of the causes of medical student distress, Neumann et al. (2011) rank mistreatment by supervisors and mentors among the stronger predictors of distress in their research.

Thus, I predict:

**Hypothesis 5c:** Abusive supervision will predict negative changes in self-efficacy across the first eight months of clinical experience.

**Daytime sleepiness.** I posit that leadership will influence trainees’ subjective daytime sleepiness by (a) affecting stimulation and engagement during the workday, and (b) influencing the quality of sleep after the workday.

First, transformational leaders can directly decrease daytime sleepiness by providing intellectual stimulation to trainees, thereby increasing engagement with work tasks (Zhu, Avolio, & Walumbwa, 2009). This is important, as cognitive engagement reduces feelings of daytime sleepiness (Boothby, 2011). Inspirational motivation may further reduce daytime sleepiness by increasing motivation (Dvir et al., 2002) toward valued goals. Paradoxically, decreased motivation is correlated with greater daytime sleepiness, due to lack of energy (Ahsberg,
Kecklund, Akerstedt, & Gamberale, 2000), suggesting the importance of motivation in preventing sleepiness.

Second, transformational leadership affects daytime sleepiness through its impact on followers’ sleep quality. Nighttime sleep quality is consistently associated with daytime sleepiness, and by improving followers’ sleep quality, transformational leadership may indirectly decrease feelings of sleepiness. Transformational leadership is expected to improve sleep quality by decreasing the stress experienced by students through its use of individualized consideration and corresponding supportive behaviours. The relationship between transformational leadership and sleep quality itself has been established in a sample of healthcare workers (Munir & Nielsen, 2009), and an exploratory study of firefighters in which supportive leadership behaviours buffered the impact of traumatic events on exhaustion, nightmares, and other sleep problems (Fullerton, McCarroll, Ursano, & Wright, 1992). Finally, a study of positive leadership behaviours (e.g., clarifying behaviours, supporting behaviours) was positively related to sleep quality in a sample of deployed military personnel (Gunia, Sipos, LoPresti, & Adler, 2015). I expect that transformational leadership will reduce feelings of daytime sleepiness in medical trainees.

**Hypothesis 6a:** Transformational leadership will predict negative changes in daytime sleepiness across the first eight months of clinical experience.

Conversely, I expect that laissez-faire leadership negatively affects daytime sleepiness, both directly from lack of engagement during the day, and from stress’s impact on nighttime sleep quality. First, laissez-faire leadership does not provide the needed stimulation to maintain or increase engagement and avoid daytime sleepiness. Thus I would expect that followers of laissez-faire leadership experience higher feelings of sleepiness at work. While the relationship between laissez-faire leadership and daytime sleepiness has not been investigated, researchers
have discussed the impact of laissez-faire leadership on engagement, suggesting that laissez-faire leadership lacks motivational power (Mester, Visser, Roodt, & Kellerman, 2003; Tims, Bakker, & Xanthopoulou, 2011). As noted earlier, daytime sleepiness is associated with decreased motivation (Ahsberg et al., 2000), so it is expected that laissez-faire will be associated with increased daytime sleepiness.

Second, laissez-faire leadership may also affect daytime sleepiness indirectly through its damage to sleep quality. As discussed, the ambiguity and absence inherent within laissez-faire leadership may create stress in students. Laissez-faire leadership has been shown to be a source of stress, and is also related to similar stress responses as decreased well-being, and interpersonal conflict, and bullying (e.g., Skogstad et al., 2007). Stress, psychological distress, and bullying predict poor nighttime sleep (e.g., Jansson & Linton, 2007; Niedhammer, David, Degioanni, Drummond, & Philip, 2009), suggesting the potential effects leadership has on sleep quality. Thus, laissez-faire leadership is expected to increase feelings of sleepiness in medical trainees.

**Hypothesis 6b**: Laissez-faire leadership will predict positive changes in daytime sleepiness across the first eight months of clinical experiences.

Finally, I would also expect abusive supervision to increase feelings of daytime sleepiness, through its effect on stress and anxiety during the day, and the effects of the stress it creates on nighttime sleep quality. First, the sustained experience of abusive supervision creates stress and anxiety on a daily basis (e.g., Martinko, Harvey, Brees, & Mackey, 2013). Such sustained experiences are expected to deplete student emotional resources, thereby increasing feelings of daytime sleepiness. Abusive supervision is associated with negative psychological states such as stress and anxiety (Tepper, 2007), which are both positively associated with sleepiness (Ohayon, Caulet, Philip, Guilleminault, & Priest, 1997; Stepanski, Markey, Zorick, &
Roth, 1990; Tepper, 2007), suggesting the potential effect of abusive supervision on daytime sleepiness.

Abusive supervision may also influence daytime sleepiness through negative effects on sleep quality, as stress is also a factor in insomnia and sleep problems (Linton, 2004). The sustained stress and anxiety experienced by students is likely to follow them home from the workplace as well. Indeed, there is a relationship between abusive supervision and insomnia (Rafferty, et al., 2010). Similarly, workplace bullying (similar to abusive supervision, but from any source, not just a leader) also negatively impacts sleep (Niedhammer et al., 2009).

Thus, I predict that:

**Hypothesis 6c: Abusive supervision will predict positive changes in daytime sleepiness across the first eight months of clinical experience.**

**Context-Relevant Experiences**

As noted earlier, an extreme context is one in which “risks of severe physical, psychological or material consequences (e.g., physical harm, devastation or destruction) to organizational members or their constituents exist” (Hannah et al., 2009, p. 897). People working in extreme contexts have unique experiences, and little is known of their psychological effects. However, as theorized by researchers examining changes in empathy in medical trainees, empathy could decline due to the unique contextual experiences faced by students, suggesting that context plays a role (e.g., Bellini & Shea, 2005; Neumann et al., 2011). Johns (2006) highlighted our need as researchers to incorporate a consideration of context into our research, as context-specific experiences can help us better understand the occurrence and meaning of relationships studied in organizational behaviour, and improve the applicability of our findings. Thus, I incorporate three specific experiences that medical trainees are expected to encounter, namely patient morbidity and mortality, beneficiary contact, and boredom. Of course
beneficiary contact and boredom are not experiences that are restricted to this particular context, but they do represent experiences that may occur within this context, and that could exert psychological consequences.

During the third year of medical school, students gain their first sustained clinical experience. Nursing students get similar exposure in the fourth year of their program. As students experience the death or on-going illness of a patient, see and question their true impact on patients, and discover their true scope of control over injury and disease, empathy, self-efficacy and daytime sleepiness might be affected. Newcomers might be especially susceptible, as these experiences may shape and influence their future cognitions and behaviours (e.g., Ashforth & Saks, 1996; Jablin, 2001). Below, I introduce the three contextual experiences that medical trainees are expected to encounter (i.e. morbidity and mortality of patients, beneficiary contact, and boredom). I then postulate how each may influence the outcome variables (i.e. empathy, self-efficacy, daytime sleepiness).

Morbidity and mortality of patients is an unavoidable experience in the medical realm, given that not everyone can be healed. Morbidity refers to a state of being diseased or unhealthy; mortality refers to death. Fullerton et al. (1992) emphasized the importance of understanding how disasters and loss of life affect rescue workers, and by capturing these experiences I hope to understand their effects on student development. I would expect that the different ways in which students experience morbidity and mortality negatively impact their empathy, self-efficacy and daytime sleepiness, and discuss each in detail below.

There are many positive aspects to the medical trainee experience. Physician and nurse trainees have the opportunity to make a positive impact on the lives of their patients. Though jobs involving frequent “customer” interactions have typically been considered depleting and a
source of burnout (e.g., Conrad & Keller-Guenther, 2006), client interactions may also be restorative and produce positive affect, self-affirmation, and a perceived prosocial impact (Lilius, 2012).

Evidence of the uplifting effect of positive interactions comes from the research of Grant on the effects of interacting with beneficiaries of one’s work (e.g., Grant, 2008a; Grant & Parker, 2009). His work on beneficiary contact has demonstrated that incorporating and promoting contact with beneficiaries, which vividly draws attention to the positive impact of one’s work, increases the desire to protect and promote others’ well-being (Grant, 2008a; Grant & Parker, 2009). Employees who form relationships with beneficiaries are better able to identify with beneficiaries and develop an affective commitment toward them (Grant, 2007). This commitment heightens prosocial motivations, resulting in greater effort and persistence, and higher levels of performance (Grant, 2008b). Grant (2008a) conceptualizes beneficiary contact along three factors: frequency of interaction, breadth of interaction with different beneficiaries, and depth of relationships formed². With a narrower focus specifically on contact with patients, I expect positive effects of beneficiary contact on empathy, self-efficacy, and daytime sleepiness.

Finally, boredom is “a transient affective state in which the individual feels a pervasive lack of interest in the current activity” (Fisher, 1991, p. 3). While the notion of boredom in extreme contexts may seem counter-intuitive, research suggests that not only does boredom occur, but medical and nursing students may be more prone to boredom, as they are likely individuals with higher capabilities (Drory, 1982) and excitement-seeking tendencies (Best & Kilpatrick, 1977). Boredom does have detrimental performance effects (e.g., Drory, 1982), and I

² My research examines each of these three components separately, but as all hypothesized relationships are in the same direction, global hypotheses are made throughout.
expect that it will also negatively affect empathy and self-efficacy during students’ initial clinical experiences, as I suggest below.

Empathy.

Morbidity and mortality. I predict that increased exposure to patient morbidity and mortality will decrease levels of empathy (a) as trainees attempt to psychologically detach themselves to cope, and (b) realize that they often cannot help patients as they expected they would.

First, psychological detachment from work has been posited as one reason for empathy decline in medical trainees (Hojat, et al., 2009), and might be a necessary condition for the decline. Some authors have suggested that detachment is a way of successfully dealing with patient mortality, and not being able to help as many patients as expected (Bellini et al., 2002; Hojat et al., 2002), both of which serve as major sources of distress for medical trainees (Neumann, et al., 2011). This assumption has some empirical support, as healthcare workers who interact with traumatized individuals suffer stress, burnout, and compassion fatigue (Collins & Long, 2003). If a strategy of detachment can prevent this, not empathizing with patients may be seen to be a rational way of dealing with stress. Thus, the more patient morbidity and mortality trainees experience, the more they may disengage, and the more their empathy would decline.

Second, feelings of helplessness may predict declines in empathy. Acts of empathy require sufficient self-control (Tangney, Baumeister, & Boone, 2004) to take on the perspectives of others and respond emotionally. Yet self-control is a limited resource (Muraven, Tice, & Baumeister, 1998), and when trainees feel that their efforts are ineffective, they may be less likely to invest that energy, and thus empathy would decline. Furthermore, both feelings of
helplessness and feelings that trainees are not making a difference are related to feelings of
cynicism (Larson & Yao, 2005; Thomas, et al., 2007).

Thus, I predict that:

**Hypothesis 7a:** Morbidity and mortality will predict negative changes in empathy across the first eight months of clinical experience.

**Beneficiary contact.** In contrast to the experiences of morbidity and mortality, contact with patients can allow medical and nursing students to see for themselves the positive impact they have had on their lives, reversing or stopping any decline of empathy, potentially even increasing empathy toward the patient.

In a theoretical article, Grant (2007) argued that contact with beneficiaries (e.g., patients) allows employees to empathize, identify with, and take the perspective of beneficiaries. The relationship between beneficiary contact and empathy has been empirically supported, as beneficiary contact is related to increased perspective-taking (Parker & Axtell, 2001). Importantly, this research was conducted in traditional contexts. The positive effects may be greater when patient health is concerned, because contact with patients highlights the positive impacts of one’s work on patients’ quality of life, not just their day-to-day routines. Given this, I expect:

**Hypothesis 7b:** Beneficiary contact (frequency, breadth, and depth) will predict increases in empathy across the first eight months of clinical experience.

**Boredom.** Finally, I expect that boredom will reduce empathy of medical trainees, through the negative effects that boredom has on affective states and stimulation.

First, displaying empathy requires attention, engagement, and patience. Yet boredom is characterized by disinterest (e.g., Kane et al., 2007), restlessness, (e.g., O’Hanlon, 1981), inattention and disengagement, and withdrawal from tasks (Spector et al., 2006). Thus, I expect
that bored students will be unlikely to invest the time required to consider another’s perspective. In the case of student physicians and nurses, boredom could result in disengagement from patients, thereby reducing empathy even further.

Second, boredom could indicate that students are not being challenged, or having the impact that they expected. Drory (1982) suggested that boredom is often caused by a mismatch between expected and experienced stimulation, and that highly skilled individuals are more susceptible to a negative mismatch. A commonly cited reason for declining empathy in medical trainees is that the practice of medicine is not what they had expected it to be (Neumann, et al., 2011). If life as a medical trainee turns out to be less stimulating than expected, and is instead experienced as boring, this could prompt disillusionment and decrease in empathy. Thus, I predict that:

**Hypothesis 7c:** Boredom will predict negative changes in empathy across the first eight months of clinical experience.

**Self-efficacy.**

**Morbidity and mortality.** Self-efficacy develops over time through performance success, vicarious experiences, verbal persuasion, and psychological states (Bandura, 1982). Bandura (1982) notes that performance accomplishments are the strongest predictor of self-efficacy. I predict that success in an extreme context, which is full of uncertainty, challenge, and risk, will positively influence self-efficacy. But being unable to cure, ease the pain, or prevent the suffering or death of a patient could be considered a performance failure and decrease a student’s self-efficacy. Exacerbating this, research suggests there is a tendency toward physicians’ self-blame for patient death, even though it may not be warranted (Collins, Block, Arnold, & Christakis, 2009). The more frequent the morbidity and mortality experiences, the more self-efficacy is likely to decrease. Thus,
**Hypothesis 8a**: Morbidity and mortality experiences will predict negative changes in self-efficacy across the first eight months of clinical experience.

**Beneficiary contact.** If frequent experiences with morbidity and mortality can be interpreted as feedback of performance failure, then contact with beneficiaries highlighting the positive effects of one’s work could be credible feedback of performance success. Receiving positive performance feedback about the effects of one’s treatment of patients, either from patients themselves or their families, would influence self-efficacy positively (Bandura, 1982). While self-efficacy has not been researched with respect to beneficiary contact, beneficiary contact does predict heightened feelings of value and competence (Penner, Dovidio, Piliavin & Shroeder, 2005) which are closely related to self-efficacy, suggesting the role beneficiary contact has on self-efficacy development.

Thus, I predict that:

**Hypothesis 8b**: Beneficiary contact (frequency, breadth, and depth) will predict positive changes in self-efficacy across the first eight months of clinical experience.

**Boredom.** It would be difficult to interpret boredom as a cue of past performance, but feelings of boredom may influence students’ assessment of their future abilities. Specifically, boredom is positively related to stress (Melamed, et al., 1995), which, as Bandura (1982) notes, is a negative psychological state expected to reduce self-efficacy. The altered emotional states that boredom creates may negatively impact the development of self-efficacy.

Thus, I predict:

**Hypothesis 8c**: Boredom will negatively impact changes in self-efficacy across the first eight months of clinical experience.

**Daytime sleepiness.**

**Morbidity and mortality.** Many medical researchers assume that medical students’ and residents’ sleepiness is simply attributable to long hours and shift work (e.g., Landrigan et al.,
This assumption has led to various policies and practices aimed at addressing student hours and scheduling (e.g., Friedman, Karani, & Fallar, 2011). However, despite these newer policies, reports suggest that medical trainee sleep is still a problem (e.g., Oakley et al., 2014). I suggest one potential reason for this is patient experiences: Difficult and negative patient interactions could also influence students’ daytime sleepiness. First, as noted earlier, a common source of daytime sleepiness is stress. Negative patient experiences are a common source of stress, and in turn could influence daytime sleepiness. One of the commonly cited sources of stress for medical caregivers is exposure to death and suffering of patients, and coinciding feelings of helplessness (Hojat, et al., 2009), which could increase feelings of daytime sleepiness.

Second, negative experiences could also affect daytime sleepiness indirectly, by compromising nighttime sleep quality (e.g., insomnia and nightmares). Encountering frequent patient morbidity and mortality could lead to rumination which impedes the onset of sleep (e.g., Zoccola, Dickerson, & Lam, 2009), and to stress and nightmares which impair quality of sleep (e.g., Neylan, et al., 2002). Traumatic events, such as witnessing significant injury or death, predict sleep disturbances (Lavie, 2001). This relationship also exists in extreme contexts, for example, as trauma predicts increased insomnia and nightmares in firefighters (Fullerton et al., 1992; Raphael, 1986), and decreased sleep quality in police officers (Neylan, et al., 2002).

Thus, I predict:

*Hypothesis 9a:* Morbidity and mortality experiences will predict negative changes in daytime sleepiness across the first eight months of clinical experience.

**Beneficiary contact.** On the other hand, I expect that beneficiary contact will decrease daytime sleepiness. Beneficiary contact could buffer many negative aspects of one’s job, reducing feelings of stress and burnout (Grant & Campbell, 2007), and by extension, daytime
sleepiness. Beneficiary contact may also lead to better sleep quality at night, and thus less daytime sleepiness at work.

First, beneficiary contact could improve daytime sleepiness by enhancing the engagement and motivation of medical trainees. Being engaged with work (and those who benefit from that work), eliciting attention, absorption, and motivation (Rothbard, 2001), is expected to prevent feelings of daytime sleepiness. One of the major predictors of engagement is meaningfulness of work (May, Gilson & Harter, 2004), which suggests that any interaction highlighting the meaning of one’s work is expected to increase engagement. Indeed, research shows that beneficiary contact increases engagement and motivation, by showing individuals that their work is valued, and that they are making a difference (Grant et al., 2007).

Beneficiary contact could also improve daytime sleepiness by improving well-being, which in turn improves nighttime sleep quality. Meaningfulness and feelings of “doing good” are expected to decrease such negative psychological states as stress, anxiety and depression. While the majority of research on the impact of beneficiary interactions is limited to performance and other organizational outcomes, some studies have investigated the influence of beneficiary contact on health and well-being. Feeling that one is benefitting others is related to lower levels of depression (Lyubomirsky, Sheldon & Schkade, 2005), and improved health and longevity (Brown, Nesse, Vinokur & Smith, 2003). Sleep quality is negatively related to depression (e.g., Tsuno, Besset & Ritchie, 2005), and positively related to health (e.g., Minowa & Tango, 2003).

Thus:

**Hypothesis 9b:** Beneficiary contact (frequency, breadth, and depth) will positively impact changes in daytime sleepiness over the first eight months of clinical experience.

**Boredom.** The effects of boredom on daytime sleepiness could potentially be bidirectional. On one hand, boredom could increase feelings of daytime sleepiness, as boredom is
characterized by a lack of motivation and meaning (Spector et al., 2006), which could be experienced as sleepiness. Boredom is also associated with stress (e.g. Melamed, et al., 1995), which could negatively affect nighttime sleep, and by extension increase daytime sleepiness. But boredom is also associated with restlessness (O’Hanlon, 1981; Robinson, 1975), which suggests an excess of energy, opposed to the lack of energy that daytime sleepiness suggests. Because of the uncertain effects of boredom on daytime sleepiness, I offer no hypothesis.

2.2 Method

Pilot study

A sample of full-time nurses and medical residents completed a pilot version of the survey to test the reliability of scales on a medical sample, and the reliability of a newly-developed scale for experienced patient morbidity and mortality for the purposes of this study. A central concern from medical and nursing school personnel was the length and time required for the survey. As such, some scales were first shortened, a few items were removed, and some adjusted to be relevant to the general hospital environment. Participants were recruited via e-mail, and completed an online version of the questionnaire. A total of 28 nurses and residents completed the pilot questionnaire (38% male, 65% nurses, M age = 27.4 years).

The scale to measure morbidity and mortality was developed using commonly-cited experiences of medical students, with input from three former medical students. A total of 23 items were developed initially for the pilot study, asking participants to rate the frequency with which they experienced events such as witnessing a patient die, reducing a patient’s pain (reverse-scored), and witnessing a patient in extreme pain. Frequency was rated on a scale of 1 = rarely, 2 = a few times/month, 3 = about once/week, 4 = a few times/week, 5 = daily. Based on the pilot data, the scale was reduced to 12 items to achieve a higher internal consistency
(Cronbach’s alpha = 0.91). Scales for *empathy*, *self-efficacy*, and *transformational leadership* were also shortened following the pilot by removing items with lower inter-item correlation, in order to shorten the total survey length.

A summary of the pilot scale reliabilities, means, and standard deviations is in Table 2-1 and the final questionnaire in Appendix A.

**Full Sample**

**Sample.** My sample consisted of third-year medical students and fourth-year nursing students at a single university in Ontario, Canada. 45% of the sample participants were nursing students. The sample was 81% female; of the nursing students 96% were female, and of the medical students 62% were females. Participants were an average age of 23.39 years old ($SD = 3.00$) at the start of the study; nursing students were an average age of 22.52 ($SD = 3.62$), and medical students were an average age of 24.42 ($SD = 2.97$).

**Recruitment.** Prior to beginning the first sustained clinical component of their program, all third-year medical students and fourth-year nursing students were invited by their respective program offices via e-mail to participate in this research, and complete an electronic survey to obtain baseline scores for all study and variables. Approximately 170 students were initially contacted through the programs. 39 students completed the first round, and were re-contacted by me by e-mail every six weeks (to approximately match medical rotation length) for a total of five follow-ups (approximately eight months total). Follow-up surveys collected all study variables to examine the hypotheses of intra-individual change for the three outcome variables. Participants could win gift cards in each of the cycles.

---

3 While students were enrolled at a single university, the practical placements and rotations did not all take place at the university’s hospital, and could be completed at any healthcare facility in the province of Ontario.
The number of participants in each cycle is as follows: 39 at T1, 34 at T2, 28 at T3, 36 at T4, and 29 at T5. This represents response rates of the entire classes of 23%, 20%, 16%, 21%, and 17%, respectively. An advantage of the multilevel modeling approach used in this research is that it accommodates for an unequal number of responses between individuals (Singer & Willett, 2003), thus even if a participant did not complete all cycles of data collection, their responses were retained for analysis. Participants must have completed at least two cycles (i.e., the baseline and at least one of the five follow-ups), and on average completed 3.69 of the follow-ups ($SD = 1.28$).

**Measures**

Means, standard deviations and reliabilities for all scales used appear in Table 2-2.

**Outcome variables.** *Empathy* was measured using the perspective-taking and empathic concern subscales of Davis’s (1980) Interpersonal Reactivity Index. The 14 items are rated on a scale of 0 (does not describe me well) to 4 (describes me very well). Example items include “I try to look at everybody’s side of a disagreement before I make the decision” for perspective-taking, and “I often have tender, concerned feelings for people less fortunate than me” for empathic concern. While the subscales were initially intended to measure two separate constructs (i.e. cognitive and affective empathy), they are usually combined to form a single measure of empathy (e.g., Joireman, et al., 2006), which was the approach taken in this study.

*Self-efficacy* was measured using a 14-item shortened version of Artino et al.'s (2012)’s 19-item scale, which assesses medical-student specific self-efficacy regarding patient care efficacy, interpersonal skills efficacy, and the practice of evidence-based medicine. I deleted items from the original scale if they were not deemed applicable for the current sample (e.g., one question asked specifically about military environments; other questions asked about health-care
cost management, which while relevant in the Canadian healthcare environment, were excluded in the interest of survey length). Participants rated their level of confidence in completing relevant tasks, such as “perform a thorough exam” and “apply knowledge of normal function to each of the major organ systems” on a 5-point Likert scale (1 = not at all confident, 5 = extremely confident).

*Daytime sleepiness* was measured using the 8-item Epworth Daytime Sleepiness scale (Johns, 1991), which asks participants to indicate the likelihood that they would fall asleep during daily activities (even if they have not completed the activities) such as watching TV, as a passenger in a car, or sitting and reading. The items are rated on a 4-point scale, from 0 = would never doze, to 3 = high chance of dozing.

**Predictor variables.** Leadership behaviours were measured using three separate scales.

*Transformational leadership* was measured using a revised version of Beauchamp et al.’s (2010) 16-item Transformational Teaching Questionnaire. The original scale includes four items for each of the components of transformational leadership, and asks participants to rate teachers on a 5-point Likert scale from 1 = strongly disagree to 5 = strongly agree. The scale was modified to refer to the “leader with whom you interacted most” instead of “teacher”, and shortened to eight items following the pilot; two items from each of the components of transformational leadership were retained while maintaining a high level of internal consistency. Example items include “My leader acts as a person I look up to” and “My leader motivates me to try my hardest”.

*Laissez-faire leadership* was measured using Hinkin and Schriesheim’s (2008) 8-item scale, which asks participants to rate their leader’s behaviours on a 7-point Likert scale (1 =
strongly disagree to 7 = strongly agree). Examples of items include “I often perform well and still receive no praise from my leader” and “When I perform poorly my leader does nothing”.

*Abusive supervision* was measured using a 10-item shortened version of Tepper’s (2000) Abusive Supervision scale, identified by Mitchell and Ambrose (2007). The scale asks participants to rate their agreement with items such as “My leader ridicules me” and “My leader invades my privacy” on a 7-point Likert scale (1 = strongly agree, 7 = strongly disagree).

Context-relevant experiences were measured using three separate instruments; one scale for morbidity and mortality, three sub-scales for beneficiary contact, and one scale for boredom.

*Morbidity and mortality* was measured using a 12-item scale developed for this study. The scale asks participants to report on the frequency of various morbidity and mortality experiences, rated on a scale of 1 = rarely, 2 = a few times/month, 3 = about once/week, 4 = a few times/week, 5 = daily. Examples of items include “Had a patient die unexpectedly” and “Were unable to help a patient”.

*Beneficiary contact* was measured using Grant’s (2008a) 9-item scale of beneficiary contact, which measures three characteristics of students’ placements in relation to patient interactions: frequency, breadth, and depth of interactions (on a scale of 1=strongly disagree to 7=strongly agree). Each factor includes three items, such as “My rotation allowed frequent communication with the people who benefit from my work” - frequency, “My rotation enabled me to meet diverse groups of people who benefit from my work” - breadth, and “My rotation enabled me to build close relationships with the people affected by my work” - depth.

*Boredom* was measured using a modified version of Drory’s (1982) 6-item scale, adjusted for the medical context. Students rated the frequency of which they experienced such feelings as monotony, time passing slowly, and nothing happening, during the previous six
weeks, on a scale of 1 = rarely, 2 = a few times/month, 3 = about once/week, 4 = a few times/week, 5 = daily.

**Control variables.** All analyses controlled for gender (0 = male, 1 = female), and specialty (nursing = 0, medical = 1). Individuals’ baseline levels of each outcome variable (i.e. empathy, self-efficacy, daytime sleepiness) were also controlled statistically. All analyses for daytime sleepiness controlled for self-reported average *hours slept* during the recent six weeks.\(^4\)

The complete questionnaire can be found in Appendix A.

### 2.3 Data Analysis

The analyses include a series of linear growth curve models for each of the study outcomes (empathy, self-efficacy, daytime sleepiness) using the multilevel modeling function in SPSS. Growth curve models were considered appropriate for this study as they allow researchers to study between-person differences in within-person change (Singer & Willett, 2003).

All variables were centered prior to analysis so that the intercept became zero, and the coefficients indicate the effects of a deviation from the mean. The “time” variable was rescaled such that the first time period is entered as zero. Binary variables (i.e. gender and specialty) were coded as 0 and 1 to ease interpretation of study results.

The analytic methodology for hypothesis testing followed procedures used by Baldwin and Hoffmann (2002) in their study of changes in adolescent self-esteem. The general goal of the current analysis was first to model the mean trajectories of empathy, self-efficacy and

---

\(^4\) Initially, my intention in all analyses for empathy was to control for agreeableness, given its association with empathy (e.g. Magalhães, Costa & Costa, 2012). Agreeableness was measured using the two agreeableness items from Gosling, Rentfrow, and Swann Jr’s (2003) short measure of the Big 5 personality variables. However due to the low reliability of the scale (see Table 2-2) I excluded agreeableness from further analyses.
daytime sleepiness over time, and then use the time-varying predictor variables (i.e. leadership, context-relevant experiences) to explain intra-individual changes. Like the current study, Baldwin and Hoffmann modeled changes in self-esteem over time, and then sought to test possible predictors of these changes. Prior to analysis, the mean levels of each outcome were plotted per period to visually explore any possible nonlinear time parameters (see Figures 2-1, 2-2, 2-3). In all cases, it appeared visually as if the mean trajectories were nonlinear, as a result of which several nonlinear functions (e.g., squared, cubed, logarithmic) were explored further. While no nonlinear hypotheses were made, the goal of this procedure was to first capture as much variance in the outcome variables as possible before adding the predictor variables. This capturing of variance makes the testing of the predictor hypotheses statistically more difficult, and more accurate.

The intraclass correlation (ICC) was calculated for each outcome. If the ICC is very low (i.e. approaching zero; Nezlek, 2008), multilevel modeling is not necessary. However, in the case of longitudinal data this is quite unlikely, as was the case in these data.

The models were built in three stages. First, the impact of time, including appropriate nonlinear terms, was included in each model to examine how empathy, self-efficacy, and daytime sleepiness changed during the duration of the study at the mean group-level. The intercept and linear time slope were treated as random effects, and allowed to vary across individuals. Note that the coefficient of the random linear time slope indicates whether or not that slope coefficient varies across individuals, and does not account for the nonlinear time coefficients. However, nonlinear terms were not treated as random effects, as growth-curve models can only accommodate a limited number of random effects (Bryk & Raudenbush, 1992). An unstructured covariance matrix was used to estimate the random effects, a diagonal
covariance matrix used for the repeated measures (Level 1), and maximum likelihood used for model estimation. At the first stage, a number of interaction terms were added to each model to investigate whether the growth trajectories differed based on any of the time-invariant variables, such as gender or medical specialty. No significant interactions emerged for empathy or self-efficacy, and so were excluded from the analysis. However, a significant time X gender interaction emerged for daytime sleepiness. Figure 2-4 presents the mean levels of daytime sleepiness by gender. These two interactions were retained for further analyses in order to control for their effects.

Second, the control variables were added to the model to examine their basic effects and assess residual variance prior to the predictor variables being added. All control variables were entered as fixed effects, as a result of which the estimated coefficients are treated and estimated as constant across individuals.

Third, the time-varying predictor variables were added to test the remainder of the hypotheses. For parsimony and ease of interpretation, separate models were estimated for leadership, and context-relevant experiences. Predictor variables were entered as fixed effects, as there were no theoretical or hypothesized reasons to suggest that their coefficients would differ between-individuals.

In all analyses, unstandardized beta coefficients can be interpreted as an upward or downward shift in the intercept for a specific period. Results were inspected for significance of the predictors, as well as changes in the fixed time terms, and residual variance of the random effects to determine if it is the predictors, and not time, which predict changes in the study variables. Finally, a measure of effect size was calculated for each predictor variable. Due to the complexities of multilevel modeling, traditional effect size calculations (e.g., Cohen’s $d$, eta) are
not possible (Peugh, 2010). Researchers have proposed a variety of potential effect sizes, but currently no consensus exists (Peugh, 2010). Proportion reduction in variance (PRV) was selected, as it is a generally accepted measure (Raudenbush & Bryk, 2002; Singer & Willett, 2003), but not comparable in the same sense as Cohen’s d or eta. As a measure of effect size, PRV describes how much an individual predictor changes the model’s residual variance, compared to a model without that predictor (Raudenbush & Bryk, 2002; Singer & Willett, 2003). In this study, the PRV reports the additional variance explained by a predictor variable (e.g., transformational leadership) compared to a model with only the time and control variables.

2.4 Results

Preliminary Analyses

A correlation table is provided in Table 2-3, however these data do not take into account the multilevel nature of the data and should only be used for a general overview of the data.

The intraclass correlations for empathy, self-efficacy, and daytime sleepiness were 0.64, 0.70, and 0.60, respectively, suggesting that multilevel modeling was appropriate. Figures 2-1 through 2-3 illustrate the changes in mean empathy, self-efficacy, and daytime sleepiness, respectively, across study periods. These graphs were used to estimate the types of nonlinear terms with which to model time. For empathy, a squared and a cubic term were added, to capture the pattern that mean empathy initially decreased, then increased, and then decreased again. For self-efficacy, a square-root and logarithmic term were added, capturing the pattern that self-efficacy decreased slightly initially, then increased, and then tapered off. For daytime sleepiness, a logarithmic term was added, to model the pattern of an initial increase in sleepiness that was followed by the gradual decrease in daytime sleepiness.

---

5 The logarithmic term was computed as log(t+1) as log(0) is incalculable.
Attrition analyses were conducted to ensure that any changes in the outcome variables was not a result of changes in the underlying sample characteristics. Tukey’s (1949) multiple comparison test showed no significant differences in terms of gender, specialty, or age of the participants across the time periods. This suggests that any fluctuation in the number of participants was not based on any descriptive characteristics, and that changes in the outcome variables was not a result of changes in the proportion of females, nurses, or the result of different mean ages.

**Growth-Curve Models**

The results of the growth-curve analysis appear in Tables 2-4 to 2-9. In each table, the effects of time on each of the outcomes are introduced in Model 1. Note that for daytime sleepiness, Model 1 is separated into 1a and 1b, to include interactions between gender and time. Model 2 includes the effects of the control variables. Finally, Model 3 adds the predictor variables of interest.

**Time trends.** Hypotheses 1 through 3 predicted overall time trends of each outcome variable prior to any other explanatory variables.

Hypothesis 1 predicted that empathy would decline throughout the study. In Table 2-4 and 2-5, the three time terms are significant under Model 1. The coefficients yielded suggest that empathy decreases initially (linear: $\beta = -0.29, p < .05$), followed by an increase (squared: $\beta = 0.24, p < .01$) and then a final decrease (cubed: $\beta = -0.05, p < .01$). Figure 2-5 graphs these terms, and shows that the initial increase is slight, and the final decrease is greater. These findings support the hypothesis that empathy decreases over time. Furthermore, the results of the Random Effects (listed at the bottom of the table) show that the random effect of linear time is
insignificant, suggesting that the linear coefficient of time is consistent across participants (but as noted earlier, random effects of nonlinear terms could not be modeled).

Hypothesis 2 predicted that self-efficacy would increase over time. Model 1 of Tables 2-6 and 2-7 show no significant time effect, and thus provide no support for the hypothesis that self-efficacy increases over time. The random effect of linear time was also insignificant, suggesting that the lack of significance for the linear effect of time are not due to differences in linear slope between individuals.

Hypothesis 3 predicted that daytime sleepiness would increase over time. Model 1a of Tables 2-8 and 2-9 show no significant time trends. However, including time X gender interactions (Model 1b) shows that changes in daytime sleepiness differ for males and females. All time parameters are significant: period ($\beta = -0.43, p < .01$), log period ($\beta = 2.39, p < .01$), period X gender interaction ($\beta = 0.56, p < .01$), and the log period X gender interaction ($\beta = -3.32, p < .01$). Plotting these interactions shows that males experience an increase in daytime sleepiness, followed by a return to baseline, while females experience a very slight decrease in sleepiness, followed by a return to baseline (Figure 2-6). Probing the interactions show that both the linear and logarithmic parameters for females are insignificant (linear: $\beta = 0.13, p > .05$; log: $\beta = -0.94, p > .05$).

**Leadership effects.** Hypotheses 4a-c predicted that leadership would explain changes in empathy. Results appear in Table 2-4. No support was found for the effects of transformational leadership (Hypothesis 4a) or abusive supervision (Hypothesis 4c). Only laissez-faire leadership (Hypothesis 4b) significantly predicted empathy ($\beta = -0.08, p < .05$; $PRV$ not interpretable$^6$).

---

$^6$ One of the potential drawbacks of using the $PRV$ is that it can be (and was in this case) negative. Due to the way in which multilevel models are constructed, when an individual predictor is negatively correlated with the group coefficient (the group being the individual-level
The empathy of medical trainees experiencing laissez-faire leadership negatively changed. It should also be noted that the cubed term for empathy was significant once the three leadership variables (i.e. transformational leadership, laissez-faire leadership, abusive supervision) are entered ($\beta = -0.03, p < .05$). This suggests that the leadership variables (likely driven by the significant effect of laissez-faire leadership) partially explain inter-individual changes in empathy.

Table 2-6 displays the results of the analysis for Hypotheses 5a-c concerning the effects of leadership on self-efficacy development. As hypothesized, transformational leadership predicted positive changes in self-efficacy ($\beta = 0.16, p < .01; PRV = 14.3\%$), supporting Hypothesis 5a; in contrast neither laissez-faire leadership (Hypothesis 5b) nor abusive supervision (Hypothesis 5c) predicted self-efficacy. In addition, once the leadership variables are entered into the model, the time terms become significant: period ($\beta = -0.38, p < .01$), square-root period ($\beta = -1.93, p < .05$), and log period ($\beta = -7.80, p < .05$). Together this shows an increase in self-efficacy, followed by a decrease (Figure 2-7), once all the leadership effects are added to the model. This suggests that leadership (largely driven by the significant effect of transformational leadership) explains significant variance around inter-individual changes in self-efficacy, and there is a mean-level pattern of self-efficacy change over time.

Hypotheses 6a-c predicted the impact of leadership on changes in daytime sleepiness. As shown in Table 2-8, both transformational leadership and laissez-faire leadership significantly predict daytime sleepiness. Transformational leadership significantly impacts daytime sleepiness, but in the opposite direction than predicted ($\beta = 0.08, p < .05; PRV = 0.0\%$); experiencing transformational leadership increases feelings of daytime sleepiness in trainees. Thus hypothesis in this case), that individual predictor increases the residual variance (Gelman & Hill, 2006). Thus the $PRV$ value here is not interpretable.
6a is not supported. (This unanticipated result is explored further in the following section of supplemental analyses.) Hypothesis 6b is supported as laissez-faire leadership positively predicts daytime sleepiness ($\beta = 0.11, p < .01; PRV = 9.8\%$). However, no support was found for the effects of abusive supervision on daytime sleepiness (Hypothesis 6c). The effects of time on the development of daytime sleepiness remain significant, but the coefficients are smaller (linear: Model 3 $\beta = -0.30, p < .01$ versus Model 2 $\beta = -0.43, p < .01$; log: Model 3 $\beta = 1.70, p < .01$ versus Model 2 $\beta = 2.40, p < .01$), suggesting that leadership explains changes in daytime sleepiness beyond the effect of time.

**Context-relevant experiences.** Hypotheses 7a-c predict the effects of morbidity and mortality experiences, beneficiary contact, and boredom, respectively, on changes in empathy. Table 2-5 displays the results of the hypothesis tests. No significant terms emerged, thus no support was found for Hypotheses 7a-c.

Hypotheses 8a-c explored the effects of context-relevant experiences on the development of self-efficacy. No support was found for the effect of morbidity and mortality experiences on self-efficacy (Hypothesis 8a). Beneficiary contact depth (but not frequency or breadth) positively predicted changes in self-efficacy ($\beta = 0.08, p < .05; PRV = not interpretable$), providing partial support for hypothesis 8b. No support emerged for the effect of boredom on self-efficacy (Hypothesis 8c).

Similar to the leadership analysis of self-efficacy, once the effects of experiences are introduced to the model, the time variables also become significant: period ($\beta = -0.32, p < .05$), square-root period ($\beta = -1.91, p < .05$), and log period ($\beta = 7.47, p < .05$) showing that self-efficacy increases, then drops (Figure 2-8). The emergence of these significant coefficients once the effects of experiences are considered, suggests that the experience variables (likely driven by
the significant effects of beneficiary contact depth) explain some of the inter-individual differences in changes, allowing an overall mean growth curve to emerge.

Last, Hypotheses 9a-b predict that morbidity and mortality and beneficiary contact will influence medical trainees’ daytime sleepiness. Morbidity and mortality increased feelings of daytime sleepiness ($\beta = 0.17, p < .01; PRV = 10.3\%$), supporting Hypothesis 9a. In terms of beneficiary contact, depth decreased feelings of daytime sleepiness ($\beta = -.06, p < .05; PRV = \text{not interpretable}$), partially supporting Hypothesis 9b. However, beneficiary contact breadth did not significantly predict daytime sleepiness, and beneficiary contact frequency positively predicted daytime sleepiness ($\beta = .11, p < .01; PRV = 8.3\%$). The more that trainees interact with patients, the sleepier they feel. In contrast, when trainees develop deeper relationships with beneficiaries, they experience less daytime sleepiness. Thus overall results for Hypothesis 9b were mixed, and are discussed further below. Finally, the coefficients of the time variables remain significant, but the effect size is smaller (linear: Model 3 $\beta = -.36, p < .01$ versus Model 2 $\beta = -.43, p < .01$; log: Model 3 $\beta = 1.94, p < .01$ versus Model 2 $\beta = 2.40, p < .01$), suggesting that the experiences help explain some of the effects of time on daytime sleepiness.

**Supplementary Analyses**

In order to (a) test the robustness of the results based on the selected covariance matrix, (b) test a rival hypothesis for changes in empathy, and to (c) further examine the results of transformational leadership on daytime sleepiness, several post hoc analyses were conducted.

**Covariance matrix sensitivity analysis.** When modeling a growth-curve, one important facet of the model is the distribution of residuals and the choice of covariance matrix to attempt to match their pattern. In some cases, the choice of covariance matrix can affect the significance of findings. The default assumption is that residuals are independent across time periods (i.e.
diagonal covariance structure). However, these assumptions are not necessarily accurate for longitudinal data, as residuals are often intercorrelated between periods (Heck, Thomas & Tabata, 2010). While there is no way to explicitly test which covariance matrix structure is appropriate, it is possible to compare models using different covariance matrices using a goodness-of-fit indicator. Tables 2-10 through 2-15 show the results of a sensitivity analysis that compares model results based on different covariance matrices. The “Final Model” in these Tables reflect the model from the main analysis, and the columns that follow compare the Final Model with three other common longitudinal covariance matrices: autoregressive, compound symmetry, and unstructured (Heck et al., 2010). The final row in the Tables lists the Akaike information criteria (AIC) for each model, a goodness-of-fit indicator that balances both variance reduction and model parsimony; lower values indicate a better fit.

Across models (Tables 2-10 to 2-15), neither the compound symmetry nor unstructured matrices produced models that converged, so their results should not be interpreted. In one case (i.e. self-efficacy and leadership, Table 2-12), the original model produced the best fit. In all other cases, the autoregressive structure was marginally better (<1% better). Finally, across all models, the significance of variables in the final model was consistent with the autoregressive model. Thus, I conclude that the results of both the final and autoregressive models demonstrate the true effects of the variables on the outcomes, and are not an artifact of assumptions about the model residuals.

**Empathy predicting attrition.** There is a possibility that participants experiencing decreases in empathy may be less likely to complete subsequent cycles of data collection,

---

7 An autoregressive structure assumes constant variance, and allows the residuals within individuals to correlate from period to period. A compound symmetry structure has constant variance and constant covariance. Finally, an unstructured covariance structure allows all terms to vary freely.
creating a bias in the results. To test this, I used multilevel regression in MPlus (as this program allows for logistic multilevel regression to predict a binary outcome) to examine whether participants’ empathy at the end of one cycle predicted their nonresponse in the next. Controlling for time, empathy in the period before did not predict whether a participant would or would not respond to the following period’s survey ($\beta = -.10$, $p > 0.05$).

**Daytime sleepiness and transformational leadership.** As noted earlier and shown in Table 2-8, transformational leadership predicted changes in daytime sleepiness, but in the opposite direction than expected. A supplemental analysis was undertaken to explore this finding further.

Based on Karasek’s (1979) and Johnson and Hall’s (1988) demand-control-support model (1979), and findings from other research that suggest females experience low control and support in the medical context (e.g., Libbus & Bowman, 1994; Nora et al., 2002; Witte, Stratton & Nora, 2006), I tested if females were more likely to experience transformational leadership as a demand, and exhibit the strain symptom of daytime sleepiness.

Since no data were collected about control or support, only the interaction between transformational leadership and gender could be tested. A significant transformational leadership $\times$ gender interaction emerged ($\beta = 0.19$, $p < .05$, $PRV = 5.1\%$; see Table 2-16, Model 4). In the presence of this significant interaction, the main effect of transformational leadership was only present for females. However, this relationship is still positive: For females, transformational leadership increased feelings of sleepiness. This finding is further explored in the discussion below.

**2.5 Discussion**
I set out to understand how different psychological experiences developed in extreme contexts. To do so, I followed a cohort of student nurses and student physicians over the first eight months of their first prolonged clinical experience. By tracking their perceptions of leadership, morbidity and mortality experiences, beneficiary contact, and boredom, I was able to gain an understanding of how empathy, self-efficacy, and daytime sleepiness develop in early-career medical professionals.

**The Development of Students’ Empathy, Self-efficacy, and Daytime Sleepiness**

My first aim in this research was to examine the overall trajectories of empathy, self-efficacy, and daytime sleepiness. The results of the time trends in empathy, self-efficacy, and daytime sleepiness are discussed below.

**Empathy.** I expected that mean empathy in medical and nursing students would decline over time. This hypothesis was supported, but the decline over time was not linear. The overall decline in empathy partially replicates previous studies on changes in empathy during medical training, which also found that empathy declines (e.g., Neumann et al., 2011; Ward et al., 2012).

My study extended previous research by also considering nonlinearity. Empathy remained high during the first few periods, and then declined during the second half of the study (Figure 2-5). The pattern of this nonlinear result is consistent with a “honeymoon-hangover” effect in job satisfaction (Boswell, Boudreau & Tichy, 2005), where a new job is initially exciting and fulfilling, but is soon followed by a drop in satisfaction. Boswell et al. (2005) argue that the initial honeymoon effect is a result of a shielding effect against unfavourable aspects of a job, and the influence of newcomer socialization, which hinders the identification of dissatisfying elements of a new job. The subsequent hangover effect is driven by disillusionment and further socialization, and a return to predisposed levels of (in their case) job satisfaction. It
is important to note, however, that the participants in Boswell et al.’s research had left their former jobs, and the honeymoon-hangover effect emerged within their new jobs, opposed to university students experiencing their first practical experience. For empathy of medical trainees, the honeymoon-hangover effect of empathy suggests that students encountering their first sustained practical experience maintain their career ideals of helping people, retaining initial levels of empathy. Here the students may be experiencing the similar shielding effect against negative aspects of the job. After a few months, the realities of the practice of medicine, and the newcomer socialization process begin influencing the decline of empathy.

Finally, results of the random model indicated that the linear time coefficient trend did not vary significantly between study participants: The mean linear trajectory illustrated by the findings was equivalent for all students, and linear changes in empathy are unlikely a result of any time-invarying variables (e.g., gender, specialty). However, as noted earlier, modeling restrictions did not allow for the examination of the random effects of the nonlinear time variables, so it remains possible that the nonlinear coefficients could vary between individuals.

Self-efficacy. I hypothesized that self-efficacy would increase over time as students emphasized successes and discounted failures in the extreme environment. However, this hypothesis was not supported. Previous research (e.g., Artino et al., 2012) showed differences in levels of self-efficacy over years of medical school, and suggested that self-efficacy would increase. But this early research was (a) correlational rather than longitudinal, and (b) examined a much greater window of time, i.e., all years of medical school. This might suggest that the eight-month period studied was not sufficient to capture overall increases in self-efficacy.

A second possibility is that diverse trajectories of self-efficacy intra-individually could “cancel-out” mean changes over time (i.e. some students had increased self-efficacy, some
students had decreased self-efficacy, which netted the mean self-efficacy to zero change), despite insignificant random linear effects (Baldwin & Hoffmann, 2002). Research on newcomer socialization suggests that self-efficacy plays a role in newcomer socialization, and that initial levels of self-efficacy moderate the socialization process and the development of subsequent self-efficacy (Saks, 1995), which suggests that self-efficacy trajectories do not have mean-level changes, but varying individual patterns. In the current data, the random effect of the linear slope coefficient was insignificant, but as noted earlier it was not possible to test whether the square-root or logarithmic coefficients of time varied across individuals. A significant random coefficient suggests between-individual variance, and would support the idea of different intra-individual patterns nullifying each other.

A third reason for not finding significant results for self-efficacy is that this particular context was not a sufficiently strong test of the hypothesis. Medical-specific self-efficacy in these students has been developing since their first day of classes (approximately 2-3 years prior to the start of this study), and likely even before then (e.g., high school, medical-related volunteer activities), as a result of which self-efficacy may have already been through changes and reached a point in its development where subsequent changes are difficult to detect.

Finally, no significant changes could be a result of the measurement, as the study of self-efficacy used a medical-specific conceptualization in this research.

**Daytime sleepiness.** Finally, I predicted that daytime sleepiness would increase over time, but this hypothesis was not supported. However, there was a significant gender X time interaction, suggesting that males (but not females) experienced an increase in daytime sleepiness, and then a return to baseline. The result for males could be explained by research showing that people adapt to new situations and sleep schedules (Horne, 2011), though
differences in this effect based on gender have not been explored. Females are more likely to experience sleep disturbances (e.g., Akerstedt et al., 2002; Tsai & Li, 2004), which would suggest that this sub-sample would increase in daytime sleepiness. The null finding could again be a symptom of diverse patterns of changes amongst females in daytime sleepiness, which leaves a mean-level net change of zero. It again could also be a weak test of the hypothesis, that may be better suited to medical residents instead of students (see for e.g., Howard et al., 2002).

**The Effects of Leadership on Changes in Empathy, Self-Efficacy & Daytime Sleepiness**

Once the mean trajectories of empathy, self-efficacy, and daytime sleepiness were examined, the second aim of this research was to explain any changes using the different leadership behaviours (i.e. transformational leadership, laissez-faire leadership, abusive supervision) that students experienced during their rotations.

**Empathy.** I initially predicted that transformational leadership would increase students’ empathy, and laissez-faire leadership and abusive supervision would decrease students’ empathy over time. Results did not support the effect of transformational leadership on empathy. Prior research has not investigated the link between transformational leadership and follower empathy, but had found relationships between transformational leadership and leaders’ own empathy (e.g., Butler & Chinowsky, 2006). In the current study, transformational leadership did not affect students’ empathy. While it is difficult to explain a null finding, this could be a result of mismatched expectations. Given the heightened focus on empathy by the medical community, and coursework prior to clinical experience that focuses on empathy, students could be expecting to see and experience empathy in practice, especially from their leaders. Actually seeing leaders exhibit empathy in practice (confirming their expectations) may not increase their empathy, but seeing behaviour that disconfirms their expectations could decrease it. This pattern was found in
research on workplace job satisfaction (Lee, 2006): Employees who expected a better workplace environment than they experienced were dissatisfied, while those whose expectations were met or exceeded did not have corresponding higher satisfaction. While empathy is not the same as job satisfaction, applying this pattern would suggest that even if students were not expecting high-quality, empathetic leadership, transformational leadership may not positively affect empathy.

Laissez-faire leadership predicted negative changes in empathy, supporting Hypothesis 4b. Prior research suggested that followers tend to mimic behaviours of leaders (e.g., Brazeau et al., 2010), and that improper treatment by supervisors was a potential cause of empathy decline (Neumann et al., 2011). The current findings support this hypothesis using a longitudinal design. This finding also extends research on the negative effects of newcomer socialization (e.g., Liu et al., 2015), as the laissez-faire behaviours of supervisors appear to be a cue to students that empathy and caring is not valued.

Finally, the hypothesis that abusive supervision would have a negative effect on changes in empathy was not supported. I would suggest that range restriction explains the non-significant finding: Abusive supervision was uncommon in this sample ($M = 1.21, SD = 0.56$, theoretical range = 1-5), and was highly skewed (skewness = 4.13; $SE = 0.19$) and kurtotic (kurtosis = 19.93; $SE = 0.34$). This range restriction and non-normality could contribute to the absence of significant results (Johns, 1996). Range restriction is a common problem in the research of abusive supervision (e.g., theoretical range of 1 to 5: $M = 1.34, SD = 0.57$, Byrne et al., 2014; $M = 1.39, SD = 0.60$, Tepper, Henle, Schurer Lambert, & Giacalone, 2008; $M = 1.27, SD = 0.47$, Tepper, Moss, & Duffy, 2011), but when researchers are able to uncover findings, their results
suggest that abusive supervision is still important to study despite its infrequent occurrence (e.g., Tepper, 2000).

**Self-efficacy.** I expected that transformational leadership would predict positive changes in self-efficacy, and that laissez-faire leadership and abusive supervision would predict negative changes. As hypothesized, transformational leadership predicted positive changes in self-efficacy, supporting Hypothesis 5a. This effect was expected because transformational leadership provides individual encouragement and inspirational motivation, which fosters individuals’ self-efficacy beliefs through verbal persuasion and social influence (Bandura, 1982). In addition, transformational leadership decreases the psychological arousal (e.g., lowered stress) necessary for individuals to feel confident in their abilities, further enhancing self-efficacy. By studying the dynamic development of self-efficacy over time, my study builds upon prior research showing a cross-sectional link between transformational leadership and follower self-efficacy (e.g., Chen & Bliese, 2002). This is an important finding as self-efficacy is related to numerous positive outcomes, both in terms of physical performance and psychological outcomes (e.g., Judge & Bono, 2001).

In contrast, no support emerged for the effects of laissez-faire leadership on the development of self-efficacy. Insignificant relationships between laissez-faire leadership and outcomes such as follower motivation (Judge & Piccolo, 2004), resilience (Harland, Harrison, Jones, & Reiter-Palmon, 2005), and burnout (Corrigan, Diwan, Campion & Rashid, 2002) are not uncommon. Thus, perhaps the absence of verbal persuasion, through appropriate feedback, by itself is not enough to have a negative impact on self-efficacy. Furthermore, as discussed earlier, many of the relationships where laissez-faire leadership was destructive (e.g., conflict, bullying), are driven by the stress that laissez-faire leadership creates (e.g., Skogstad et al.,
The stress created by laissez-faire leadership was also expected to negatively affect self-efficacy, as stress represents negative psychological arousal. But psychological arousal is the least important of Bandura’s (e.g., 1982) antecedents, expected to have the least impact on self-efficacy. Either way, the effect (or null effect) of laissez-faire on self-efficacy remains in need of an explanation.

Finally, the hypothesis that abusive supervision would negatively affect self-efficacy was not supported. This insignificant result is likely a result of the range restriction and floor effect, as abusive supervision was not common in this sample, as discussed earlier.

**Daytime sleepiness.** I hypothesized that transformational leadership would predict decreases in daytime sleepiness, and that laissez-faire leadership and abusive supervision would predict increases in daytime sleepiness. Transformational leadership did indeed predict changes in daytime sleepiness, but in the opposite direction than I had predicted. Notwithstanding my initial hypotheses, recent research suggests that the current findings could be accurate: Syrek and Antoni (2014) studied the effects of unfinished tasks at work on rumination and sleep loss at night, and found that the relationship was strongest when paired with a supervisor with high performance expectations. The authors liken these high performance expectations to those of transformational leadership, and theorize that these behaviours create increased stress for employees, which in turn is detrimental to sleep. With respect to the current study, early in a students’ career, high performance expectations while simultaneously adapting to a new environment could be negatively affecting sleep.

Despite the support for a positive relationship between transformational leadership and daytime sleepiness provided by Syrek and Antoni (2014), I conducted supplementary analyses to further probe these results. As Syrek and Antoni (2014) suggest, the relationship between
transformational leadership and rumination may be significant because high performance expectations are a stressor. Following this line of reasoning, and applying Karasek’s demand-control-support model (1979), the positive relationship between transformational leadership (with high performance expectations as a demand) and daytime sleepiness (a strain) would only be expected by those experiencing low control or support.

While data on control or support were not collected, considering the unique sample at hand (and the available variables), it might be expected that female students experience less control (as I will discuss below), and may be more likely to negatively react to transformational leadership. Female medical students are more likely than male students to experience discrimination and sexual harassment (Nora et al., 2002; Witte, Stratton & Nora, 2006). This discrimination and sexual harassment can be interpreted as a lack of support for female medical students, and may also result in feelings of loss of control, as no one else recognizes or attempts to remedy the situation. Nurses in general are exposed to discrimination and harassment by both male co-workers and patients (Libbus & Bowman, 1994), and experience higher levels of workplace bullying than other hospital staff (Quine, 2001)\(^8\). Again, these experiences represent a lack of control and support in their environments.

The disproportionately high number of females in the sample could be responsible for the positive relationship between transformational leadership and daytime sleepiness. Thus I explored the moderating effects of gender on the relationship between transformational leadership and daytime sleepiness. Testing the moderating effects of gender in this sample

\(^8\) While these studies do not look at interactive effects of gender due to the dominance of females in the field, it has been suggested that male nurses tend to enjoy advantages when working in female-dominated fields, opposed to the negative effects of females working in male-dominated fields (Williams, 1992). The advantages represent greater control and support for male nurses than female nurses.
produced significant results, showing that the positive relationship between transformational leadership and daytime sleepiness was only significant for females (with an insignificant relationship for males). This supports the notion that the relationship between transformational leadership and daytime sleepiness differs for males and females, perhaps due to the lower control and support for females (but this requires further investigation).

Moving away from Karasek’s model to account for the gender effects, I would suggest that students who are sleepier at the beginning of a rotation may react negatively to transformational leadership, and thus experience even greater sleepiness, while those who begin a rotation less sleepy react to transformational leadership in the predicted way. The basis for this is that when individuals are ego-depleted from sleepiness (Wagner, Barnes, Lim & Ferris, 2012), I would expect that transformational leadership further depletes followers, as transformational leadership in the presence of low resources could be stressful. Individuals with low levels of daytime sleepiness do not react to transformational leadership as a stressor, and instead are invigorated, resulting in decreased daytime sleepiness. However, attempts to test this model with the current study data were unsuccessful as the models did not converge.

Laissez-faire leadership predicted daytime sleepiness, supporting hypothesis 6b. Because of its absence of stimulation, and because it is experienced as stressful, laissez-faire leadership negatively affects students’ daytime sleepiness9. This extends prior findings that laissez-faire leadership can detrimentally impact followers.

---

9 While the potential stress caused by laissez-faire leadership was not detrimental to self-efficacy (as psychological states are the least important of Bandura’s antecedents), stress is a salient predictor of sleep problems (e.g. Akerstedt et al., 2002; Hall et al., 2004; Linton, 2004).
Finally, there were no significant effects of abusive supervision on daytime sleepiness. As I explained earlier, range restriction might be the most plausible explanation, as the relative infrequency of abusive supervision in the current sample could contribute to these findings.

**Context-Relevant Experiences**

My final goal in this research was to investigate the impact of different experiences associated with working in extreme contexts, namely morbidity and mortality experiences, beneficiary contact, and boredom, on empathy, self-efficacy, and daytime sleepiness were explored.

**Empathy.** While I predicted that the morbidity and mortality, beneficiary contact, and boredom that students endured during their placements would influence changes in empathy, no significant results emerged. These experiences were theorized as reasons for empathy decline previously (e.g., patient death, not being able to help patients; Neumann et al., 2011; Hojat, et al., 2009), but no empirical research on causes of empathy decline had been undertaken.

Stress and adaptive hardening to protect oneself were common reasons used to explain empathy decline (e.g., Neumann et al., 2011), and so the morbidity and mortality scale was developed to capture these experiences. However, the results showed that exposure to morbidity and mortality does not impact changes in empathy. As suggested by Karasek (1979), decreased empathy might only occur where students experience morbidity and mortality in the face of low control and low support.

None of the beneficiary contact factors (i.e. frequency, breadth, depth) predicted changes in empathy. Research suggested that beneficiary contact would increase empathy, especially since beneficiary contact has been shown to increase perspective-taking (Grant, 2008b). However that research was conducted in a traditional organization, and it is possible that this
represents a case where findings do not translate into a more extreme environment. In traditional organizations, beneficiary contact and helping others is not often a job requirement. Because it is not a core job requirement, perhaps this is why beneficiary contact has strong effects on employees – because it is a deviation from the norm. In the case of patient care, beneficiary contact is more embedded into the expectations of a job. Perhaps this nuance can help explain why beneficiary contact may not facilitate perspective-taking in medicine, but further research is needed to understand this.

Finally, I expected that boredom would decrease empathy, through the affective consequences of boredom, and the lack of meaningfulness of students’ work. No significant effect emerged. However, this insignificant main effect might be explained by recent research that shows that boredom is multidimensional, comprised of five factors: time passing slowly, low arousal, inattention, high arousal, and disengagement (Baratta & Spence, 2015). With one factor capturing high arousal, and another capturing low arousal, perhaps it is due to these very different experiences and responses to boredom that main effects could not be discerned. In my earlier hypothesizing, I focused largely on disengagement, and partially on high arousal (i.e., restlessness), and by not being able to separately study the different dimensions was not able to adequately discern boredom’s true effects.

**Self-efficacy.** I predicted that morbidity and mortality experiences and boredom would change self-efficacy negatively, while beneficiary contact would change self-efficacy for the better.

Morbidity and mortality experiences predicted negative changes in self-efficacy, (Hypothesis 8a), likely because morbidity and mortality experiences constitute negative feedback about performance failures, which then deplete self-efficacy. While the dynamic changes in self-
efficacy had not been studied in prior research, nor the possible causes of these changes, these results support and extend Bandura’s original theory (1977) of self-efficacy.

Of the three factors of beneficiary contact (i.e. frequency, breadth, depth), depth predicted changes in self-efficacy, partially supporting Hypothesis 8b, but frequency and breadth of patient experiences did not. Thus, being able to form deeper relationships with beneficiaries yields positive benefits for self-efficacy. This supports and extends Bandura’s original theory (1977) on self-efficacy, as deeper, higher quality relationships are more likely to be characterized by open and honest performance feedback. It also extends research on the effects of interactions with beneficiaries (e.g., Grant, 2008b): These results suggest that it is not important whether a student interacts with many beneficiaries, or a variety of different beneficiaries, (which, as noted earlier, is likely a role expectation of students), but it is important to be able to cultivate deeper relationships with beneficiaries.

Finally, no support was found for the effects of boredom on self-efficacy. I theorized that boredom would decrease self-efficacy, as boredom could represent a lack of challenge, and also represents a source of stress. Boredom does represent negative emotional arousal, but as noted earlier, psychological states are the weakest predictor of self-efficacy, and alone may not be enough to detrimentally impact self-efficacy. The multidimensional nature of boredom (Baratta & Spence, 2015), discussed above, could also be contributing to the null finding.

**Daytime sleepiness.** I expected that morbidity and mortality experiences would increase daytime sleepiness, and beneficiary contact would decrease daytime sleepiness. Both variables predicted changes in the daytime sleepiness of medical trainees.

Morbidity and mortality experiences increased daytime sleepiness, supporting Hypothesis 9a. The stress and depletion resulting from morbidity and mortality experiences increases
daytime sleepiness. While the specific effects of these negative experiences had not been explored in prior research, evidence did suggest that stress (e.g., Rafferty et al., 2010) and traumatic experiences (e.g., Lavie, 2001) would negatively impact sleep quality. This finding extends previous research, demonstrating the effect of morbidity and mortality experiences on the dynamic development of daytime sleepiness over an extended period of time.

Beneficiary contact also predicted daytime sleepiness. Of the three factors of this variable, frequency and depth (not breadth) significantly predicted changes in daytime sleepiness. Depth of beneficiary contact decreased daytime sleepiness, as predicted in Hypothesis 9b. Lilius (2012) suggested that the quality of the relationship with customers leads not only to less burnout, but to greater invigoration. As she explained, interactions with clients (in this case patients) have the potential to be depleting or restorative. Whether they are depleting or restorative depends on (a) how many regulatory resources are required, and (b) whether they generate three resources shown to counteract or buffer resource drain: positive affect, self-affirmation, and perceived prosocial impact. An examination of what the depth of beneficiary contact captures (i.e. building close relationships with people affected by one’s work, forming emotional connections with people who benefit from one’s work, and having meaningful communications with people benefitting from one’s work) points to its potential restorative nature.

Conversely, and contrary to Hypothesis 9b frequency of beneficiary contact predicted increases in daytime sleepiness. This finding again supports Lililus’s (2012) theory: In the absence of restorative resources, client interactions can be experienced as depleting. This finding is also supported by Karasek’s (1979) demand-control model. Frequent interactions could be considered equivalent to quantitative role overload – a job demand – and in the absence
of control, result in a strain response of increased daytime sleepiness. Supporting this, high workload is considered a workplace stressor (e.g., Krantz, Berntsson & Lundberg, 2005) and places a demand on resources (e.g., Crawford, LePine & Rich, 2010).

**Strengths and Limitations of the Current Study**

This study benefitted from several conceptual and methodological strengths. Conceptually, this study was the first to examine changes in empathy, self-efficacy, and daytime sleepiness. Historically, researchers have often neglected to account for time, and the need to study whole events and processes (Johns, 2006). Within the framework of newcomer socialization, the current research shows that empathy, self-efficacy, and daytime sleepiness developed dynamically over time, considering the first eight-months of prolonged patient contact. This study then incorporated possible predictors of that change. Even research using growth-curve modeling tends to place a strong emphasis on the effect of time, as the majority of growth-curve models look at differences in trajectories based on differences in static variables (e.g., personality and previous experiences; Chan & Schmitt, 2000; race and gender; Willson, 2003). Of course this type of model answers important questions and should not be discounted, but by incorporating time-varying variables to explain change, the dynamic development of variables can be more accurately modeled, enhancing the ecological validity of such models. This also allows researchers to incorporate the effects of time (e.g., the honeymoon effect; Helmreich, Sawin, & Carsrud, 1986), while also incorporating predictors of change beyond the effects of time.

Methodologically, I used a longitudinal, multilevel analysis to examine change and its potential predictors, extending the focus past most research, which has been rooted in cross-sectional questions. Finally, potential biases relating to attrition were ruled out.
Interpretations and implementation of study results cannot be made without regard for possible threats to their validity. First, all the data were based on participants’ self-report. Self-report data are common as they are often the easiest to collect, but are threatened by a variety of response biases (Donaldson & Grant-Vallone, 2002). Supplemental sources of data would have been beneficial, for example, objective measures of patient morbidity and mortality, or electronic data on sleep quality, in order to exclude plausible rival hypotheses based on mono-source bias. However self-report surveys were the most appropriate and feasible method for this study, especially for variables such as self-efficacy, which is a personal, subjective expectation of one’s future performance, and empathy, which is also an internal cognitive and affective process.

Second, one particular response bias warrants further consideration: The effect of demand characteristics on participants’ responses. Tendencies to hypothesis-guess, and general information-seeking by newcomers to infer what is “normal” could result in responses that are supportive of the hypotheses. To combat hypothesis guessing, no specific details about the aims of the study were shared with participants. Replicating the between-person model over multiple time periods also limits the threat of this bias (Rubin, Paolini, & Crisp, 2010). As for the within-person changes found for the outcome variables, it is possible that respondents were influenced by information from their environments about “usual” changes over time. Empathy and its decline are discussed in medical school, and anecdotally, medical students and residents take and compare the Epworth Sleepiness Scale for entertainment. In terms of empathy, suggesting that its overall decline is a function of a demand characteristic is somewhat consistent with my earlier supposition that empathy decline is a function of culture, and that students takes cues from their environments on how to behave. Whether the decline demonstrated is truly how respondents feel, or they are just responding how they are “supposed to” is difficult to discern, however
participants answered these (approximately) ten-minute surveys every six weeks, and I believe that given the demands of their daily lives, it is unlikely that participants would remember how they answered in the previous cycle (and they did not have access to their previous responses), let alone be able to emulate expected changes over time with their responses. Thus I do not believe this is a legitimate threat.

Third, the sample was drawn from one university only, which could limit the generalizability of study findings. However, students’ placements were not restricted to the university, and the sample reflected institutions across the province. Fourth, the study used a retrospective longitudinal design, asking participants about the previous six weeks’ experiences, and then their empathy, self-efficacy, and daytime sleepiness that day. While it would have been optimal to measure predictors and outcomes separately at different points in time, this form of repeated measures has commonly been used in psychological research when predictors are time-varying (e.g., Baldwin & Hoffmann, 2002; Galambos, Barker, & Krahn, 2006; Raudenbush & Chan, 1993).

Fifth, responses in this study could be affected by non-response bias, where the responses and effects for non-responders differ from those included in the study. While this is a threat, and inferences and applications of findings should be made with this in mind, it should also be noted that the gender proportions and mean ages are consistent with that of similar programs in the United States (e.g. Association of American Medical Colleges, 2014; National League for Nursing, 2012). Furthermore, recent studies show that the expense of increasing response rate is often not justified given the difference in survey accuracy (e.g. Curtain, Presser & Singer, 2000; Keeter Kennedy, Dimock, Best & Craighill, 2006).
Sixth, the sample was disproportionately female (81%), which could affect the generalizability of results. However, I suggest that this does not threaten generalizability within this context because (a) this is similar to medical and nursing programs elsewhere and (b) gender did not significantly predict the outcome variables in the final models. However, generalizability to other contexts may be limited.

Seventh, while data collection began prior to students’ first sustained clinical experiences in an effort to control for potential confounding factors, the period studied does not capture students’ first patient interactions. As part of their programs, students have the opportunity to interact with patients (in a more limited manner), and it is also not unreasonable to expect that students with an interest in medicine would also have engaged in volunteer activities before starting their programs, and that these prior patient experiences could have already affected their empathy, self-efficacy, and daytime sleepiness. This could weaken the strength of the tests. Similarly, it is also possible that limiting the focus to the first eight months excluded later developmental experiences.

Eighth, multiple hypotheses were tested in the same analysis (e.g., three leadership behaviours in a single regression), which suggests the possibility of experiment-wise error (or family-wise error), and thus risk of type I error. Historically, researchers have adjusted for this problem by widening confidence intervals (or reducing p-values), but recent research suggests that the common adjustments to prevent such errors are not necessary when using multilevel modeling: Multilevel modeling produces more efficient estimates that eliminate the need for multiple comparison adjustments (Gelman, Hill & Yajima, 2012). As the authors explained, multilevel modeling makes use of partial pooling in its estimation, by using a weighted combination of group-level effect (i.e., the individual across time periods), and the population-
level effect (i.e., the entire sample). The result is shrinkage in the variance of the estimation, and a movement of that estimate to its accurate value. Illustrated, this would provide much tighter confidence intervals than those produced by non-multilevel techniques. The authors concluded that corrections for multiple hypotheses are not necessary for the already efficient estimates.¹⁰

Ninth, insufficient variance in leadership variables could limit the ability to find significant relationships. However, while it is difficult to determine what is sufficient, all of the standard deviations in this study are similar to those in other research: The standard deviation of transformational leadership in this study was 0.85 ($M = 4.01$), compared to other research using a similar scale (range = 1 – 5), with a standard deviation of approximately 0.90 ($M = 3.30 – 3.71$) for each of the four sub-factors (Beauchamp et al., 2014), and 0.75 ($M = 3.61$) for the full scale (Johannson, & Rajha Olsson, 2014); the standard deviation of laissez-faire ($SD = 1.23, M = 2.73$, theoretical range = 1 – 7) was similar to that in other research using the same scale as this study (e.g., $SD = 1.37, M = 3.71$, reward omission, $SD = 1.83, M = 2.50$, punishment omission, Hinkin & Schreisheim, 2009; $SD = 0.97, M = 2.03$, Saboe, 2012), and the laissez-faire portion of the Multifactor Leadership Questionnaire (e.g., range = 0 – 4: $SD = 1.10, M = 2.46$, Chaudhry, & Javed, 2012; $SD = 0.48, M = 0.52$, Skogstad et al., 2007). As noted earlier, the standard deviation of abusive supervision from this study was low, but was similar to that found by others (e.g., theoretical range of 1 to 5: $SD = 0.57, M = 1.34$, Byrne et al., 2014; $SD = 0.60, M = 1.39$, Tepper, Henle, Schurer Lambert, & Giacalone, 2008; $SD = 0.47, M = 1.27$, Tepper, Moss, &

¹⁰ For informational purposes, I have added additional notation to each table identifying whether a significant relationship was also below the Bonferroni correction p-value (e.g. Dunn, 1959). But given that this correction is conservative (e.g. Perneger, 1998), and the findings by Gelman, Hill and Yajima (2012) discussed above, I am confident that the original estimates presented are accurate.
Duffy, 2011). Thus, variance is not likely a problem that threatens the strength of the tests (with the exception of the range restriction of abusive supervision).

Finally, while a longitudinal design was used and allowed an examination of change over time, causality cannot be inferred. There remains a possibility that, for example, these variables simply move in the same direction, and the underlying cause of the change over time lies elsewhere. Causation must therefore await experimental manipulation.

**Future Directions**

This study paves the way for future research. Conceptually, my findings suggest some avenues for future research. First, what other experience or leadership variables should be considered when studying both extreme and traditional contexts? For example, only one variable (i.e. laissez-faire leadership) explained changes in empathy. Changes in empathy are not random, and given the importance of empathy in the area of medicine in terms of patient health (e.g., patient compliance, patient-physician information exchange; Kim et al., 2004), should be explored further. In addition, sleepiness is an important predictor of outcomes at work, including emotional (e.g., Ouweneel, Le Blanc, Schaufeli, van Wijhe, 2012), and self-control-related (e.g., Barnes, 2012) variables. Sleepiness might also predict empathy decline. Recent research by Guadagni, Burles, Ferrara, and Iaria (2014) showed in an experimental manipulation that sleep deprivation negatively affected emotional empathy, extending findings that sleep is detrimental to emotional processing, and highlighting an interesting avenue for future research.

Second, on a related note, is the seemingly widespread assumption that empathy decline is unfavourable. However, given the potential negative effects of having high empathy (e.g., negative reactions to trauma; Regehr et al., 2002), perhaps some empathy decline is functional.
Future research would benefit from taking a curvilinear approach to the study of empathy and investigate potential optimal levels of empathy that benefit both patients and medical caregivers.

Third, the counterintuitive finding about the relationship between transformational leadership and daytime sleepiness warrants further research. Post-hoc analysis showed that this relationship held only for females, possibly due to uniquely stressful experiences facing women in healthcare (and potentially, other contexts, too). Future research should further explore this effect, as discussed earlier.

Fourth, findings from this research point to the importance of the depth of beneficiary contact. Grant (e.g., Grant, 2008a, Grant & Parker, 2009) showed that beneficiary contact is an important variable, but did not separate the variable into its three components; Lilius (2012) did so, and suggested that beneficiary contact depth and quality are perhaps the most important factors of contact. My findings add to this growing body of research by showing that depth of interactions with beneficiaries can increase feelings of self-efficacy, and decrease feelings of daytime sleepiness in medical trainees. At the same time, frequency of beneficiary contact increased daytime sleepiness – the opposite effect to that of depth. This highlights the importance of separately studying the components, and the need to understand each component’s unique effects. Future research should further investigate the role of beneficiary interactions.

Fifth, while I only examined the direct effects of leadership, morbidity and mortality, beneficiary contact, and boredom, future research might consider any compounding or attenuating effects of interactions between these variables (e.g., Byrne et al., 2014). For example, Bellé (2013) found that the relationship between transformational leadership and performance was strongest in the presence of beneficiary contact, and the importance of these finding are enhanced because they were derived from an experimental field study. Might this
interaction also predict daytime sleepiness or empathy? Might beneficiary contact depth reduce the negative effects of morbidity and mortality on sleep? Might a combination of the variables (e.g., morbidity and mortality experiences by frequency of patient interactions) help explain changes in empathy?

Finally, my study investigated changes during the first year of sustained clinical experience in medical training. This is a critical career juncture due to the potential long-term effects of newcomer socialization on their behaviour and cognition (e.g., Scott & Myers, 2005), and it would be important to continue to examine the development of empathy, self-efficacy, and daytime sleepiness. Research examining change beyond this early period is sparse, yet support for continuing to focus on this development derives from Figure 2-5, which shows a steep decline in empathy starting at approximately four months into students’ clinical experience and continuing to the end of the studied eight months. Research shows that empathy declines during residency as well (Neumann et al., 2011), but theoretically empathy should at least plateau (by design, as our scales are measured on a point scale) at some point, but no research examining change in practicing physicians and nurses exists. It might be expected that self-efficacy will overall increase and also plateau at some stage during students’ early careers, especially the medical-specific self-efficacy that was studied in this study, but again no evidence exists that points to when this might happen. In terms of sleepiness, the residency portion of training for physicians is especially concerning (e.g. Whetsell, 2003), so understanding how it develops, and how physicians adapt and recover once full-time is important. Correlational research suggests that approximately one-quarter of practicing physicians have abnormally high daytime sleepiness (Chen et al., 2008): Gaining a larger picture of the development of these important variables might give us a better understanding of how certain individuals became who they did.
Methodologically, the current findings raise some possibilities for future research. First, expanding the sources of data would augment study results, and avoid potential mono-source and mono-method bias. Supervisor-ratings or patient-ratings of study variables such as empathy, and objective measures of sleep quality could be useful. Second, my study was limited in the number of possible random effects, which prohibited me from examining the random effects of the nonlinear time variables. Future research should investigate whether these nonlinear time effects differ between individuals, and can accomplish this by increasing the number of time points. Third, methodological advances have made the possibility of daily studies a reality. Examining the experiences of medical trainees on a more consistent basis could uncover the more subtle changes and their causes. Fourth, similarly, statistical advances in the development of cross-lagged structural equation models could be applied to similar longitudinal research, and help better-answer questions about causation (e.g. Burkholder & Harlow, 2003). Finally, varying the context to other medical programs, other countries, or other caregiving professions, would improve the external validity of study findings, and the widespread applicability.

**Practical Implications**

The significant results from this study point to implications for the training of medical professionals. First, declines in empathy were evident. To combat this, it might be useful to increase empathy before students enter their first practical experience. While this may not prevent the declines, increasing the initial intercept of empathy may make declines less problematic with respect to patient care. Programs that focus on empathy are being tested and implemented (e.g., Riess et al., 2012), and should be augmented before and during patient interactions, to build psychological resources (i.e. mental and physical well-being).
Second, my results showed that transformational leadership and laissez-faire leadership significantly affect medical trainees. Research has shown that leadership can be trained (e.g., Brown & May, 2012), and medical programs would benefit from leadership training for staff in supervisory positions to improve the well-being of medical trainees. Furthermore, mentorship can be a positive resource for newcomers that facilitates adjustment (e.g., Allen, et al., 1999; Ostroff & Kozlowski, 1993). Nursing and medical programs may benefit from the formal implementation of mentoring programs using senior students, residents, or junior staff to supplement the leadership that students may (or may not) be receiving from supervisors.

Third, morbidity and mortality experiences negatively influenced self-efficacy and daytime sleepiness. To combat these negative effects, medical and nursing programs could provide resources and support for dealing with these negative experiences (e.g., counselors). Implementing resilience training prior to patient interactions may also help buffer medical and nursing students from morbidity and mortality and leave them better equipped to continue to provide high quality care. Resilience training has been successfully implemented in a sample of physicians, creating lowered stress and higher quality of life (e.g., Sood, Prasad, Schroeder, & Varkey, 2011).

Finally, beneficiary depth predicted changes in both self-efficacy and daytime sleepiness. Medical and nursing programs could focus on giving trainees opportunities to develop deeper relationships with patients to foster growth of self-efficacy and better daytime sleepiness. Such programs should turn the focus away from frequency of patient contact, which increases daytime sleepiness.

Conclusion
My study was the first to examine potential causes in the changes of empathy, self-efficacy, and daytime sleepiness in medical trainees during their first sustained practical experience. Using a multilevel longitudinal design, I found that the transformational and laissez-faire leadership behaviours of immediate supervisors, as well as personal experiences in morbidity and mortality, and beneficiary contact, contributed to changes in trainee empathy, self-efficacy, and daytime sleepiness. Areas of future research were discussed, as well as practical implications of the study findings for the training of our future medical caregivers.
2.6 References


Table 2-1: Scale reliabilities and descriptive statistics from pilot sample

<table>
<thead>
<tr>
<th>Scale</th>
<th>Number of items</th>
<th>Reliability</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis’s perspective taking</td>
<td>7</td>
<td>0.84</td>
<td>3.62</td>
<td>0.64</td>
</tr>
<tr>
<td>Davis’s empathic concern</td>
<td>7</td>
<td>0.79</td>
<td>3.91</td>
<td>0.62</td>
</tr>
<tr>
<td>Medical-specific self-efficacy, shortened</td>
<td>14</td>
<td>0.91</td>
<td>3.74</td>
<td>0.54</td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td>8</td>
<td>0.80</td>
<td>2.30</td>
<td>0.53</td>
</tr>
<tr>
<td>TFL, shortened</td>
<td>8</td>
<td>0.95</td>
<td>3.59</td>
<td>0.97</td>
</tr>
<tr>
<td>Laissez-faire leadership</td>
<td>8</td>
<td>0.89</td>
<td>3.03</td>
<td>1.25</td>
</tr>
<tr>
<td>Abusive supervision</td>
<td>10</td>
<td>0.92</td>
<td>1.33</td>
<td>0.56</td>
</tr>
<tr>
<td>Boredom</td>
<td>6</td>
<td>0.79</td>
<td>1.90</td>
<td>0.51</td>
</tr>
<tr>
<td>Morbidity &amp; mortality</td>
<td>12</td>
<td>0.91</td>
<td>2.17</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Table 2-2: Descriptive statistics and scale reliabilities

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Scale Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empathy</td>
<td>166</td>
<td>3.92</td>
<td>0.51</td>
<td>.83</td>
</tr>
<tr>
<td>Efficacy</td>
<td>166</td>
<td>3.56</td>
<td>0.66</td>
<td>.93</td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td>166</td>
<td>2.00</td>
<td>0.53</td>
<td>.80</td>
</tr>
<tr>
<td>Sleep hours</td>
<td>166</td>
<td>6.63</td>
<td>1.11</td>
<td>NA</td>
</tr>
<tr>
<td>Transformational leadership</td>
<td>166</td>
<td>4.08</td>
<td>0.85</td>
<td>.96</td>
</tr>
<tr>
<td>Laissez-faire leadership</td>
<td>166</td>
<td>2.73</td>
<td>1.23</td>
<td>.93</td>
</tr>
<tr>
<td>Abusive supervision</td>
<td>166</td>
<td>1.21</td>
<td>0.56</td>
<td>.95</td>
</tr>
<tr>
<td>Morbidity &amp; mortality experience</td>
<td>166</td>
<td>1.83</td>
<td>0.54</td>
<td>.85</td>
</tr>
<tr>
<td>Beneficiary contact- frequency</td>
<td>166</td>
<td>5.40</td>
<td>1.40</td>
<td>.98</td>
</tr>
<tr>
<td>Beneficiary contact- breadth</td>
<td>166</td>
<td>5.04</td>
<td>1.48</td>
<td>.95</td>
</tr>
<tr>
<td>Beneficiary contact- depth</td>
<td>166</td>
<td>4.91</td>
<td>1.64</td>
<td>.97</td>
</tr>
<tr>
<td>Boredom</td>
<td>166</td>
<td>2.02</td>
<td>0.70</td>
<td>.90</td>
</tr>
<tr>
<td><strong>Level 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>39</td>
<td>0.81</td>
<td>0.40</td>
<td>NA</td>
</tr>
<tr>
<td>Specialty</td>
<td>39</td>
<td>0.45</td>
<td>0.50</td>
<td>NA</td>
</tr>
<tr>
<td>T1 Empathy</td>
<td>39</td>
<td>4.00</td>
<td>0.48</td>
<td>.82</td>
</tr>
<tr>
<td>T1 Efficacy</td>
<td>39</td>
<td>3.40</td>
<td>0.58</td>
<td>.92</td>
</tr>
<tr>
<td>T1 Sleepiness</td>
<td>39</td>
<td>1.94</td>
<td>0.46</td>
<td>.70</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>28</td>
<td>5.34</td>
<td>0.93</td>
<td>-.25</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>1.</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Gender</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Specialty</td>
<td>.08</td>
<td>-.48**</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>T1 Empathy</td>
<td>-.04</td>
<td>.17*</td>
<td>-.14</td>
</tr>
<tr>
<td>5.</td>
<td>T1 Efficacy</td>
<td>-.04</td>
<td>.10</td>
<td>-.56**</td>
</tr>
<tr>
<td>6.</td>
<td>T1 Sleepiness</td>
<td>-.08</td>
<td>-.02</td>
<td>.05</td>
</tr>
<tr>
<td>7.</td>
<td>Empathy</td>
<td>-.05</td>
<td>.12</td>
<td>-.17*</td>
</tr>
<tr>
<td>8.</td>
<td>Efficacy</td>
<td>.06</td>
<td>.10</td>
<td>-.36**</td>
</tr>
<tr>
<td>9.</td>
<td>Sleepiness</td>
<td>-.13</td>
<td>-.08</td>
<td>.07</td>
</tr>
<tr>
<td>10.</td>
<td>Sleep hours</td>
<td>-.06</td>
<td>.06</td>
<td>-.09</td>
</tr>
<tr>
<td>11.</td>
<td>TFL</td>
<td>.12</td>
<td>.17*</td>
<td>-.14</td>
</tr>
<tr>
<td>12.</td>
<td>LF</td>
<td>.05</td>
<td>-.31**</td>
<td>.21**</td>
</tr>
<tr>
<td>13.</td>
<td>AS</td>
<td>-.11</td>
<td>-.16*</td>
<td>.01</td>
</tr>
<tr>
<td>14.</td>
<td>Mortality</td>
<td>.32**</td>
<td>-.27**</td>
<td>.39**</td>
</tr>
</tbody>
</table>

Note: TFL = Transformational leadership; LF = Laissez-faire leadership; AS = Abusive supervision; BC = Beneficiary contact

* p < .05, ** p < .01
Table 2-4: Growth-curve models of empathy with leadership

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>PRV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.04 (.08)</td>
<td>.05 (.22)</td>
<td>.18 (.15)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-0.29* (.12)</td>
<td>-.12 (.18)</td>
<td>-.15 (.15)</td>
<td></td>
</tr>
<tr>
<td>Time^2</td>
<td>0.24** (.08)</td>
<td>.15 (.12)</td>
<td>.16 (.09)</td>
<td></td>
</tr>
<tr>
<td>Time^3</td>
<td>-0.05** (.01)</td>
<td>-.03 (.02)</td>
<td>-.03* (.01)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>.01 (.18)</td>
<td>-.14 (.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specialty</td>
<td>-.14 (.15)</td>
<td>-.12 (.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T₁ Empathy</td>
<td>.46* (.08)</td>
<td>.61** (.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFL</td>
<td>.04 (.05)</td>
<td>n.i.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF</td>
<td>-.08* (.03)</td>
<td>n.i.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>.04 (.05)</td>
<td>0.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.18** (.05)</td>
<td>.05 (.05)</td>
<td>.04 (.03)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.00 (.00)</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>.073** (.01)</td>
<td>.075** (.02)</td>
<td>.073** (.01)</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05; **p < .01

a p < Bonferroni correction

PRV = proportion reduction in variance

n.i. = not interpretable
Table 2-5: Growth-curve models of empathy with experiences

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>PRV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.04 (.08)</td>
<td>.05 (.22)</td>
<td>.17 (.16)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-.29* (.12)</td>
<td>-.12 (.18)</td>
<td>-.18 (.16)</td>
<td></td>
</tr>
<tr>
<td>Time^2</td>
<td>.24** (.08)</td>
<td>.15 (.12)</td>
<td>.16 (.10)</td>
<td></td>
</tr>
<tr>
<td>Time^3</td>
<td>-.05** (.01)</td>
<td>-.03 (.02)</td>
<td>-.03* (.02)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>.01 (.18)</td>
<td>-.06 (.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specialty</td>
<td>-.14 (.15)</td>
<td>-.14 (.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T_1 Empathy</td>
<td>.46* (.08)</td>
<td>.59** (.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morbidity &amp; mortality experience</td>
<td>-.05 (.07)</td>
<td></td>
<td>0.4%</td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact- frequency</td>
<td>-.00 (.05)</td>
<td></td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact- breadth</td>
<td>.00 (.05)</td>
<td></td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact- depth</td>
<td>.02 (.04)</td>
<td></td>
<td>0.8%</td>
<td></td>
</tr>
<tr>
<td>Boredom</td>
<td>-.04 (.05)</td>
<td></td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.18** (.05)</td>
<td>.05 (.05)</td>
<td>.05 (.03)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>.073** (.01)</td>
<td>.075** (.02)</td>
<td>.073** (.01)</td>
<td></td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$  
\* $p$ < Bonferroni correction  
PRV = proportion reduction in variance  
n.i. = not interpretable
Table 2-6: Growth-curve models of self-efficacy with leadership

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>PRV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-.05 (.10)</td>
<td>-.17 (.19)</td>
<td>-.14 (.20)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-.26 (.15)</td>
<td>-.27 (.15)</td>
<td>-.38** (.14)</td>
<td></td>
</tr>
<tr>
<td>Square-root time</td>
<td>-1.53 (.90)</td>
<td>-1.60 (.89)</td>
<td>-1.93* (.83)</td>
<td></td>
</tr>
<tr>
<td>Log time</td>
<td>6.08 (3.24)</td>
<td>6.31 (3.23)</td>
<td>7.80* (3.02)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>.08 (.16)</td>
<td>.01 (.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specialty</td>
<td>.06 (.16)</td>
<td>.07 (.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1 efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFL</td>
<td>.72** (.12)</td>
<td>.68** (.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.33** (.09)</td>
<td>.13** (.05)</td>
<td>.14** (.02)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>.122** (.02)</td>
<td>.122** (.02)</td>
<td>.104** (.01)</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05; ** p < .01

a p < Bonferroni correction

PRV = proportion reduction in variance

n.i. = not interpretable
Table 2-7: Growth-curve models of self-efficacy with experiences

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>PRV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-.05 (.10)</td>
<td>-.17 (.19)</td>
<td>-.15 (.18)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-.26 (.15)</td>
<td>-.27 (.15)</td>
<td>-.32* (.15)</td>
<td></td>
</tr>
<tr>
<td>Square-root time</td>
<td>-1.53 (.90)</td>
<td>-1.60 (.89)</td>
<td>-1.91* (.92)</td>
<td></td>
</tr>
<tr>
<td>Log time</td>
<td>6.08 (3.24)</td>
<td>6.31 (3.23)</td>
<td>7.47*</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>.08 (.16)</td>
<td>.02 (.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specialty</td>
<td>.06 (.16)</td>
<td>.08 (.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T&lt;sub&gt;1&lt;/sub&gt; efficacy</td>
<td>.72** (.12)</td>
<td>.63** (.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morbidity &amp; mortality experience</td>
<td></td>
<td>-19* (.07)</td>
<td>n.i.</td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact - frequency</td>
<td>.00 (.05)</td>
<td>n.i.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact - breadth</td>
<td>-.03 (.04)</td>
<td>n.i.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact - depth</td>
<td>.08* (.03)</td>
<td>n.i.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom</td>
<td>.02 (.05)</td>
<td>0.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.33** (.09)</td>
<td>.13* (.05)</td>
<td>.08 (.04)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>.122** (.02)</td>
<td>.122** (.02)</td>
<td>.125**</td>
<td></td>
</tr>
</tbody>
</table>

*<i>p < .05</i>; **<i>p <.01</i>

<sup>a</sup><i>p < Bonferroni correction</i>

PRV = proportion reduction in variance

*n.i.* = not interpretable
Table 2-8: Growth-curve models of daytime sleepiness with leadership

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 1a</th>
<th>Model 1b</th>
<th>Model 2</th>
<th>Model 3</th>
<th>PRV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.08 (.08)</td>
<td>-.06 (.17)</td>
<td>-.03 (.15)</td>
<td>-.04 (.15)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-.00 (.08)</td>
<td>-.43** (.16)</td>
<td>-.43** (.15)</td>
<td>-.30* (.15)</td>
<td></td>
</tr>
<tr>
<td>Log time</td>
<td>-.17 (.44)</td>
<td>2.39** (.86)</td>
<td>2.40** (.85)</td>
<td>1.70* (.84)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>.18 (.20)</td>
<td>.12 (.15)</td>
<td>.17 (.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specialty</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>T&lt;sub&gt;1&lt;/sub&gt; sleepiness</td>
<td>- .04 (.10)</td>
<td>- .06 (.10)</td>
<td>- .04 (.10)</td>
<td>- .06 (.10)</td>
<td></td>
</tr>
<tr>
<td>Sleep hours</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>TFL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.08* (.04) 0.0%</td>
</tr>
<tr>
<td>LF</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.11** (.03) 9.8%</td>
</tr>
<tr>
<td>AS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.08* (.05) 0.0%</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*gender</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Log(time)*gender</td>
<td>-3.32** (.99)</td>
<td>-3.28** (1.00)</td>
<td>-2.50* (.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.20** (.05)</td>
<td>.19** (.05)</td>
<td>.06* (.03)</td>
<td>.05* (.02)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>.00 (.00)</td>
<td>.01 (.00)</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>.095** (.05)</td>
<td>.084** (.01)</td>
<td>.081** (.01)</td>
<td>.075** (.01)</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05; **p < .01

<sup>a</sup> p < Bonferroni correction

PRV = proportion reduction in variance

<em>n.i.</em> = not interpretable
Table 2-9: Growth-curve models of daytime sleepiness with experiences

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 1</th>
<th>Model 2a</th>
<th>Model 2b</th>
<th>Model 3</th>
<th>PRV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.08 (.08)</td>
<td>-.06 (.17)</td>
<td>-.03 (.15)</td>
<td>.05 (.15)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-.00 (.08)</td>
<td>-.43** (.16)</td>
<td>-.43** (.15)</td>
<td>-.36* (.14)</td>
<td></td>
</tr>
<tr>
<td>Log time</td>
<td>-.17 (.44)</td>
<td>2.39** (.86)</td>
<td>2.40** (.85)</td>
<td>1.94* (.79)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>.18 (.20)</td>
<td>.12 (.15)</td>
<td>.11 (.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speciality</td>
<td>-.04 (.10)</td>
<td>-.12 (.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T&lt;sub&gt;1&lt;/sub&gt; sleepiness</td>
<td>.72** (.10)</td>
<td>.75** (.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep hours</td>
<td>.00 (.02)</td>
<td>.00 (.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morbidity &amp; mortality experience</td>
<td></td>
<td>.17**&lt;sup&gt;a&lt;/sup&gt; (.06)</td>
<td>10.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact frequency</td>
<td>.11**&lt;sup&gt;a&lt;/sup&gt; (.04)</td>
<td>8.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact breadth</td>
<td>-.04 (.04)</td>
<td>1.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beneficiary contact depth</td>
<td>-.06* (.03)</td>
<td></td>
<td>n.i.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom</td>
<td>.56** (.18)</td>
<td></td>
<td></td>
<td></td>
<td>n.i.</td>
</tr>
<tr>
<td><strong>Interactions</strong></td>
<td>-3.32** (.99)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*gender</td>
<td>.56** (.17)</td>
<td>.37* (.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(time)*gender</td>
<td>.19** (.05)</td>
<td>-3.28**</td>
<td>-2.31* (.91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td>.01 (.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.20** (.05)</td>
<td>-.01 (.01)</td>
<td>.06* (.03)</td>
<td>.07* (.03)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>.00 (.00)</td>
<td>.084** (.01)</td>
<td>.00 (.00)</td>
<td>.01 (.00)</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>.095** (.05)</td>
<td>.092** (.01)</td>
<td>.081** (.01)</td>
<td>.065** (.01)</td>
<td></td>
</tr>
</tbody>
</table>

*<i>p</i> < .05; **<i>p</i> < .01

<sup>a</sup><i>p</i> < Bonferroni correction

PRV = proportion reduction in variance

<i>n.i.</i> = not interpretable
Table 2-10: Sensitivity of covariance matrix on results of empathy and leadership model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th>AR(1)</th>
<th>CS&lt;sup&gt;1&lt;/sup&gt;</th>
<th>UN&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercepts</td>
<td>.18 (.15)</td>
<td>.19 (.15)</td>
<td>.18 (.15)</td>
<td>.18 (.13)</td>
</tr>
<tr>
<td>Time</td>
<td>-.15 (.15)</td>
<td>-.15 (.14)</td>
<td>-.15 (.15)</td>
<td>-.21 (.12)</td>
</tr>
<tr>
<td>Time^2</td>
<td>.16 (.09)</td>
<td>.16 (.09)</td>
<td>.16 (.09)</td>
<td>.18* (.07)</td>
</tr>
<tr>
<td>Time^3</td>
<td>-.03* (.01)</td>
<td>-.03* (.01)</td>
<td>-.03* (.01)</td>
<td>-.03** (.01)</td>
</tr>
<tr>
<td>Gender</td>
<td>-.14 (.13)</td>
<td>-.14 (.13)</td>
<td>-.14 (.13)</td>
<td>-.04 (.12)</td>
</tr>
<tr>
<td>Specialty</td>
<td>-.12 (.11)</td>
<td>-.12 (.11)</td>
<td>-.12 (.11)</td>
<td>-.14 (.11)</td>
</tr>
<tr>
<td>T&lt;sub&gt;1&lt;/sub&gt; Empathy</td>
<td>.61** (.09)</td>
<td>.62** (.09)</td>
<td>.61** (.09)</td>
<td>.60** (.09)</td>
</tr>
<tr>
<td>TFL</td>
<td>.04 (.05)</td>
<td>.05 (.05)</td>
<td>.04 (.05)</td>
<td>.05 (.04)</td>
</tr>
<tr>
<td>LF</td>
<td>-.08* (.03)</td>
<td>-.07* (.03)</td>
<td>-.08* (.03)</td>
<td>-.05* (.03)</td>
</tr>
<tr>
<td>AS</td>
<td>.04 (.05)</td>
<td>.04 (.06)</td>
<td>.04 (.05)</td>
<td>.03 (.04)</td>
</tr>
<tr>
<td>AIC</td>
<td>95.3</td>
<td>94.4</td>
<td>97.3</td>
<td>132.9</td>
</tr>
</tbody>
</table>

<sup>1</sup>Convergence not achieved  
*p < .05; **p < .01
Table 2-11: Sensitivity of covariance matrix on results of empathy and experiences model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th>AR(1)</th>
<th>CS¹</th>
<th>UN¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.17 (.16)</td>
<td>.18 (.16)</td>
<td>.17 (.16)</td>
<td>.19 (.15)</td>
</tr>
<tr>
<td>Time</td>
<td>-.18 (.16)</td>
<td>-.19 (.15)</td>
<td>-.18 (.16)</td>
<td>-.21 (.13)</td>
</tr>
<tr>
<td>Time^2</td>
<td>.16 (.10)</td>
<td>.16 (.09)</td>
<td>.16 (.10)</td>
<td>.17 (.08)</td>
</tr>
<tr>
<td>Time^3</td>
<td>-.03* (.02)</td>
<td>-.03* (.02)</td>
<td>-.03* (.02)</td>
<td>-.03* (.01)</td>
</tr>
<tr>
<td>Gender</td>
<td>-.06 (.13)</td>
<td>-.06 (.13)</td>
<td>-.06 (.13)</td>
<td>.01 (.12)</td>
</tr>
<tr>
<td>Specialty</td>
<td>-.14 (.13)</td>
<td>-.14 (.13)</td>
<td>-.14 (.13)</td>
<td>-.19 (.12)</td>
</tr>
<tr>
<td>T₁ Empathy</td>
<td>.59** (.10)</td>
<td>.59** (.10)</td>
<td>.59** (.10)</td>
<td>.58** (.09)</td>
</tr>
<tr>
<td>Morbidity &amp; mortality</td>
<td>-.05 (.07)</td>
<td>-.05 (.07)</td>
<td>-.05 (.07)</td>
<td>-.01 (.05)</td>
</tr>
<tr>
<td>BC- frequency</td>
<td>-.00 (.05)</td>
<td>-.02 (.05)</td>
<td>-.00 (.05)</td>
<td>-.00 (.04)</td>
</tr>
<tr>
<td>BC- breadth</td>
<td>.00 (.05)</td>
<td>.01 (.05)</td>
<td>.00 (.05)</td>
<td>.01 (.04)</td>
</tr>
<tr>
<td>BC - depth</td>
<td>.02 (.04)</td>
<td>.02 (.04)</td>
<td>.02 (.04)</td>
<td>-.01 (.03)</td>
</tr>
<tr>
<td>Boredom</td>
<td>-.04 (.05)</td>
<td>-.04 (.05)</td>
<td>-.04 (.05)</td>
<td>-.04 (.04)</td>
</tr>
<tr>
<td>AIC</td>
<td>104.8</td>
<td>104.3</td>
<td>106.8</td>
<td>149.2</td>
</tr>
</tbody>
</table>

¹Convergence not achieved
*p < .05; **p < .01
Table 2-12: Sensitivity of covariance matrix on results of self-efficacy and leadership model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th>AR(1)</th>
<th>CS(^1)</th>
<th>UN(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.14 (.20)</td>
<td>-.15 (.20)</td>
<td>-.14 (.20)</td>
<td>-.03 (.16)</td>
</tr>
<tr>
<td>Time</td>
<td>-.38** (.14)</td>
<td>-.39** (.13)</td>
<td>-.38** (.14)</td>
<td>-.35** (.11)</td>
</tr>
<tr>
<td>Square-root time</td>
<td>-1.93* (.83)</td>
<td>-2.01* (.82)</td>
<td>-1.93* (.83)</td>
<td>-1.88* (.65)</td>
</tr>
<tr>
<td>Log time</td>
<td>7.80* (3.02)</td>
<td>8.10** (2.93)</td>
<td>7.80* (3.02)</td>
<td>7.53* (2.31)</td>
</tr>
<tr>
<td>Gender</td>
<td>.01 (.17)</td>
<td>.02 (.17)</td>
<td>.01 (.17)</td>
<td>-.05 (.14)</td>
</tr>
<tr>
<td>Specialty</td>
<td>.07 (.16)</td>
<td>.07 (.16)</td>
<td>.07 (.16)</td>
<td>-.02 (.14)</td>
</tr>
<tr>
<td>T1 Efficacy</td>
<td>.68** (.12)</td>
<td>.68** (.13)</td>
<td>.68** (.12)</td>
<td>.70** (.10)</td>
</tr>
<tr>
<td>TFL</td>
<td>.16** (.05)</td>
<td>.16** (.05)</td>
<td>.16** (.05)</td>
<td>.16** (.04)</td>
</tr>
<tr>
<td>LF</td>
<td>-.03 (.03)</td>
<td>-.04 (.03)</td>
<td>-.03 (.03)</td>
<td>-.01 (.03)</td>
</tr>
<tr>
<td>AS</td>
<td>.01 (.06)</td>
<td>.02 (.06)</td>
<td>.01 (.06)</td>
<td>-.03 (.05)</td>
</tr>
<tr>
<td>AIC</td>
<td>202.8</td>
<td>202.9</td>
<td>204.8</td>
<td>268.8</td>
</tr>
</tbody>
</table>

\(^1\)Convergence not achieved

*\(p < .05\); **\(p <.01\)
Table 2-13: Sensitivity of covariance matrix on results of self-efficacy and experiences model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th>AR(1)</th>
<th>CS(^1)</th>
<th>UN(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-.15 (.18)</td>
<td>-.17 (.18)</td>
<td>-.15 (.18)</td>
<td>-.08 (.15)</td>
</tr>
<tr>
<td>Time</td>
<td>-.32* (.15)</td>
<td>-.34* (.14)</td>
<td>-.32* (.15)</td>
<td>-.28* (.11)</td>
</tr>
<tr>
<td>Square-root time</td>
<td>-1.91* (.92)</td>
<td>-2.06* (.91)</td>
<td>-1.91* (.92)</td>
<td>-1.70* (.63)</td>
</tr>
<tr>
<td>Log time</td>
<td>7.47* (3.33)</td>
<td>8.06* (3.21)</td>
<td>7.47* (3.33)</td>
<td>6.79* (2.26)</td>
</tr>
<tr>
<td>Gender</td>
<td>.02 (.15)</td>
<td>.03 (.15)</td>
<td>.02 (.15)</td>
<td>-.06 (.13)</td>
</tr>
<tr>
<td>Specialty</td>
<td>.08 (.15)</td>
<td>.08 (.15)</td>
<td>.07 (.15)</td>
<td>.03 (.13)</td>
</tr>
<tr>
<td>T(_1) Efficacy</td>
<td>.63** (.11)</td>
<td>.62** (.12)</td>
<td>.63** (.11)</td>
<td>.68** (.10)</td>
</tr>
<tr>
<td>Morbidity &amp; mortality</td>
<td>-1.19* (.07)</td>
<td>-.21* (.08)</td>
<td>-.19* (.08)</td>
<td>-.18** (.07)</td>
</tr>
<tr>
<td>BC- frequency</td>
<td>.00 (.05)</td>
<td>-.00 (.05)</td>
<td>.00 (.05)</td>
<td>.03 (.04)</td>
</tr>
<tr>
<td>BC- breadth</td>
<td>-.03 (.04)</td>
<td>-.02 (.04)</td>
<td>-.03 (.04)</td>
<td>-.07* (.04)</td>
</tr>
<tr>
<td>BC- depth</td>
<td>.08* (.03)</td>
<td>.07* (.03)</td>
<td>.08* (.03)</td>
<td>.07** (.03)</td>
</tr>
<tr>
<td>Boredom</td>
<td>.02 (.05)</td>
<td>.04 (.05)</td>
<td>.02 (.05)</td>
<td>.03 (.04)</td>
</tr>
<tr>
<td>AIC</td>
<td>214.9</td>
<td>214.0</td>
<td>216.9</td>
<td>297.7</td>
</tr>
</tbody>
</table>

\(^1\)Convergence not achieved

\(^*\)p < .05; \(^**\)p < .01
Table 2-14: Sensitivity of covariance matrix on results of daytime sleepiness and leadership model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th>AR(1) $^{\dagger}$</th>
<th>CS $^{\dagger}$</th>
<th>UN $^{\dagger}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-.04 (.15)</td>
<td>-.03 (.14)</td>
<td>-.04 (.15)</td>
<td>-.04 (.15)</td>
</tr>
<tr>
<td>Time</td>
<td>-.30* (.15)</td>
<td>-.31* (.15)</td>
<td>-.30* (.15)</td>
<td>-.30* (.15)</td>
</tr>
<tr>
<td>Log time</td>
<td>1.70* (.84)</td>
<td>1.70* (.84)</td>
<td>1.70* (.84)</td>
<td>1.70* (.84)</td>
</tr>
<tr>
<td>Gender</td>
<td>.17 (.15)</td>
<td>.16 (.14)</td>
<td>.17 (.15)</td>
<td>.17 (.15)</td>
</tr>
<tr>
<td>Specialty</td>
<td>-.06 (.10)</td>
<td>-.06 (.10)</td>
<td>-.06 (.10)</td>
<td>-.06 (.10)</td>
</tr>
<tr>
<td>T1 Sleepiness</td>
<td>.71** (.10)</td>
<td>.71** (.10)</td>
<td>.71** (.10)</td>
<td>.71** (.10)</td>
</tr>
<tr>
<td>Sleep hours</td>
<td>.00 (.02)</td>
<td>.00 (.02)</td>
<td>.01 (.02)</td>
<td>.00 (.02)</td>
</tr>
<tr>
<td>TFL</td>
<td>.08* (.04)</td>
<td>.08* (.04)</td>
<td>.08* (.04)</td>
<td>.08* (.04)</td>
</tr>
<tr>
<td>LF</td>
<td>.11** (.03)</td>
<td>.11** (.03)</td>
<td>.11** (.03)</td>
<td>.11** (.03)</td>
</tr>
<tr>
<td>AS</td>
<td>.08 (.05)</td>
<td>.08 (.05)</td>
<td>.08 (.05)</td>
<td>.08 (.05)</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*gender</td>
<td>.41* (.17)</td>
<td>.41* (.17)</td>
<td>.41* (.17)</td>
<td>.41* (.17)</td>
</tr>
<tr>
<td>Log(time)*gender</td>
<td>-2.50* (.96)</td>
<td>-2.49* (.96)</td>
<td>-2.50** (.96)</td>
<td>-2.50* (.96)</td>
</tr>
<tr>
<td>AIC</td>
<td>145.3</td>
<td>144.8</td>
<td>147.3</td>
<td>218.0</td>
</tr>
</tbody>
</table>

$^{\dagger}$Convergence not achieved

*p < .05; **p < .01
Table 2-15: Sensitivity of covariance matrix on results of daytime sleepiness and experiences model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Final Model</th>
<th>AR(1)</th>
<th>CS†</th>
<th>UN†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.05 (.15)</td>
<td>.06 (.15)</td>
<td>.05 (.15)</td>
<td>.16 (.12)</td>
</tr>
<tr>
<td>Time</td>
<td>-.36* (.14)</td>
<td>-.37* (.15)</td>
<td>-.36* (.14)</td>
<td>-.37** (.13)</td>
</tr>
<tr>
<td>Log time</td>
<td>1.94* (.79)</td>
<td>2.04* (.84)</td>
<td>1.94* (.79)</td>
<td>2.05** (.62)</td>
</tr>
<tr>
<td>Gender</td>
<td>.11 (.15)</td>
<td>.09 (.15)</td>
<td>.11 (.15)</td>
<td>.04 (.12)</td>
</tr>
<tr>
<td>Specialty</td>
<td>-.12 (.10)</td>
<td>-.13 (.10)</td>
<td>-.12 (.10)</td>
<td>-.25** (.10)</td>
</tr>
<tr>
<td>T1 Sleepiness</td>
<td>.75** (.10)</td>
<td>.74** (.10)</td>
<td>.75** (.10)</td>
<td>.79** (.10)</td>
</tr>
<tr>
<td>Sleep hours</td>
<td>.00 (.02)</td>
<td>.00 (.02)</td>
<td>.00 (.02)</td>
<td>.00 (.02)</td>
</tr>
<tr>
<td>Morbidity &amp; mortality</td>
<td>.17** (.06)</td>
<td>.15* (.06)</td>
<td>.17** (.06)</td>
<td>.18** (.05)</td>
</tr>
<tr>
<td>BC- frequency</td>
<td>.11** (.04)</td>
<td>.12** (.04)</td>
<td>.11** (.04)</td>
<td>.13** (.03)</td>
</tr>
<tr>
<td>BC- breadth</td>
<td>-.04 (.04)</td>
<td>-.04 (.04)</td>
<td>-.04 (.04)</td>
<td>-.05 (.03)</td>
</tr>
<tr>
<td>BC- depth</td>
<td>-.06* (.03)</td>
<td>-.06* (.03)</td>
<td>-.06* (.03)</td>
<td>-.05* (.02)</td>
</tr>
<tr>
<td>Boredom</td>
<td>.05 (.04)</td>
<td>.05 (.04)</td>
<td>.05 (.04)</td>
<td>.04 (.03)</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*gender</td>
<td>.37* (.17)</td>
<td>.39* (.18)</td>
<td>.37* (.17)</td>
<td>.37* (.15)</td>
</tr>
<tr>
<td>Log(time)*gender</td>
<td>-2.31* (.91)</td>
<td>-2.40* (.96)</td>
<td>-2.31* (.91)</td>
<td>-2.39** (.72)</td>
</tr>
<tr>
<td>AIC</td>
<td>147.2</td>
<td>146.8</td>
<td>149.2</td>
<td>204.1</td>
</tr>
</tbody>
</table>

†Convergence not achieved
*p < .05; **p < .01
Table 2-16: Post-hoc analysis examining leadership and daytime sleepiness

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-.04 (.15)</td>
<td>-.08 (.15)</td>
</tr>
<tr>
<td>Time</td>
<td>-.30* (.15)</td>
<td>-.28 (.15)</td>
</tr>
<tr>
<td>Log time</td>
<td>1.70* (.84)</td>
<td>1.61 (.81)</td>
</tr>
<tr>
<td>Gender</td>
<td>.17 (.15)</td>
<td>.21 (.15)</td>
</tr>
<tr>
<td>Specialty</td>
<td>-.06 (.10)</td>
<td>-.06 (.10)</td>
</tr>
<tr>
<td>T1 sleepiness</td>
<td>.71** (.10)</td>
<td>.71** (.10)</td>
</tr>
<tr>
<td>Sleep hours</td>
<td>.00 (.02)</td>
<td>.01 (.02)</td>
</tr>
<tr>
<td>TFL</td>
<td>.08* (.04)</td>
<td>-.07 (.07)</td>
</tr>
<tr>
<td>LF</td>
<td>.11** (.03)</td>
<td>.11** (.03)</td>
</tr>
<tr>
<td>AS</td>
<td>.08 (.05)</td>
<td>.04 (.05)</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*gender</td>
<td>.41* (.17)</td>
<td>.37* (.17)</td>
</tr>
<tr>
<td>Log(time)*gender</td>
<td>-2.50* (.96)</td>
<td>-2.38* (.93)</td>
</tr>
<tr>
<td>Gender*TFL</td>
<td></td>
<td>.19* (.08)</td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.05* (.02)</td>
<td>.06* (.03)</td>
</tr>
<tr>
<td>Time</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
</tr>
<tr>
<td>Residual</td>
<td>.075** (.01)</td>
<td>.070** (.01)</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01
Figure 2-1: Mean empathy across study periods
Figure 2-2: Mean self-efficacy across study periods
Figure 2-3: Mean daytime sleepiness across study periods
Figure 2-4: Mean daytime sleepiness across study periods by gender
Figure 2-5: Results of empathy growth-curve analysis
Figure 2-6: Results of daytime sleepiness growth-curve analysis by gender

Note: The female curve is not significantly different from zero.
Figure 2-7: Results of growth-curve analysis of self-efficacy in leadership model
Figure 2-8: Results of growth-curve analysis of self-efficacy in experiences model
Chapter 3

Scalpel! Leadership Behaviours and their Consequences in and Beyond the Operating Room

Abstract

Leadership research has often been conducted with little to no regard for context. This study attempts to remedy this by examining leadership in an extreme context—in which “risks of severe physical, psychological or material consequences [e.g., physical harm, devastation or destruction] to organizational members or their constituents exist” (Hannah et al., 2009, p. 897), namely the context of the operating room. I predicted that transformational leadership, laissez-faire leadership, and over-controlling leadership would influence short and long-term surgical outcomes directly, and indirectly through the meditational variables of psychological safety and boredom. Further, I predicted that the meditational paths would be moderated non-linearly by surgical complexity. Using a combination of observational, survey, and health record data, hypotheses were tested on 150 surgeries spanning specialties. Results failed to support the mediational hypotheses, but did yield significant results for the direct impact of leadership in operating room: laissez-faire leadership and over-controlling leadership predicted blood loss, laissez-faire leadership and over-controlling leadership predicted complications at discharge, and transformational leadership and over-controlling leadership predicted complications at follow-up. Surgical complexity consistently moderated these effects.
3.1 Introduction

In a context in which successful performance could mean life or death, studying the potential impact of leadership behaviours and the mechanisms through which they exert their influence is critical. Study 2 of my dissertation examines the influence of leadership behaviours in the operating room on both immediate and longer-term surgical outcomes. By examining both proximal outcomes (e.g., errors, blood loss), as well as distal outcomes (e.g., complications during recovery, complications at a month follow-up), a more complete picture of the importance of leadership during surgery can be obtained. Furthermore, I also examine whether leadership exerts its effects through psychological safety and boredom, and whether surgical complexity moderates these effects.

By studying the leadership behaviours of surgeons in the operating room, I will begin to uncover the dynamics of leadership in the under-researched area of extreme contexts (a context in which “risks of severe physical, psychological or material consequences [e.g., physical harm, devastation or destruction] to organizational members or their constituents exist” Hannah, Uhl-Bien, Avolio, & Cavaretta, 2009, p. 897). First however, I will propose a model of leadership in extreme contexts, drawing on research from traditional contexts on performance, safety, and errors, and limited research in other extreme contexts. I then extend this model, incorporating two mediators (i.e. psychological safety, boredom) and one moderator (i.e. complexity) of the leadership-surgical outcomes relationships. This model is subjected to empirical testing in the operating room – with real surgeons, real patients and real outcomes. I will conclude with a discussion of the results, possible future avenues for research, and practical implications.

Leadership During Surgery
The goal of my research is to examine the effects of leadership during surgery on proximal and distal patient outcomes. Proximal outcomes include errors, and blood loss, and distal outcomes include complications prior to hospital discharge (referred to as “complications at discharge” in this paper), and longer-term complications within the month following discharge (referred to as “complications at follow-up” in this paper). To enhance ecological validity, these measures of surgical outcomes were selected by practicing surgeons and anesthesiologists as important metrics of successful surgery.

As discussed in the introductory chapter of this dissertation, an extreme context is comprised of five factors, namely the temporal state of the organization (i.e. preparation, response, recovery), the magnitude of potential consequences, the probability of those consequences occurring, the proximity of organizational members (both physical and psychological distance), and the form of threat (i.e. physical, psychological, material). These factors vary both between and within extreme contexts, but what distinguishes extreme from traditional contexts is the form of threat, magnitude of potential consequences, and the probability of those consequences occurring. An extreme context involves outcomes of more significant magnitude, with a greater probability of occurring.

Two factors render surgery an appropriate setting in which to study leadership in extreme contexts, namely the training of surgeons, and the formal roles in the operating room. First, unlike other extreme contexts such as the military where leadership training may be standard, the surgeon is a nominal leader, with no formal leadership training, so a wide variance of behaviours can be expected. Second, surgeries involve highly prescribed team member roles and behaviors (Galinsky, Chou, Halevy, & Van Kleef, 2012). These “strong situations” (e.g., Cooper & Withey, 2009) can be expected to enhance team effectiveness (e.g., Mathieu, Heffner, Goodwin,
Salas & Cannon-Bowers, 2000), but also leave less flexibility for team member initiative (e.g., Apker, Propp, & Zabava Ford, 2005). Along with these highly prescribed roles which characterize extreme contexts such as military, law enforcement, or fire-fighting, surgeries invariably involve transient teams – teams that come together for a brief task, with little to no likelihood of that same team performing together again (Harvey & Butcher, 1998). Trust and performance challenges may be higher in transient teams versus traditional, stable teams. I believe that together, this means that extreme contexts will provide a more stringent test of leadership, potentially increasing (a) the importance of any lessons learned, and (b) the generalizability of any findings.

**Leadership Behaviours: Transformational, Laissez-faire, & Over-controlling**

Before a discussion and hypothesis development of the impacts of leadership in extreme contexts and the development of a theoretical model, a discussion of the specific behaviours of interest is needed. I begin by describing the leadership behaviours of interest in this research: transformational leadership, laissez-faire leadership, and over-controlling leadership. I then offer a refinement of the conceptualization of leadership, namely the need to consider consistency of leadership behaviours in relation to other leadership behaviours, as opposed to the dominant model, which examines leadership behaviors in isolation from each other.

**Transformational Leadership**

Transformational leadership is the most widely-studied theory of leadership, associated with numerous positive organizational outcomes, such as organizational effectiveness, financial performance, and employee attitudes (Barling, Christie & Hoption, 2010). Transformational leadership is comprised of four categories of behaviours: Intellectual stimulation, individualized consideration, inspirational motivation, and idealized influence (e.g., Bass, 1990). Intellectual
stimulation involves behaviours meant to challenge employees’ assumptions and help them see their work from a different perspective. Individualized consideration involves behaviours that demonstrate the leader understands the employees as individuals, with different sets of skills and needs. These behaviours include showing empathy and trust towards employees. Inspirational motivation involves challenging employees to strive for and achieve more than they thought possible, by helping in setting high and realistic goals. Finally, idealized influence involves behaviours that demonstrate a leader’s focus on the employees and long-term success of the group, by showing a collective sense of mission, ethics and integrity. While each of these categories are described as separate behaviours, transformational leadership is most-often studied as a unidimensional construct (Barling et al., 2010)

**Laissez-faire Leadership**

Laissez-faire leadership is characterized by non-leadership. This type of leader is avoidant and unwilling to take ownership or responsibility (Barling, Christie & Hoption, 2010). More specifically, laissez-faire leadership is characterized by providing neither praise nor feedback for positive performance, nor punishment or feedback for negative performance (Hinkin & Schreisheim, 2008). Furthermore, in this lack of interest and avoidance, laissez-faire leadership has been described as not only a lack of leadership, but a destructive style of leadership in terms of the outcomes it affects (e.g., bullying, interpersonal conflict; Skogstad, Einarsen, Torsheim, Aasland & Hetland, 2007).

**Over-controlling Leadership**

Over-controlling leadership is a widely recognized behaviour that many leaders exhibit. First introduced by Dupre and Barling (2006) as “supervisory control over work performance”, over-controlling leadership comprises behaviours that constrain or limit subordinates’ control in
performing their job. Behaviours include exerting control over employees’ time (e.g., when they can take breaks, how fast they have to work), and how they do their jobs (e.g., exerting influence over decisions, constantly watching for errors). Practitioners or followers might refer to this type of leader as a “micromanager”.

While there is virtually no research specifically examining over-controlling leadership, over-controlling behaviours have been studied as part of larger theories of negative leadership, including destructive leadership (e.g., Einarsen, Aasland & Skogstad, 2007; Schyns & Schilling, 2013) and toxic leadership (e.g., Steele, 2011), both of which make mention of leader rigidity, obstructiveness, autocracy, physical behaviours, and micromanagement. However none of these types of leadership have gained much empirical attention. Furthermore, these types of leadership include many different behaviours (e.g., behaviours that would fall under abusive supervision, over-controlling, and authoritarian), despite the theoretically and empirically distinct consequences of such behaviours. Finally, Schyns and Schilling (2013) included leaders’ destructive intentions in their definition of destructive leadership; over-controlling leadership does not consider intent.

Further information on the construct can be gleaned from both the wider literature on control in the workplace, and from literature on the effects of over-controlling parenting. Within the study of job design (Hackman, Oldham, Janson & Purdy, 1975), autonomy reflects the amount of flexibility and control that employees have to decide when and how they do their work. Leaders have the authority to change the amount of control and flexibility that employees have; an over-controlling leader limits flexibility. Similarly, when examining parental control, Grolnick and Pomerantz (2009) distinguished between over-controlling parenting, characterized by pressure, intrusion, and domination, and simply providing parental guidance. This is an
important distinction when discussing the relationship between a leader and employee as well, as findings from the parenting literature suggest that over-controlling parenting is associated with children delinquent behaviour problems (e.g., Pettit, Laird, Dodge, Bates, & Criss, 2001), aggression and hyperactivity (e.g., Stormshak, Bierman, McMahon, & Lengua, 2000), and anxiety (e.g., Van Der Bruggen, Stams, & Bogels, 2008), while simple guidance (or structuring) behaviour is associated with reduced child misbehaviour (e.g., Fletcher, Steinberg, & Williams-Wheeler, 2004), higher academic competence (e.g., Skinner, Johnson, & Schneider-Wheeler, 2004), and less risk behaviours (e.g., drug use; DiClemente et al., 2001). Over-controlling leadership does not include guidance or structuring behaviours (e.g., contingent reward leadership behaviours; e.g., Podsakoff, Todor, & Skov, 1982).

Extending all of this to leadership behaviours, over-controlling leadership creates an environment low in flexibility, by placing restrictions on when employees perform their job, and closely monitoring how they perform their job.

**Refining the Conceptualization of Leadership: Consistency**

Historically, leadership behaviours have been considered in isolation. Leaders are measured (almost always with scales) by judgments on how often or how likely they are to behave a certain way, whether in terms of transformational, abusive, or charismatic behaviors, for example. Leaders are given a score to reflect, for example, how transformational they are, and then these scores are used in research to test theories. This conceptualization of leadership has allowed researchers to make advances in leadership theory, but I believe a significant piece of the puzzle is missing, namely the consistency with which such behaviors are expressed by the same leader. It is possible for a single leader to exhibit multiple different leadership behaviours, and capturing this is important. For example, transformational leadership and laissez-faire
leadership, two seemingly “opposite” behaviors, have been negatively correlated in some cases (e.g., $r = -0.24$; Furtner, Baldegger & Rauthmann, 2013), and positively correlated in others (e.g., $r = 0.30$; Arnold, Connelly, Walsh, Martin Ginis, 2015). This suggests that leadership behaviours are neither categorical nor mutually-exclusive: Leaders are not either transformational or laissez-faire, but likely a combination of each. Capturing these inconsistencies in behaviours could more accurately reflect the true nature of leadership.

The notion of measuring consistency (or inconsistency) of leadership has been discussed by Hannah, Sumanth, Lester and Cavarretta (2014) and Mullen, Kelloway and Teed (2011). In a review of “new genre” positive leadership behaviours (e.g., transformational leadership, authentic leadership), Hannah et al., (2014) concluded that in order to influence followers, leaders must gain and maintain credibility, and this is done by acting consistently. Leaders have minimal leeway to act contrary to the positive behaviours that followers expect, before leaders begin to lose positive influence. This suggests that the more consistently that leaders behave in a positive manner (e.g., transformational), the better the outcomes will be. Mullen et al. (2011) also recognized the need to study inconsistency, and suggested that positive leadership behaviours can be undermined as leaders also exhibit negative behaviours. The authors examined the effects of inconsistent leadership, investigating the interactive effects of safety-specific transformational leadership and safety-specific passive leadership. Their results demonstrated empirically that any positive effects of transformational leadership were attenuated as passive leadership increased.

My study introduces the notion of consistency of leadership, accounting for the fact that leadership behaviours are not perfectly consistent. Consistency in this study is operationalized as the proportion of any surgeons’ leadership behaviors that belong to any of the three leadership
behaviours studied (i.e. transformational, laissez-faire, over-controlling)\textsuperscript{11}. I believe that this allows a more accurate representation of leadership, but this does not change the way in which leadership is discussed or the hypotheses are developed, and as such all hypotheses refer to the consistency of leadership behaviours affecting the outcomes.

**Leadership During Surgery and Surgical Outcomes**

In this section, I develop hypotheses about the impact of transformational, laissez-faire, and over-controlling leadership on proximal (i.e., errors, blood loss) and distal surgical outcomes (i.e., complications at discharge and one-month follow-up). To do so, I will draw on research in traditional organizations, and limited research in other extreme contexts.

**Performance, Safety, & Errors**

Due to the limited research on how leadership affects surgical outcomes, to develop hypotheses on the effect of leadership behaviours on surgical outcomes, I will draw upon three relevant areas of research in other organizations (both traditional and extreme) in my development of the hypotheses for the extreme context of the operating room. Below I briefly introduce the areas of performance, safety behaviours, and errors. The goal in the following subsections is to simply introduce the nature of performance, safety and errors – a discussion of how they relate to leadership follows later. None of these areas parallel surgical outcomes, but point to the potential effects of leadership on surgical outcomes.

**Performance.** Performance can be separated into task performance and contextual performance (Motowildo & Van Scotter, 1994). Task performance involves activities that are directly related to the execution and maintenance of the organization’s technical core, often stipulated in an employee’s formal job description. In contrast, contextual performance is not

\textsuperscript{11} Please note that this is not consistency across time and surgeries, but rather within-person, within-surgery across the three different behaviours.
directly related to the technical core, but instead consists of activities that support the broader organizational environment (e.g., cooperation, organizational citizenship behaviours). However both types of performance uniquely contribute to overall organizational performance (Motowildo & Van Scotter, 1994).

Motowildo and Van Scotter (1994) identify “performing surgery in a hospital” as a form of task performance. In contrast, they identify categories of contextual behavior as (a) helping with extra-role activities, (b) persistence and extra enthusiasm, (c) helping and cooperation, (d) following rules and procedures even when they are inconvenient, and (e) supporting organizational objectives. The development of hypotheses relating leadership behaviours to surgical outcomes will draw in part on the impact of leadership on task performance.

**Safety.** A second relevant area of organizational research is that of safety. Safety could be considered a sub-component of task performance, as the safe execution of company objectives is core to organizations’ goals. This is especially true in the operating room, where keeping all the team members, and the patient, safe is a top priority. Safety could also be considered contextual performance, as Motowildo and Van Scotter (1994) identify following rules and procedures even when inconvenient as a category of contextual performance. However, it is important to make the distinction that the surgical outcomes studied in this paper should not be considered as safety performance, but instead, indicators of safety performance.

Safety in traditional organizations is often reflected in safety outcomes such as accidents and injuries, which is closer to the surgical outcomes studied in the current research (e.g., errors, blood loss). Safety is often also conceptualized as a positive safety climate (the perceptions of an organization’s safety priorities). Safety climate is often used as a mediator, as it is a predictor of workplace accidents and injuries (e.g., Barling, Loughlin, & Kelloway, 2002; Zohar, 2000).
There is rich body of research concerning safety, which is especially relevant to this study: Traditional organizations can face situations and events similar to those in extreme contexts, where employee (and constituent) health and safety is on the line. Employees in traditional organizations face the possibility of injury and death. What separates traditional organizations from an extreme context is the frequency and probability of extreme events occurring, as employees and constituents (or patients in this case) in extreme contexts face a higher frequency and probability of these events. Applying evidence of how leadership impacts safety in these instances provides evidence of leadership’s effects in more turbulent environments.

**Errors.** Finally, the study of errors in organizations is relevant to the current study, especially the proximal occurrence of errors, but also the occurrence of blood loss (as excess blood loss could be an indication of an error), and the distal outcomes of complications (as a complication during recovery can be a sign that an error occurred during the procedure). Understanding what is known about errors in other organizations will help inform the potential effects of leadership on surgical outcomes.

An error is defined as an action performed by an individual that threatens the achievement of a goal (Bauer & Mulder, 2007), and is considered without regard to whether it had a negative consequence of not (Eubanks & Mumford, 2010). The error could be detected and remedied, could have no significant impact, or could result in a negative outcome, or what in the medical arena is called an adverse event – patient injury resulting from medical (mis)management rather than a function of the patient’s disease or pre-existing condition (Zhang, Patel, Johnson, & Shortliffe, 2004).
Surgical teams have started to attract the attention of organizational and medical researchers, with a specific focus on patient safety and medical errors. In the last decade, errors have become a frequent topic in the medical field. In a study at the turn of the millennium, it was estimated that nearly 100,000 patients in the United States die each year due to one or more errors in the health system, and that this is the eighth leading cause of death in the United States (Henry, 2000). More recently it was estimated that approximately 24,000 Canadians die each year from medical errors (Giese, 2012). These statistics do not capture the number of near misses (where death almost occurred), or the reduced quality of life as a result of medical errors. These statistics have sparked an increase in attention of the medical community (not to mention the whole country) on the causes of medical errors, and the potential ways to prevent them.

Vashdi, Bamberger and Erez (2014) investigated the use of team reflexivity (i.e. briefing prior to surgery and debriefing after surgery) as a possible way to prevent errors, and increase learning from errors. In moderately complex surgeries, briefing and debriefing activities were associated with fewer errors. However these strategies were not effective for error prevention in lower or higher complexity surgeries. Perhaps leadership could fill this gap. Evidence will be drawn from this literature to inform hypotheses on the potential effects of leadership on surgical errors, and by extension, the indirect effects of leadership on the remaining surgical outcomes.

**Translating Findings into Extreme Contexts**

Findings from traditional contexts can be extended to extreme contexts for the purposes of hypothesis development. When translating findings from traditional to extreme contexts, two factors present in extreme contexts could differentially impact any effects of leadership namely, environmental uncertainty, and strong situations.
Uncertainty and ambiguity. Uncertainty and ambiguity, which vary between and within extreme contexts (Baran & Scott, 2010), could affect the effectiveness of leadership. I predict that leadership may be more important under uncertain and ambiguous conditions (versus stable and predictable conditions), as followers need more direction, sense-making, and adaptation (Baran & Scott, 2010).

Evidence from research on environmental uncertainty support the notion that leadership is more important under conditions of greater uncertainty. Environments rated as highly uncertain are those viewed as “highly risky, as contexts in which a few erroneous decisions could result in severe trouble and possibly put the survival of the organization at risk” (Waldman, Ramirez, House & Puranam, 2001, pg. 136), and the relationship between charismatic leadership and company performance is highest under conditions of high environmental uncertainty (Hoogh et al., 2004; Waldman, et al., 2001). Similar moderating effects were found for the effects of intellectual stimulation on firm performance (Waldman, Javidan & Varella, 2004). In both these cases, zero (e.g., Waldman et al, 2001) or weaker relationships (e.g., Hoogh et al., 2004) emerged in less ambiguous situations.

Strong situations. The notion of strong and weak situations was developed to explain when personality or situations would have stronger effects on behaviour. Strong situations are “those in which everyone knows what to do and why and how to do it” (Cooper & Withey, 2009, pg. 63). One recent study showed that personality variables are stronger predictors of job performance under weak circumstances (e.g., Judge & Zapata, 2015). The operating room is an example of a strong situation, as team members have prescribed roles, and surgeries often have tightly developed procedures. Extreme contexts in general reflect strong situations, and the strong situation theory has been applied to explain phenomena within them (e.g., Bearman,
As Hannah et al. (2009) explained, people are expected to be better prepared for extreme events, and are more closely directed by policies and procedures designed to improve their response. A defining characteristic of extreme contexts (e.g., firefighting, law enforcement) is that they have clear procedures so that people know what they are expected to do, how to do it, and why.

Strong situations could attenuate the need for, and effects of leadership under such conditions. As mentioned earlier, strong situations are positively related to team effectiveness (e.g., Mathieu et al., 2000), but also reduce variability of individual behaviours (e.g., Apker et al., 2005). I predict that leadership will exert a greater influence on behaviour under weak situations. In situations in which roles and procedures are highly prescribed, there may be less opportunity for leadership to have a meaningful impact on followers. However, Cooper and Withey (2009) had difficulty finding convincing evidence of the strong situation hypothesis.

In summary, I believe that leadership will matter under conditions of uncertainty and ambiguity, and potentially matter more in weak situations. Thus I conclude any effects of leadership in extreme contexts will be no weaker than those yielded in traditional contexts.

**Transformational Leadership and Surgical Outcomes**

Surgeons’ transformational leadership behaviours are expected to positively influence both proximal and distal surgical outcomes, achieving fewer errors, less blood loss, and fewer complications at both discharge and long-term follow-up. Transformational leadership elicits such positive effects by changing the way that followers view themselves, their relationships, and their work (Barling, 2014).

By showing individualized consideration of the needs and abilities of each team member, and by using inspirational motivation, I expect that transformational leadership will affect the
way that surgical team members view themselves. Demonstrations that transformational leaders believe in the capabilities of team members (e.g., giving residents the opportunity to perform a new procedure, asking for input from nursing staff) are likely to elicit greater levels of confidence in each team member’s ability to achieve optimal surgical outcomes. This individual efficacy is expected to contribute to actual achievement of positive surgical outcomes (e.g., Stajkovic & Luthans, 1998).

Transformational leadership also changes the way that team members view their relationships with each other by reducing perceived social distances (e.g., Antonakis & Atwater, 2002), and developing a safe, trusting environment. Reducing perceived social distance may be especially important in the operating room, where hierarchies are strongly ingrained in the culture (e.g., Kellogg, 2012). By recognizing the unique contributions of each team member, and the need for all team members to play their part in the success of the team, perceived social distance is diminished, which may increase the power of a leader (e.g., Antonakis & Atwater, 2002), and increase effort toward goals (e.g., Avolio, 1999). Transformational leadership also fosters trust in the leader (e.g., Podsakoff, MacKenzie, Moorman, & Fetter, 1990) by using individualized consideration and idealized influence. Trust in the leader increases leaders’ influence, and develops a team environment where members are more collaborative and creative (e.g., Barczak, Lassk, & Mulki, 2010), further improving surgical outcomes. Together, the lower social distance and higher trust in leadership provides an environment where members work together and collaborate to yield optimal and creative solutions, thus enhancing surgical outcomes.

Finally, and perhaps most importantly, transformational leadership is associated with the way in which team members view their work. As medical caregivers progress in their careers,
there is a tendency to empathize less with patients, and respond more habitually (e.g., Marcum, 2005). Through idealized influence, transformational leadership refocuses followers on the greater purpose of their work: Giving patients a better quality of life through solving their medical problems. By changing the way followers view their work, followers may be less likely to take shortcuts, and be less focused on the daily operational constraints of the operating room (e.g., following the schedule) and more attentive to doing the right things to promote full recovery of the patient. Furthermore, by viewing each patient as a unique individual, transformational leadership stimulates team members to find new, or at least better-suited, methods for treating patients, and challenge their common assumptions and biases of patient care (e.g., Groopman & Prichard, 2007). The resulting motivation is expected to improve both proximal and distal surgical outcomes.

While there is no direct evidence of the effect of transformational leadership on the surgical outcomes of interest, support for the impact of transformational leadership on these surgical outcomes comes from research in traditional contexts in relation to performance, and safety climate. There is also limited evidence of the relationship in extreme contexts as well. These are discussed in detail below.

**Performance.** While performance is not directly transferrable to surgical outcomes (as performance is a behaviour, and surgical outcomes are consequences of behaviour), the effect of leadership on performance provides clues to the impact of leadership on surgical outcomes. In traditional contexts, transformational leadership is consistently positively associated with individual subordinate performance, team performance, and organizational performance (Barling et al., 2010; Wang, Oh, Courtright, & Colbert, 2011). Transformational leaders often achieve greater performance by increasing the intrinsic value of performance to employees, eliciting
greater levels of motivation (Seibert, Wang & Courtright, 2011), and by affecting the efficacy beliefs of followers (Pillai & Williams, 2004) improving their confidence in their own abilities. Transformational leadership has been related to such outcomes as objective sales performance (MacKenzie, Podsakoff, & Rich, 2001), customer service (Liao & Chuang, 2007), and business unit goal achievement (Howell & Avolio, 1993). These findings suggest that transformational leaders influence how followers see themselves and their work, as theorized above, and support the expected relationship between transformational leadership and surgical outcomes.

**Safety.** The role of leadership in fostering safety climates has also been investigated, specifically transformational leadership. Transformational leadership facilitates a positive safety culture and fosters safety consciousness (safety knowledge and behaviours) amongst employees (Barling et al., 2002; Zohar, 2002a), and creates an environment of openness to having safety issues raised (Mullen, 2005). Transformational leaders achieve these safety-related attitudes through each of the four components of transformational leadership: idealized influence places emphasis on safety as a core value, discouraging a short-term focus created by everyday pressures; inspirational motivation encourages followers to work for the collective good; intellectual stimulation helps followers think about new ways to improve safety; individualized consideration demonstrates care for followers’ well-being, including their safety. The rationale for the first three components supports the theorized way that transformational leaders in the operating room change how team members view relationships and their work. The safety attitudes elicited by transformational leaders are also negatively related to accidents and injuries in the workplace (Barling et al., 2002; Zohar, 2002a; 2002b), supporting the potential effects of transformational leadership on surgical outcomes.
**Extreme contexts.** While limited, there is some support for the effect of transformational leadership on patient safety in the medical context, as well as evidence of the effects of transformational leadership in other extreme contexts.

First, a positive safety climate has been suggested as one way in which the incidence of errors in the medical environment could be reduced, and change “unhelpful cultures” to “safety cultures” – open cultures that promote learning from errors and not placing blame for errors (Joyce, Boaden, & Esmail, 2005). These safety cultures have been negatively related to treatment errors in hospitals (Katz-Navon, Naveh, & Stern, 2005), similar to how safety climate is negatively related to injuries in traditional contexts (e.g., Barling et al., 2002). Given the positive relationship between transformational leadership and safety culture in traditional contexts (e.g., Barling et al., 2002), I would expect that transformational leadership will exert similar effects in the operating room.

Second, data from other extreme contexts on the effect of transformational leadership on performance further support of the effect of transformational leadership on surgical outcomes. Bass, Avolio, Jung and Berson (2003) showed that transformational leadership had a positive impact on unit performance in the military. While this research was done using a simulation exercise, this does suggest the transferability of transformational leadership’s positive effects to employees in extreme contexts, and the potential impact of transformational leadership on surgical outcomes. Within the extreme context of aviation, crew training encouraging the First Officer (second in command to the pilot) to question the pilot’s decisions (which is consistent with transformational leadership) have been introduced to reduce errors (e.g., Helmreich & Clayton, 1993; Li, Harris, & Chen, 2007). Similar practices have been successfully applied within anesthesiology (Fletcher, McGeorge, Flin, Glavin, & Maran, 2002), and more recently to
improve safety adherence of surgical teams (France, Leming-Lee, Jackson, Feistritzer, & Higgins, 2008), further suggesting the transferability of transformational leadership into the operating room.

Given the positive effect of transformational leadership on performance and safety in traditional contexts, and the positive effects in extreme contexts, I predict that:

**Hypothesis 1a:** A greater consistency of transformational leadership behaviours will be related to fewer errors, less blood loss, and fewer complications at discharge and follow-up.

**Laissez-faire Leadership and Surgical Outcomes**

I expect that laissez-faire leadership will have a detrimental impact on both proximal and distal surgical outcomes. Like transformational leadership, laissez-faire leadership is expected to influence the way that team members view themselves, their relationships, and their work, but negatively.

Laissez-faire leadership may negatively impact how team members view themselves, as laissez-faire leadership (a) provides no feedback about individual capability, and (b) the experience of “non-leadership” is stressful (e.g., Skogstad et al., 2007). Feedback, whether positive or negative, is important for self-efficacy judgments (Bandura, 1982). Without such feedback, which would typify the failure to reward or punish behaviours appropriately (Hinkin & Schreisheim, 2008), individuals may question their capabilities. Stress may also interfere with individuals’ feelings of efficacy, as psychological states could negatively influence performance expectations (e.g., Bandura, 1977), leaving followers questioning their capabilities. Together, this leaves team members uncertain about their own performance capabilities, and unsure of how they are to achieve outcomes, which is expected to translate into poor surgical outcomes.

In terms of how followers view their relationships, the stress created by laissez-faire leadership is likely to result in a team environment characterized by agitation, aggression, and
conflict (e.g., Skogstad et al., 2007). Skogstad et al. (2007) also showed an increased incidence of bullying between co-workers as a result of laissez-faire leadership. This environment is detrimental to the achievement of positive surgical outcomes, as team members do not work together to achieve goals, and instead are antisocial and combative.

Finally, laissez-faire leadership may negatively impact how followers view their job. Laissez-faire leadership communicates indifference: toward team members, surgical outcomes, and patients. This shifts followers’ motivation from bettering patient safety and quality of life, to performing at minimal standards during surgery, or undermining coworkers (e.g., bullying; Skogstad et al., 2007), all to the detriment of surgical outcomes.

There is evidence in support of this rationale from the areas of research on performance and safety in traditional organizations, on errors in extreme contexts, and on other outcomes in extreme contexts, each of which is discussed in turn below.

**Performance.** Research on laissez-faire leadership remains scarce, but in traditional organizational contexts laissez-faire leadership is not only the absence of “good” leadership, it is a “destructive” form of leadership – as demonstrated by the undesirable outcomes laissez-faire leadership is associated with (Skogstad et al., 2007). Laissez-faire leadership is negatively related to individual (Hinkin & Schriesheim, 2008; Bass & Yammarino, 1991) and group task performance (Howell & Avolio, 1993). Laissez-faire leadership is also negatively related to other factors that influence performance, such as role clarity, satisfaction with supervisor (Hinkin & Schriesheim, 2008), and subordinate motivation (Judge & Piccolo, 2004), and positively related to role conflict, role ambiguity, and coworker conflict (Skogstad, et al., 2007). By failing to provide positive or negative feedback or support, employees will lack the goals, direction,
motivation, or tools to improve task performance. These findings in traditional organizations support the proposition that laissez-faire leadership will negatively affect surgical outcomes.

**Safety.** One area in which laissez-faire leadership has received attention in traditional contexts is safety, and the effects of “turning a blind eye” toward safety issues. The theoretical rationale underlying the predicted negative effects of laissez-faire leadership on safety rests on the idea that inattention or negativity toward safety issues communicates to followers that safety is not important, and that followers should also treat it as such. Laissez-faire leadership is negatively related to safety outcomes, supporting the expected negative impact in the operating room. Zohar (2002b) showed that laissez-faire leadership behaviours were negatively related to safety climate, and indirectly related to workplace accidents and injuries. Kelloway, Mullen and Francis (2006) introduced the idea of safety-specific passive leadership (i.e. ignoring safety concerns until intervention becomes critical; composed of both laissez-faire and passive management-by-exception behaviours). Safety-specific passive leadership was negatively related to safety climate and safety consciousness, and mediated the relationship between leadership and safety-related incidents. Finally, Mullen et al. (2011) showed that safety-specific passive leadership negatively affected safety behaviours, and also further attenuated any positive effects of safety-specific transformational leadership. This further supports the negative effects that laissez-faire leadership may have on surgical team values, and the ultimate negative effect on surgical outcomes.

**Errors and extreme contexts.** Shapell and Wiegmann (1997) proposed that poor leadership could contribute to the incidence of errors, noting that unsafe supervision is a potential cause of errors. Unsafe supervision results from either a loss of supervisory awareness of the situation, inadequate policies and procedures, lack of leadership, or failure to correct
known problems. Failure to correct known problems further suggests the effect that laissez-faire leadership might have on errors. While there seems to be no research on the effect of laissez-faire leadership on errors, findings from other areas indirectly support its potential role.

Advances in research on the antecedents of errors have been made in the extreme context of aviation, with some findings suggesting that laissez-faire leadership will be positively related to errors. In a review of 523 accidents from the Republic of China Air Force, multiple errors were shown to contribute toward an accident (Li & Harris, 2006). This compounding effect has also been found in other areas of error and disaster research (e.g., Perrow, 1984). Providing support for the link between leadership behaviours and errors, Li and Harris noted that inadequate supervision (defined as failure to “provide guidance and operational doctrine to pilots” pg. 1060) was a key factor in aviation errors, suggesting that laissez-faire leadership behaviours could negatively influence errors. The operating room has been compared to the cockpit in terms of how it is managed and how errors can be avoided (e.g., Helmreich & Davies, 1996), which suggests that similar causes of errors in the operating room might apply.

**Extreme contexts.** Finally, while there is no research on laissez-faire leadership in any extreme context (beyond that discussed above), evidence from the research presented in Chapter 2 of this dissertation support the rationale that laissez-faire leadership negatively impacts how followers view their work, and that laissez-faire leadership may negatively influence surgical outcomes. In that context, laissez-faire leadership negatively predicted follower empathy. While not directly related to performance outcomes, this finding suggests that laissez-faire leadership changes the way that followers view their patients, and by extension their work. As theorized earlier, this uncaring disposition could in turn affect how followers treat those patients, ultimately affecting surgical outcomes.
Thus, I predict that:

**Hypothesis 1b:** A greater consistency of laissez-faire leadership behaviours will be related to more errors, more blood loss, and more complications at discharge and follow-up.

**Over-controlling Leadership and Surgical Outcomes**

The final leadership behavior in my study, over-controlling leadership, is expected to detrimentally affect surgical outcomes, again by negatively influencing the way followers view themselves, their relationships, and their work.

While not necessarily intentional, over-controlling leadership communicates distrust of followers. As leaders control exactly how followers perform their work, they communicate that followers cannot adequately perform without assistance, and cannot be trusted. This may cause followers to question whether they can successfully achieve positive surgical outcomes, which in turn is expected to negatively influence actual achievement. Furthermore, the pressures that over-controlling leadership places on performance and time may be considered stressors, as followers have no autonomy, and are pressured to work a certain way (e.g., Frone, Russell, & Cooper, 1995). This stress further depletes followers’ confidence and efficacy, further impeding the achievement of positive surgical outcomes.

Over-controlling leadership may also foster a team environment emphasizing hierarchy, rigidity, and social distance (e.g., Antonakis & Atwater, 2002), as over-controlling leadership leaves little room for followers to question the leader, or provide alternative suggestions for how to optimize surgical outcomes. This limits the potential for new and optimal performance solutions to be discussed or used, as the leader instead dictates what is to be done. Furthermore, given the historically hierarchical nature in the way in which surgical teams function, followers may be less likely to raise concerns that cast doubt on the leader’s decisions, or affect how long the task could take, for fear of being reprimanded. Finally, Dupre and Barling (2006) showed...
that over-controlling behaviours increased psychological aggression toward supervisors, and this relationship was mediated by interpersonal injustice. This suggests that over-controlling leadership creates a suboptimal team environment. The effect of over-controlling leadership on interpersonal relationships in the operating room is thus expected to impede optimal surgical outcomes.

Finally, over-controlling leadership influences how followers view their work. Over-controlling leadership may create compliance in followers with what the leader wants, by closely monitoring their performance, which is unlikely to translate to positive surgical outcomes for two reasons. First, while higher compliance may ensue, the behaviours elicited are likely directed toward what is being rewarded: Fast performance in compliance with leader expectations, which could be counter to the goals of positive patient outcomes. Second, focusing on minute details (e.g., sticking to the schedule, following all procedures) shifts focus away from the broader, long-term view of the patient’s well-being. This short-term and narrow view of performance may negatively impact surgical outcomes, as followers take short-cuts or risks to meet the expectations of the leader, to the detriment of the patient.

While over-controlling leadership is not widely researched, evidence from similar constructs (i.e. destructive and toxic leadership, job control and autonomy) provides support for the predicted relationship between over-controlling leadership and proximal and distal surgical outcomes. This support comes from research on performance and safety in traditional organizations, and related variables (e.g., toxic leadership) in extreme contexts.

**Performance.** As discussed above, the excessive control placed on employees limits the options they have to find and use performance solutions. The behaviours inherent in over-controlling leadership also demonstrate a lack of confidence in employees’ abilities to
successfully perform. Control, from any source, has long been treated as an important variable in organizational research (e.g., Spector, 1986). Employees’ lack of control is negatively related to performance (e.g., Dupre, Barling, & LeBlanc, 2004), and positively related to undesirable workplace outcomes such as turnover intentions (e.g., Barling & Kelloway, 1996), absenteeism (e.g., Spector, 1986), and aggression (e.g., Dupre & Barling, 2006). When these over-controlling behaviours emanate from a superior, I would expect that these same relationships will exist, and could even be stronger due to the power difference in a leader-subordinate relationship (e.g., Clements & Washbush, 1999). Furthermore, employees with greater autonomy feel more committed to their work and are intrinsically motivated to perform (Hackman, et al., 1975; Spector, 1986). Giving employees the authority to make decisions demonstrates trust and gives employees a sense of responsibility toward the organization (Ryan & Deci, 2000). Over-controlling leadership tells subordinates that they are not trusted, cannot perform on their own, and are not valuable organizational members.

**Safety.** While over-controlling leadership has not been studied in the context of safety climate, the effect of general control on safety in the workplace has been subject to research. Findings suggest that over-controlling leadership will have a detrimental impact on safety climate. Specifically, personal control is related to proactive safety behaviours (Geller, Roberts & Gilmore, 1996), and autonomy is positively related to safety compliance (Parker, Axtell & Turner, 2001), and propensity to undertake safety initiatives (Simard & Marchand, 1995). While autonomy is not interchangeable with control, it does represent greater control. Over-controlling leadership is expected to limit autonomy, and thus I would expect that an environment with an over-controlling leader will be limited in these positive safety behaviours.
**Extreme contexts.** While over-controlling leadership has not been studied in extreme contexts, as mentioned earlier, a related construct, namely toxic leadership, has been examined and provides support for the potential effect of over-controlling leadership on surgical outcomes.

As noted earlier, toxic leadership encompasses aspects of over-controlling leadership (e.g., inflexible, authoritarian), and results of this research point to the possible detrimental impact of over-controlling leadership on surgical outcomes. Toxic leadership in the military is negatively related to follower job satisfaction and organizational commitment (Gallus, Walsh, van Driel, Gouge, & Antolic, 2013), and follower ethics, confidence in the leader, creativity, and problem-solving (Steele, 2011). These negative relationships reflect how toxic leadership negatively influences how followers see their work, and suggest that over-controlling leadership could elicit similarly negative results.

Thus, I predict that:

**Hypothesis 1c:** A greater consistency of over-controlling leadership behaviours will be related to greater errors, more blood loss, and more complications at discharge and follow-up.

**Mediators of the Leadership-Surgical Outcome Relationships**

The influence of leadership on organizational outcomes is rarely direct (Barling et al., 2010). In the operating room, I propose that two variables (i.e. psychological safety and boredom) will mediate the relationship between leadership and surgical outcomes. Each of these mediators is discussed in turn below, and the proposed model can be found in Figure 3-1.

**Psychological Safety**

Psychological safety is an important variable in all organizations, but especially those with a high concern for safety (i.e. the operating room). Psychological safety was defined by Edmondson (1999) as “a shared belief that the team is safe for interpersonal risk taking” (p. 354),
and is often implicit, rather than discussed explicitly by group members. Groups with high levels of perceived psychological safety will leave members feeling comfortable taking interpersonal risks within the group, such as dissenting, voicing their own opinions, and engaging in learning behaviours such as sharing information, discussing errors, asking for help, and experimenting (Edmondson, 1999). Team members would not be afraid of being judged as incompetent, punished, or rejected should they speak up, make a mistake or ask for help. This is important, because an unwillingness or fear to take an interpersonal risk in an extreme context has the potential to cause physical risks, as the ability to admit to or speak up when a mistake has been made is often critical to avoid catastrophe (Edmondson, 2003).

Psychological safety is often a mediating variable, used to explain the effect of a behaviour on a variety of outcome variables (e.g., learning from failures; Carmeli & Gittell, 2009). Psychological safety is an explanatory construct (that helps answer the question of “why” two variables are related; Edmondson, 2003), so understanding the factors that influence it is important. One such predictor of psychological safety is leadership behaviour.

**Leadership and psychological safety.** In her early work on psychological safety, Edmondson (e.g., 1999) identified leaders as crucial in influencing teams’ psychological safety. She suggests that leaders influence psychological safety through role-modeling, and supportive coaching behaviours. She expects that a supportive, encouraging leader, who is non-defensive when responding to questions and non-punishing of desirable learning behaviours, is likely to make team members feel part of a safe environment. Conversely, an authoritarian leader who often uses punishment to influence team member behaviour is likely to make members feel unsafe to take interpersonal risks.
The positive behaviours described by Edmondson are similar to those of transformational leadership, and suggest that these behaviours will influence psychological safety. Transformational leadership encourages followers to think for themselves and question assumptions, and fosters an environment in which team members feel free to challenge each other and their leader, and raise any safety and performance concerns.

Several studies support the prediction that transformational leadership will positively influence psychological safety. Schaubroeck, Lam, and Peng (2011) showed that transformational leadership and servant leadership improved psychological safety, which in turn improved performance (measured by supervisory ratings of team effectiveness). Walumbwa and Schaubroeck (2009) found a relationship between ethical leadership and psychological safety. Ethical leadership, which is closely related to the idealized influence component of transformational leadership, involves setting value-based examples for followers through role modeling, two-way communication, and positive reinforcement. This is much the same way a transformational leader uses idealized influence and individualized consideration. In a separate study, dyadic discovery (getting to know the attitudes, beliefs, and values of individual team members) predicted greater levels of team psychological safety, which in turn predicted greater team performance (Roussin, 2008). Finally, leaders’ behavioural integrity for safety, which is consistent with idealized influence, improves team psychological safety and reporting of errors (Leroy et al., 2012).

Thus it is expected that:

**Hypothesis 2a:** A greater consistency of transformational leadership will be related to higher psychological safety

In contrast, I expect that laissez-faire leadership will negatively influence psychological safety. Laissez-faire leadership would not involve any of the behaviors that Edmonson identified
as crucial to the development of psychological safety. Furthermore, the passive nature of laissez-faire leadership could be interpreted as an unwillingness or fear on the part of the leader to take interpersonal risks, and therefore implicitly convey to team members that such behaviours should be avoided.

Empirical research on the influence of leadership on psychological safety has been limited to positive forms of leadership. No evidence exists on the relationship between laissez-faire leadership and psychological safety, though laissez-faire leaders would not be expected to foster a psychologically safe environment, and given its effects on behaviors such as bullying (Skogstad et al., 2007), could even negatively impact its development. First, a laissez-faire leader creates ambiguity in what is expected of followers (e.g., Skogstad et al., 2007). Laissez-faire leadership results in uncertainty and ambiguity about how to perform jobs, and expected behaviours. With no cues to followers as to whether they will be rewarded or reprimanded for taking interpersonal risks, followers may be disinclined to take risks. Furthermore, by not taking interpersonal risks, leaders signal to the rest of the team that they are expected to act similarly. Second, as discussed earlier, ambiguity created by laissez-faire leadership is also a source of stress that elicits negative outcomes such as coworker conflict, and bullying (e.g., Skogstad et al., 2007), indicative of the absence of psychological safety.

Given this, I expect that:

**Hypothesis 2b:** A greater consistency of laissez-faire leadership will be related to lower psychological safety.

Finally, over-controlling leadership imposes control over employee actions and demonstrates distrust in employees, creating a psychologically unsafe environment for employees. While there is no empirical research linking over-controlling leadership behaviours to psychological safety, I would suggest that over-controlling leadership will have a negative
effect on psychological safety. First, there is a negative relationship between follower lack of control and trust in leadership (Ryan & Deci, 2000; Steele, 2011). A team that does not trust the leader would be expected to feel less psychologically safe, and less likely to take interpersonal risks such as questioning the leader, asking for help, or admitting to mistakes, for fear of reprimand. Second, there is a positive relationship between destructive leadership (which includes some over-controlling behaviours) and employees feeling that they are penalized for honest mistakes (Steele, 2011). One characteristic of a psychologically safe environment is the ability to raise concerns or admit to mistakes without fear of being penalized. The positive relationship between destructive leadership and follower feelings of being penalized for these behaviours suggests that such leadership behaviours negatively affect psychological safety.

Thus I expect that:

**Hypothesis 2c:** A greater consistency of over-controlling leadership will be related to lower psychological safety.

**Psychological safety and surgical outcomes.** As described earlier, teams that are high in psychological safety have members who feel comfortable questioning others’ decisions, voicing their opinions, sharing information, discussing and learning from mistakes, and asking for help (Edmondson, 1999). In the operating room, psychological safety will help positively impact proximal surgical outcomes, as team members are (a) more willing to mention when an error has been made (e.g., when the sterile field has been broken), (b) ask for help (e.g., when blood loss is beginning to escalate), and (c) be open to discussing the optimal methods with which to treat a patient. These behaviours would also affect distal surgical outcomes: By speaking up when errors are made (or speaking up before they are made) teams can then take precautions to remedy the error and compensate for any possible negative effects, thus reducing the instance of longer-term complications. By openly sharing information, asking for help, and
voicing opinions, a wide range of methods and treatments specific to a patient may be developed, considering multiple perspectives.

There is some support for the potential effect of psychological safety on surgical outcomes. Empirical findings show that psychological safety predicts team performance in a manufacturing company (Edmondson, 1999), financial performance and goal-achievement across a range of mid-size companies (Baer & Friese, 2003), and successful implementation of new technology in a healthcare setting (Edmondson, 2003). The positive effects of psychological safety on these organizational outcomes indicate that psychological safety may positively affect surgical outcomes. Furthermore, psychological safety is related to other desirable organizational outcomes such as work engagement (Nembhard & Edmondson, 2006), goal achievement, innovation, and creativity (Carmeli, Reiter-Palmon, & Ziv, 2010). Most importantly, psychological safety is related to learning from mistakes to prevent future errors (Carmeli & Gittell, 2009), suggesting that psychological safety will affect errors in the operating room.

Given this, I predict that:

**Hypothesis 2d:** Psychological safety will be positively related to surgical outcomes.

**Hypothesis 2e:** Psychological safety will mediate the relationship between consistency of transformational leadership and surgical outcomes.

**Hypothesis 2f:** Psychological safety will mediate the relationship between consistency of laissez-faire leadership and surgical outcomes.

**Hypothesis 2g:** Psychological safety will mediate the relationship between consistency of over-controlling leadership and surgical outcomes.

**Boredom**

I suggest that boredom will mediate the relationship between consistency of leadership and performance. Boredom is defined as “an unpleasant, transient affective state in which the
individual feels a pervasive lack of interest in and difficulty concentrating on the current activity . . . [and] feels that it takes conscious effort to maintain or return attention to that activity” (Fisher, 1993, pg. 396). Van Tilburg and Igou (2012) showed that boredom is a distinct affective experience, where people feel they are not challenged, and their situation and actions meaningless. While little empirical research exists on boredom, I suggest that leadership may influence feelings of boredom, which in turn could hinder performance.

**Leadership and boredom.** Several factors reduce feelings of boredom in traditional contexts. Specifically, challenging goals reduce boredom (at least on simple tasks), and outside stimulation such as music or white noise reduces boredom on visual tasks (Locke & Bryan, 1967; McBain, 1970; Warm & December, 1986). Job design (e.g., greater responsibility/autonomy, job rotation) and task delegation also reduce boredom (Fisher, 1991). What is of importance from the perspective of an understanding of extreme contexts is that most of these methods for reducing boredom are at least somewhat under the control of a leader (when situational policies and procedures allow): Setting goals, creating the environment, finding meaning, and job design and delegation. Thus, I expect that leadership will impact the boredom experienced by team members.

Transformational leadership is expected to decrease feelings of boredom. Intellectual stimulation challenges followers to view their work from different perspectives, a cognitive activity that would reduce boredom. Griffin (1983) showed that leaders can be trained to point out interesting aspects of the job to their employees, thus reducing boredom. Inspirational motivation and idealized influence would reduce feelings of boredom by providing meaning and purpose to followers’ work. Indeed, transformational leadership enhances the extent to which employees perceived their work to be meaningful (e.g., Arnold, Turner, Barling, Kelloway &
McKee, 2007; Nielsen, Randall, Yarker & Brenner, 2008; Neilsen, Yarker, Brenner, Randall & Borg, 2008). Finally, through individualized consideration, transformational leaders recognize idiosyncratic needs for stimulation and encourage engagement where needed, and recent research (Guglielmi, Simbula, Mazzetti, Tabanelli, & Bonfiglioli, 2013) showed a negative relationship between individualized consideration and boredom. As a result, I predict:

**Hypothesis 3a:** A greater consistency of transformational leadership will be related with lower team member ratings of boredom.

In contrast, laissez-faire leadership may contribute to feelings of boredom. The ambiguity and lack of direction that characterize laissez-faire leadership could contribute to feelings of boredom. Furthermore, without goals, employees could have difficulty finding meaning in their work, reducing motivation and increasing feelings of boredom. This is consistent with Van Tilberg and Igou’s (2012) findings that a lack of challenge and meaning contribute to boredom.

Given this, I predict that:

**Hypothesis 3b:** A greater consistency of laissez-faire leadership will be related with higher team member ratings of boredom.

Finally, the effects of over-controlling leadership on boredom are uncertain. On one hand, an over-controlling leader engaging in micromanaging and performance pressure could reduce boredom (albeit in an undesirable way). On the other hand, over-controlling leadership may limit opportunities to prevent boredom (e.g., by forbidding music, conversation, reducing opportunities for employees to think for themselves). In job design research, autonomy is negatively related to boredom (Fisher, 1991); having autonomy taken away could theoretically increase boredom. However, given the potential bi-directional effect, no hypothesis is offered.
**Boredom and surgical outcomes.** I expect that boredom will have a detrimental impact on surgical outcomes. Employees who are bored would not be paying full attention to their jobs, which could increase the incidence of the proximal outcomes of errors and blood loss, and distal outcomes of complications. Furthermore, those experiencing a lack of challenge and meaning in their work would not be motivated toward achieving optimal surgical outcomes.

While there is no research on the outcomes of boredom in extreme contexts, the results in traditional contexts show that boredom negatively affects organizational outcomes. In one study of long-haul truck drivers, performance (measured as absenteeism, and property damage to the truck and unloading stations) suffered as boredom increased (Drory, 1982). These demonstrate boredom’s effects on withdrawal from work (represented by absenteeism), and safety behaviours (represented by damage due to carelessness). When “performance” involves safely performing surgery on a patient, any decline in that performance could be hazardous. Boredom results in attention lapses, increasing the chance of error (Cox, 1980; Drory, 1982; O'Hanlon, 1981). Smith-Jentsch et al. (2012) found that astronauts had difficulty transitioning quickly from boredom to “hyper-vigilance”, suggesting a slower reaction time when action was necessary and urgent. Boredom also predicts workplace injuries among young workers (Frone, 1998).

Additionally, boredom is associated with other affective responses such as agitation, restlessness, and feeling upset (Robinson, 1975). In a study of pilots, hostility increased after long, monotonous (boring) flights (O'Hanlon, 1981). Because boredom elicits withdrawal, a lack of concern for safety, and issues of attention, I propose that:

**Hypothesis 3c:** Boredom in the operating room will negatively impact surgical outcomes.
Considering the relationship between boredom and performance and the hypothesized relationships between two of the leadership behaviours and boredom, a mediated relationship is proposed, in which leadership affects surgical team performance through boredom:

**Hypothesis 3d:** Boredom will mediate the relationship between consistency of transformational leadership and surgical outcomes.

**Hypothesis 3e:** Boredom will mediate the relationship between consistency of laissez-faire leadership and surgical outcomes.

**The Moderating Effects of Complexity**

Thus far the model I have developed could apply within most organizations. However, as Johns (2006) reminds us, context matters—and this is especially true of extreme contexts. As I now suggest, surgical complexity will influence the expression and consequences of leadership. Surgical complexity is a team-rated judgment of the potential magnitude and probability of adverse outcomes, considering the type of surgery, the patient condition, and other situational factors (e.g., need for multiple specialties). A low complexity surgery, for example, would be a surgery on an otherwise healthy patient, a procedure the surgeon has completed numerous times, with low risks of adverse outcomes. A removal of a kidney stone by laser would fall into this category. A high complexity surgery, on the other hand, may involve a patient with many co-morbidities, multiple specialties and surgeons, with a greater probability of adverse outcomes. An example of this might be an obese victim of a car accident, who requires the removal of a spleen, repairs to other internal organs, and bones – with patient death being a high likelihood. Since it captures the magnitude and probability of adverse outcomes, two of the elements of an extreme context, surgical complexity is a relevant contextual variable to the study of extreme contexts. Below I explore the role of complexity in the model I have developed relating leadership and surgical outcomes (see Figure 3-1).
Complexity, Leadership and Surgical Outcomes

The effects of leadership on surgical outcomes is expected to vary based on the degree of surgical complexity. For the purposes of explanation, this discussion is separated into the effects of leadership on outcomes under low, mean, and high levels of complexity.\(^\text{12}\)

**Low complexity.** Within low complexity contexts, leaders must help team members overcome complacency (Hannah et al., 2009). Even though these surgeries are likely to be standardized and routine, I expect that leadership is more important in lower complexity surgeries than in surgeries of average complexity. Positive forms of leadership that stimulate employees, ensure proper preparations are undertaken, and due diligence is undertaken despite the low probability of negative outcomes will be critical in low complexity surgeries. Without such leadership, team members may be susceptible to thinking that it is unlikely that something bad will happen, and may loosen their adherence to protocols. This will be reflected in a stronger relationship between leadership and positive surgical outcomes in low complexity surgeries (i.e. a stronger positive relationship for transformational leadership), and stronger negative relationships for laissez-faire and over-controlling leadership. The effect of complacency and low probability of consequence in the absence of good leadership can be found in the safety literature, where it has been shown that complacency from over-reliance on technology increases accidents (Singh, Molloy, & Parasuramen, 1993), and complacency is a potential cause of the *Challenger* disaster (Vaughan, 1996). Vaughan made the point that when operators have been doing the same work for a while without consequence (despite exposure to hazards and risks), they become complacent, and the likelihood of an accident increases.

\(^\text{12}\) While the discussion is organized categorically for descriptive purposes, complexity is measured as a continuous variable.
Positive leadership in the operating room will be important to combat this, and nonleadership or negative leadership could exacerbate it.

Evidence of the effect of surgical complexity as a moderator emerges from a recent study by Vashdi et al. (2013), in which complexity moderated the relationship between team reflexivity (i.e. briefing and debriefing performance events for the purposes of learning) and errors. Specifically, in low complexity surgeries, team reflexivity had no effect on errors – “simple” surgeries with less potential for learning and improvement did not benefit from debriefing strategies and were left susceptible to complacency. The opposite effect is expected with respect to leadership, in others words, I expect that transformational leadership will have a stronger positive effect, and that laissez-faire leadership and over-controlling leadership will have a stronger negative effect on surgical outcomes.

**Moderate complexity.** As complexity increases, but procedures are still relatively familiar and standardized, team members are more likely to engage with the task and others more intently (Hannah et al., 2009). At a mean level of complexity, the impact of leadership is likely to be less than at lower complexity surgeries due to the engagement of team members. Leadership is still expected to exert significant effects on valued outcomes, as in traditional contexts, but any effects would not be as strong as low complexity surgeries. Evidence for the moderating effects at medium complexity is again shown in Vashdi et al. (2013), as the greatest impact of team reflexivity on errors emerged at medium levels of complexity.

**High complexity.** Finally, in higher complexity situations, though team members are still likely to be highly engaged and stimulated, leadership will become more important to the achievement of surgical outcomes, as procedures are less familiar, and are uncommon. In high complexity surgeries, leadership will be necessary to maintain positive emotions, self-efficacy
beliefs, and sense-making. The result is expected to be a stronger relationship between leadership and surgical outcomes than at average levels of complexity. Vashdi et al’s, (2013) findings are again suggestive, as team reflexivity was ineffective in predicting errors in high complexity situations, presumably because highly complex surgeries are inherently unpredictable and abnormal, and anything learnt from a surgery of this nature will be difficult to generalize and apply to subsequent surgeries. On the other hand, as described above, I expect that leadership will exert a stronger influence on surgical outcomes in high complexity cases.

Thus, I predict that complexity will moderate the relationship between leadership behaviours and performance outcomes in a curvilinear fashion: a strong moderating effect for lower complexity surgeries, a mild effect for medium complexity surgeries, and the same strong effect again for higher complexity surgeries.

Thus:

**Hypothesis 4:** Complexity will be a curvilinear moderator of the relationship between leadership behaviours and surgical outcomes, with leadership having a stronger impact in lower and higher complexity situations than in medium complexity situations.

**Psychological Safety and Complexity**

I posit that complexity will moderate the paths from leadership to psychological safety, and from psychological safety to surgical outcomes. Hannah et al. (2009) suggest that some psychological variables, such as trust, may function differently in an extreme context (Hannah et al., 2009). I propose the role of psychological safety would function differently under differing conditions of complexity.

First, I expect that the relationship between leadership behaviors and psychological safety will be moderated by complexity. As complexity increases, it will be more difficult for transformational leadership to foster psychological safety, and laissez-faire leadership and over-
controlling leadership will be more detrimental to the development of psychological safety (i.e. a stronger negative relationship).

As Hannah et al. (2009) theorized, as the magnitude of probability of consequence increases, followers are expected to reassess their leaders’ competence. Extending this to psychological safety, when surgeries are most complex, team members may reevaluate the safety of their environment for interpersonal risk-taking. The same positive leadership behaviours (i.e. transformational) that foster psychological safety in less complex situations, may not be sufficient to demonstrate that questioning decisions or admitting to mistakes, are acceptable under higher complexity situations. The same negative leadership behaviours (i.e. laissez-faire, over-controlling) that produce negative effects on psychological safety in low complexity situations may have much stronger negative effects in high complexity cases. This reevaluation of interpersonal factors may occur as followers deem the skills and character of leaders to be more important in more extreme situations (Sweeney, 2010). Sweeney (2010) found that subordinates reevaluated the trust they had in their military commanders prior to going into combat (i.e., facing a greater risk of injury or death), giving overall lower ratings of trust. This suggests that as the magnitude of probability of consequences increased, followers placed less trust in leaders, which is the expected effect of complexity on the relationship between leadership and psychological safety.

Sweeney points to interdependence theory for this rationale (Thibaut & Kelley, 1959), which when applied to leader-follower relationships, suggests that the more dependent followers’ outcomes are on leaders’ behaviours, the more likely followers are to monitor and make inferences about the leaders’ abilities and traits. For example, over-controlling leadership under normal conditions restricts followers, and places social sanctions on those who “step out of line”.

189
As complexity increases and followers evaluate the leader and team environment more intensely, the restrictive behaviors involved in over-controlling leadership are expected to affect psychological safety more negatively. This change occurs as the situation becomes more ambiguous, the rules and procedures become less clear (i.e. the situation becomes weaker), and the followers are at the mercy of the leader.

Thus I predict that:

**Hypothesis 5a**: Complexity will moderate the relationship between consistency of transformational leadership and psychological safety.

**Hypothesis 5b**: Complexity will moderate the relationship between consistency of laissez-faire leadership and psychological safety.

**Hypothesis 5c**: Complexity will moderate the relationship between consistency of over-controlling leadership and psychological safety.

With respect to any effects of psychological safety on surgical outcomes, a similar reasoning to the effects of leadership on surgical outcomes may be applied. Under low complexity surgeries, psychological safety will be important to overcome the seeming simplicity of the procedures: When spotting mistakes, participants may second-guess speaking up for fear of looking stupid, or be afraid to ask for help for such a simple procedure. High psychological safety in such situations may help overcome some of these situations.

As complexity increases, and team members are more engaged and cautious in the slightly more complex (but still familiar) procedures, psychological safety may not be as critical in affecting their behaviour and achievement of surgical outcomes.

Finally, as complexity reaches the higher end, and the magnitude of consequences increase, psychological safety becomes critical for superior performance. As Hannah, Campbell and Matthews (2010) explain, “in a dangerous context it is entirely plausible that factors like
trust may not be optional, and that if a sufficient threshold level is not met, leadership influence may not even occur” (pg. S169). This suggests that the path between psychological safety and surgical outcomes will be moderated by complexity. As Hannah et al. (2010) suggested, low psychological safety at high levels of complexity, may result in lost coordination, confusion, and chaos. It is worth noting that Hannah et al. (2010) drew these conclusions based on research in a military setting, where presumably a more dangerous context puts soldiers’ lives at risk, opposed to a surgical setting where it is the patients’ lives on the line. I believe that while the differences in effect sizes may not be quite as drastic, similar patterns will emerge in surgical contexts.

Thus:

**Hypothesis 5d:** Complexity will moderate the relationship between psychological safety and surgical outcomes.

### 3.2 Method

The study examines leadership in operating rooms at a teaching hospital in Eastern Ontario. Using a combination of observation, survey, and archival measures of variables, I tested the effect of leadership behaviours (i.e. transformational, laissez-faire, over-controlling) on surgical outcomes, through the mediators of psychological safety and boredom, and moderated by complexity. The method is described in detail below, with descriptions of the participants, observational training, study procedure and data collection.

**Participants**

**Patient surgeries.** The sample consisted of 150 randomly selected patient surgeries. Patients were 53% male, with an average age of 51 years at the time of their surgery. The sample consisted of 84% pre-scheduled, and 16% emergency procedures\(^\text{13}\). 60% of the cases were classified as “B” cases or lower, in order to obtain patient consent. A “B” case is a case that is prioritized to be completed within 24 hours of admission to the hospital.
were from the General Surgery department, 27% were from the Obstetrics and Gynecology department, and 13% were from the Urology department. Orthopedic surgeries involving open joints were excluded due to sterility concerns.

**Surgical teams.** Teams were composed of surgeons, nurses, anesthesiologists, and residents in different specialties. A total of 42 different surgeons are represented in the sample. Not all surgeons opted to disclose demographic information, but of those who reported this information, the mean age was 50.4 ($SD = 9.4$), mean tenure at the hospital was 13.4 years ($SD = 10.1$), and mean professional tenure was 23.1 years ($SD = 9.7$). 93% of the surgeons were male. The rest of each team was composed of one of 32 anesthesiologists, one to three of 39 nurses, and varying numbers of the 39 surgical and anesthesiology residents. Of those disclosing demographic information, the anesthesiologists were 66% male, had a mean age of 44.4 ($SD = 6.8$), a mean tenure at the hospital of 11.5 years ($SD = 6.7$), and mean professional tenure of 18.7 years ($SD = 6.4$). The nurses were 5% male, were a mean age of 43.0 ($SD = 9.6$), had a mean tenure at the hospital of 12.3 years ($SD = 7.1$), and mean professional tenure of 14.8 years ($SD = 7.8$). Residents were 54% male, a mean age of 30.2 ($SD = 3.3$), had a mean tenure at the hospital of 2.7 years ($SD = 1.35$), and mean professional tenure (not including medical school) of 2.9 years ($SD = 1.5$).

**Observer/Research Assistant Hiring and Training**

Observers/Research Assistants\(^{14}\) were required for observing and recording leadership behaviours and errors, administering paper surveys to team members, and collecting information from patient records after the surgery was completed. Observers also assisted in follow-up patient phone calls.

\(^{14}\) While students were responsible for being both Observers and Research Assistants, the term “Observer” will be used through the remainder of this document for simplicity.
Medical and nursing students were recruited through advertisements on the medical school and nursing program websites. Students wishing to apply were asked to send a resume to the study investigators. A total of 10 submitted resumes. All applicants were then asked to complete a 20-minute interview with two study investigators, in which questions were asked about teamwork, patient interactions and consent, and interacting with surgical team members. Two second-year medical students and three third-year nursing students were hired for the positions.

Observers attended a full day of training; half the day focused on leadership behaviours, and the remainder of the day on medical errors, and study methodology. Training on leadership behaviours was led by Dr. Julian Barling and myself, and provided students with information on the three leadership behaviours (i.e. transformational leadership, laissez-faire leadership, over-controlling leadership), along with examples of how these behaviours might manifest in an operating room environment.

The leadership training concluded by completing three sample observations using the observation sheet and videos of three different leaders, to ensure an adequate level of interrater reliability. Interrater reliability of the observations across the three sample observations during the training, calculated using the weighted Cohen’s kappa (Cohen, 1968) was 0.77 for transformational leadership, 0.33 for laissez-faire leadership, and 0.25 for over-controlling leadership\(^{15}\). Following the viewing and rating of each video, we discussed the observers’ ratings as a group, and discussed any of the discrepancies.

\(^{15}\) Please note that these data were calculated on the basis of \(N = 3\), and thus are only reported for informative purposes, and for the trainers to guide their subsequent discussion about identifying leadership behaviours.
The medical error training was led by Dr. Darren Beiko, who, as a practicing surgeon and medical school professor, has expertise in medical errors. His training involved identifying any adverse events, complications or unexpected deviations from the normal anesthetic or surgical procedure. A similar training procedure was successfully used by Vashdi et al. (2013). Dr. Beiko’s training also included a tour of the OR department, and demonstrated the use of the electronic patient record system where the students would login to collect additional information. Training included showing observers the research forms they would be required to complete, and observers were familiarized with the observation forms (Appendix B). Specific study hypotheses were not disclosed to observers. Materials (e.g., PowerPoint slides) from the training can be found in Appendix C.

**Procedure**

**Consent.** The study required consent by all surgical team members, as well as the patient. Consent from surgical team members was obtained at various weekly departmental meetings prior to the start of the study, where all potential participants were informed about the study. Of all medical staff, most agreed to participate\(^\text{16}\).

Patient consent was obtained by the observers on the day of the surgery. Patients were briefed and given the chance to have any questions answered.

**Surgical sampling procedure.** Surgeries were randomly selected using a random number generator in Microsoft Excel, on the morning of a designated observation day from the pool of surgeries with consenting staff members. Selected surgeries were arranged in scheduled order, and assigned among the observer teams. Observers were instructed to observe the first surgery, and upon completion proceed to the next scheduled surgery that had not yet begun.

\(^{16}\) To respect confidentiality, no further specific details (e.g. demographic, profession) details of non-consenting staff are included.
**Observer scheduling.** Observers were present in the OR for three day and two evening shifts each week. Observers worked in pairs, and there were two pairs working during the day, and one pair in the evenings. When possible, the pairs were composed of one medical student and one nursing student. Observers rotated through the evening shifts. As observers did not know each other prior to this study, pairs were initially assigned by the investigator, and then adapted to the different schedules (i.e. as rotation through the night shifts dictated).

**Observation materials.** An observation sheet was developed to record pre- and post-operative data from team members and patient records, as well as intraoperative observations of leadership behaviours and errors (Appendix B). The intraoperative observation sheet was developed based on observational methods used in sports research (Darst, Zakrajsek, & Mancini, 1989) and developmental psychology research (Frick, Barry, & Kamphaus, 2009), which rely more heavily on observational data than do organizational research. The form was designed to ease the “paper-and-pen” process of event recording observations (Frick et al., 2009), and required observers to write the time of the observation, who performed the behaviour, circle whether it was a leadership behaviour or an error, and if it was a leadership behaviour, to circle the corresponding type of leadership behaviour.

**Measures.**

**Predictor variables.** The predictor variables of transformational leadership, laissez-faire leadership, and over-controlling leadership were collected through observation. The number of observations of each leadership behaviour was tallied for each surgery, and used to calculate a percentage. The final variables represent the proportion of total leadership behaviours (between 0 and 1) in a specific surgery that were transformational leadership, laissez-faire leadership, or over-controlling leadership. For interpretation, a score of 0 transformational leadership
represents a surgery in which no transformational behaviors occurred, a score of 0.50 represents a surgery where half of a surgeon’s leadership behaviours were transformational and the other half were a combination of other leadership behaviours, and a score of 1.0 represents a surgery where all of the surgeon’s leadership behaviours were transformational.¹⁷

**Mediating variables.** The mediating variables were measured using surveys completed by surgical team members after surgery. Psychological safety was measured using Edmondson’s (1999) 7-item scale (rated on a scale of 1=strongly disagree to 7=strongly agree), and included items such as “It is safe to take a risk on this team,” and “It is difficult to ask other members of this team for help” (reversed scored). Boredom was measured using a 6-item version of Drory’s (1982) boredom frequency scale, adapted for the surgical context. The questionnaire asked participants to rate the frequency during the surgery they felt, for example, “bored”, or “wanting the surgery to be over” on a scale of 1=never/almost never to 5=almost all the time. The complete questionnaire can be found in Appendix D.

**Moderating variable.** The moderator variable of surgical complexity was collected from the lead surgeon, lead anesthesiologist and lead nurse prior to each surgery, each of whom was asked separately to provide an overall rating of complexity of the surgery given the procedure and patient condition, on a scale of 1 to 10 (1=lowest, 10=highest). The three ratings were then averaged across the raters to create the linear complexity variable. Curvilinear complexity was then created by first centering the complexity variable, and then squaring it.

¹⁷ Observers were also asked to complete survey measures of each leadership variable following each surgery. They completed Beachamp, Barling, Morton, Keith, & Zumbo’s (2010) transformational leadership scale, Podsakoff et al.’s (1990) scale for laissez-faire leadership, and Dupre & Barling’s (2006) scale for over-controlling leadership. The correlations between these scale variables and the proportional measures are as follows: 0.38 (p < .01) for transformational leadership, 0.31 (p < .01) for laissez-faire leadership, and 0.37 (p < .01) for over-controlling leadership. These correlations all represent moderate correlations between the pairs of measures (Cohen, 1988).
**Outcome variables.** The outcome variables include the proximal surgical outcomes of errors, and blood loss, and the distal surgical outcomes of complications at discharge, and complications at follow-up. The outcome variables were selected by a group of surgeons and anesthesiologists helping with this project (who were not part of our sample). The team identified the outcomes as common metrics that (a) hospitals and practitioners would care about (enhancing ecological validity), and (b) where possible, included data that were collected and reported in patient files by default, or were not difficult to collect.

Errors (e.g., multiple attempts to insert an IV, cut resulting in blood loss, breaking the sterile field) were collected by observation (Appendix B). A medical error was defined as “an act of omission or commission in planning or execution that contributes or could contribute to an unintended result” (Grober & Bohnen, 2005, pg. 42); this definition separates the process of committing a medical error from an adverse event. Historically, researchers have relied on adverse events (e.g., Gawande, Thomas, Zinner & Brennan, 1999; Vashdi et al., 2013), but as Grober and Bohnen (2005) discuss, this omits those cases where no adverse event occurred as a result of error (e.g., breaking of the sterile field not resulting in infection from contamination), which contributes to the underrepresentation of actual medical errors. Thus, following Vashdi et al. (2013), observers were instructed to apply the definition above when identifying an error. Errors for each surgery were tallied, and then divided by the length of the surgery to create a rate of errors per hour.

Blood loss was computed as actual blood loss in cubic centimeters (cc’s), collected from patient records, minus the expected blood loss as reported by the surgeon prior to the surgery (Appendix B). Actual blood loss is routinely recorded and entered to patient records. Estimated blood loss, while not officially recorded, is a common estimate with high ecological validity that
is made by surgeons in estimating procedure risk, and when ordering blood prior to procedures. The difference was divided by the length of the surgery to create a rate of actual blood loss versus expected blood loss per hour, and divided by 10 to ease strain on the model. While this variable reflects “blood loss versus actual blood loss, per hour”, for simplicity it will be referred to as “blood loss”.

Complications at discharge (e.g., need for antibiotics, need for specialist consultation) represent early post-operative outcomes of concern to the medical community, and are indicators of performance during surgery. These were obtained from patient records (using the form in Appendix E), and summarized as a count variable.

Finally, complications at follow-up were collected via phone calls to patients (using the form in Appendix F) at least 30 days after their surgery (\( M = 55.1 \) days after, \( SD = 49.4 \)), and summarized as a binary variable of whether or not the patient experienced any complications (e.g., was prescribed antibiotics, made an unplanned visit to a medical professional) since they were discharged from the hospital (0 = no complication, 1 = complication).

**Control variables.** Numerous variables were collected for the purpose of statistical control. These include both patient-specific variables, and procedure-specific variables. Patient variables include gender (0 = male, 1 = female), year of birth, and ASA rating (a score on a scale of 1 to 5, which reflects the patient’s physical status, where 1 is a normal healthy patient, and 5 is a patient who is not expected to survive without the operation; American Society of Anesthesiologists, 1995). Procedure-specific variables include whether it was an emergency procedure (0 = regular, 1 = emergency), the type of anesthesia used (0 = general anesthesia, 1 = spinal or local anesthesia), and the department (general surgery, urology, obstetrics and gynecology). The department was binary coded into two variables, one for urology (1 = urology
department), and one for obstetrics and gynecology (1 = obstetrics and gynecology department), with a coding of zero for both representing a general surgery department case. Length of stay (in days) was collected to control for complications at discharge, and length of time between surgery and follow-up phone call (in days) was collected as a control variable for complications at follow-up.

Data collection. To understand how measures were amalgamated, the data collection procedure is detailed in chronological order.

Prior to the surgery, observers gathered the following information from patient records: Gender, year of birth, and ASA rating. Details of the surgery were also collected from records as to whether it was an emergency, the type of anesthesia used, and the department. The date of the procedure was noted. The surgeon was asked about the expected blood loss in cubic centimeters (cc’s) for the procedure. Finally, the lead surgeon, lead anesthesiologist and lead nurse were asked to provide their ratings of complexity.

Intra-surgery, leadership behaviours (i.e. transformational leadership, laissez-faire leadership, over-controlling leadership) and errors were recorded by observers. To maintain independence, observers were asked to stand far enough apart during the surgery so that they could view the surgery but not see what the other was recording (Darst et al., 1989). Observers noted the time, who performed the behaviour, and circled the behaviour (i.e. transformational leadership, laissez-faire leadership, over-controlling leadership, or error) they witnessed. The start and end times of the surgery were recorded.

Immediately following the surgery, observers distributed surveys to team members to collect information on the mediating variables (i.e. psychological safety, boredom). Information
on actual blood loss was collected from patient records (with the aid of the surgeon where necessary).

Date of discharge and data on complications during their hospital stay were recorded from patient records.

Finally, observers began to contact patients by phone 30 days following their procedure. No voicemails were left in accordance with hospital privacy policies, as a result of which in some cases (N = 48) observers attempted to contact patients numerous times without success. When contact was made, observers asked patients about possible complications they could have experienced. The date the follow-up was completed was also recorded.

### 3.3 Data Analysis

**Interrater Agreement and Reliability**

Prior to any analysis, all survey variables (i.e. the mediating variables) were first tested for interrater agreement, computed using $r_{WG(J)}$ (James, Demaree & Wolf, 1993). This statistic tests the variation around scale ratings given by different individuals (and incorporates a penalty for chance agreement) to ensure that individual ratings can be averaged to the group-level. The interrater agreement is reported in Table 3-1. Using the critical value benchmarks outlined by Dunlap, Burke and Smith-Crowe (2003), all interrater agreement values were of sufficient magnitude to justify averaging to the group-level. Scales were then tested for internal consistency using Cronbach’s alpha (Cronbach, 1951). All alphas are reported in Table 3-1, and were acceptable (i.e. all $\alpha > .70$). Scale items were then averaged together for the final variables for the analysis.

Complexity ratings were also tested for agreement prior to averaging together, using Brown and Hauenstein’s (2005) $a_{WG}$, reported in Table 3-1. The $a_{WG}$ is similar to the $r_{WG(J)}$, in
that it measures agreement compared to chance agreement, but is more appropriate for ratings of
a single stimulus (i.e. complexity) rather than multiple ratings (i.e. scale items). While the
authors warn against the use of cutoff points, they do note that values less than 0.59 should be
considered unacceptable, especially for averaging separate ratings together. The $a_{WG}$ in this case
was above 0.59 (i.e. $a_{WG} = 0.79$), and so the three ratings were averaged together for a single
complexity variable.

Finally, inter-observer agreement of observations of leadership and errors were analyzed,
using Cohen’s weighted kappa (Cohen, 1968). Because the variables were ordered count data,
the weighted kappa is appropriate as it places a greater penalty on counts that are further apart
(Cohen, 1968). The final kappas are reported in Table 3-1. Results suggest that there was
“substantial agreement” (i.e. $kappa > 0.60$ (Landis & Koch, 1977) between observers for each of
the observed variables (transformational leadership = .67; laissez-faire leadership = .89; over-
controlling leadership = .73; errors = .78). Thus the counts by each observer were averaged
together for each variable, prior to computing the final proportion variables.

Analytic Strategy

Because the sample included the same surgeons in multiple surgeries, intraclass
correlations (ICC) of the mediating and outcome variables were calculated to examine whether a
portion of the variance in those variables could be explained by a “leader” effect. The ICCs are
reported in Table 3-1. An ICC that approaches zero suggests that multilevel modeling is not
necessary (e.g., Nezlek, 2008). The ICCs for the mediator variables deviated from zero,
indicating that part of their variance is explained by the grouping effect of the lead surgeon. The
ICC$s$ of many of the outcome variables were lower, but because all models include the
mediators, it is important to account for the higher-level variance in all models. Thus the
“complex” modeling option was selected in Mplus to control for a possible leader effect. The “complex” command is similar to multi-level modeling, and models variance at multiple levels of the data, thereby controlling for the grouped nature of the data (in this case the surgeon-level and then patient-level); it differs from pure multilevel modeling in that it does not allow specification of the level-two model beyond stating the clustering variable. The “complex” analysis is appropriate for multilevel data where the multilevel nature of the data is not conceptually relevant and therefore not a formal part of the design, but does need to be recognized and controlled statistically (MacKinnon, 2008; Preacher, 2015).

After accounting for clustering of the data, moderated mediation analyses were conducted to test all study hypotheses. All variables were centered prior to each analysis, except for binary (e.g., patient gender, follow-up complications) and count variables (i.e. discharge complications). All models controlled for seven variables: (1) whether the surgery was an emergency or not, (2) urology surgery or not, (3) obstetrics and gynecology case or not, (4) the patient’s gender, (5) patient’s year of birth, (6) patient’s ASA rating, and (7) whether the patient was under general anesthesia or anesthesia that left the patient awake. Additionally, the models for complications at discharge controlled for the patient’s length of stay, and the models for complications at follow-up controlled for the number of days that had passed since the patient’s procedure and any complications at discharge.

I used linear regression with maximum likelihood estimation, with robust standard errors. Poisson regression models were used for models including discharge complications (as this was a count variable), and logistic regression models were used for follow-up complications (as this was a binary outcome). Models were run separately for each mediating variable, and for each leadership variable, to increase the interpretability of the results.
Each model included two simultaneous regressions: The first represents the a-path of the mediation (the regression of the leadership variable, control variables, and interaction terms on the mediating variable; Figure 3-1) and the second represents the b- and c’-path of the full model (the regression of the leadership variable, mediating variable, control variables and interaction terms on the outcome variable; Figure 3-1). Where both mediating paths were significant, the indirect moderated mediation effect was tested at different levels of complexity. This was accomplished using a model constraint in Mplus that computed and tested the indirect effect of leadership on the outcome variable at the mean level of complexity, and at one standard deviation above and below the mean level of complexity.

Next, where possible, the proportion reduction in variance (PRV) was calculated for the leadership effects, and the combined leadership and complexity effects. PRV is a generally accepted effect size measure for use in multilevel modeling (Raudenbush & Bryk, 2002; Singer & Willett, 2003), calculated by comparing the residual variance in the model before and after a predictor variable is included. While it is an accepted measure, it should be noted that since estimating effect sizes in multilevel models is not straightforward, the results of this analysis are not comparable in the same sense that traditional effect size measures are (e.g., Cohen’s $d$, eta) (Peugh, 2010).

Finally, all significant moderation effects were probed. As explained by Hayes and Matthews, “when an interaction is found, it should be probed in order to better understand the conditions (i.e., the values of the moderator) under which the relationship between the focal predictor and the outcome is strong versus weak, positive versus negative, and so forth” (2009, p. 924). Probing was accomplished by modeling the interactions at a variety of levels of

---

18 Since the models for complications at discharge and complications at follow-up were modeled using a poisson and logistic distribution, respectively, it is impossible to calculate PRV.
complexity, again using a model constraint in Mplus, and the simple slopes were plotted. This procedure modeled total effect sizes at different levels of complexity, and tested the significance of these effect sizes.

3.4 Results

Preliminary Analyses

Descriptive statistics of all study variables (prior to centering) can be found in Table 3-1. A correlation table is reported in Table 3-2, but please note that this does not account for the multilevel leader-effect, and should only be used for a general overview of the data.

While the data gathered from surgeries and discharge is complete, follow-up information collected via patient phone-calls is missing approximately one-third of participants. To ensure that these data were not biased in any way, an attrition analysis was performed to test whether attrition at follow-up was random. Using a one-way ANOVA, patients choosing not to participate at follow-up, or who could not be contacted, did not differ from respondents for whom complete data were available, in terms of complexity, ASA rating, errors, blood loss, and complications at discharge. In addition, no significant results emerged with respect to the leadership behaviors to which the two groups had been exposed. However, both discharge and follow-up complications data would be underestimated as two patients were still hospitalized at follow-up (and their data would likely be outliers).

Main Analysis

Results of the final analysis are presented in Tables 3-3 through 3-5. A cursory examination of the results highlighted a potential issue relating to boredom. No leadership behaviour was related to boredom (Table 3-6), and thus boredom could not mediate the relationship between leadership and surgical outcomes (Hypothesis 4). One probable reason for
this is range restriction in the boredom variable. The low mean (1.59 on a scale of 1 to 5), low standard deviation (0.30), and low range (1.00 - 2.47) suggest a high degree of range restriction, and a “floor effect”, among individuals within teams. Tellingly, only five of the over 800 staff surveys reported an average boredom greater than 3. For comparison, boredom in the first sample had a higher mean and standard deviation ($M = 2.02$, $SD = 0.7$), and while conducted in very different contexts, other research suggests boredom follows a normal distribution with a centered mean (e.g., Drory, 1982; Gugliemi, et al., 2013). Thus, interpretation of the results relating to boredom is not appropriate, and these results have been removed from the main analysis tables. Thus, each table shows the results of one leadership behaviour on each of the surgical outcomes (i.e. the proximal outcomes of errors and blood loss; the distal outcomes of complications at discharge and complications at follow-up) through psychological safety. Each hypothesis is tested against these final models.

**Leadership and surgical outcomes.** Hypothesis 1 predicted the direct effect of leadership on proximal and distal surgical outcomes$^{19}$. Hypothesis 4 predicted that complexity would moderate the relationship between leadership and surgical outcomes, with the effect of leadership on surgical outcomes stronger for low complexity and high complexity surgeries. This is illustrated in the $c'$ path in Figure 3-1. Given that these hypotheses complement each other, and an overall relationship cannot be discerned without considering all moderated effects, they are reviewed simultaneously.

**Transformational leadership and surgical outcomes.** Hypothesis 1a predicted that transformational leadership would positively predict surgical outcomes. Results of the analysis

---

$^{19}$ Please note that this study’s conceptualization of leadership incorporated consistency of leadership behaviours. As such, all significant results can be interpreted as “the more consistently a surgeon behaved as [transformational/laissez-faire/over-controlling], the [better/worse] the surgical outcome”.

205
Transformational leadership did not predict either of the proximal surgical outcomes (i.e. errors, blood loss), nor the distal outcome of complications at discharge. Transformational leadership only predicted complications at follow-up. The interaction between transformational leadership and linear complexity predicted complications at follow-up ($\beta = -3.50, p < .01$). Upon probing, this relationship was only significant at lower levels of complexity (see Fig. 3-2 and Table 3-7); transformational leadership was positively related to complications at follow-up at lower levels of complexity, this result for low complexity is contrary to what was predicted. In summary, there was mixed support for Hypothesis 1a and Hypothesis 4.

**Laissez-faire leadership and surgical outcomes.** Table 3-4 reports the results of the effects of laissez-faire leadership on proximal and distal surgical outcomes. No support was found for the relationship between laissez-faire leadership and the proximal outcome of errors. However, consistent with Hypothesis 1b, laissez-faire negatively predicted the proximal outcome of blood loss ($\beta = -17.87, p < .01; \text{PRV} = 1.6\%$), and this relationship was moderated by both linear complexity ($\beta = 11.81, p < .05$), and curvilinear complexity ($\beta = 11.11, p < .01$; total $\text{PRV} = 7.4\%$). The total effect of the interaction is presented in Figure 3-3. Probing this relationship (Table 3-8) shows that for lower to mean levels of complexity, there is a negative relationship between laissez-faire leadership and blood loss; however, at high levels of complexity there is a positive relationship between laissez-faire leadership and blood loss, thus partially supporting Hypothesis 1b and Hypothesis 5.

Laissez-faire leadership also predicted the distal surgical outcome of complications upon discharge, and linear complexity moderated this relationship ($\beta = -3.16, p < .05$). Figure 3-4 illustrates the relationship and Table 3-9 probes the effects. Results show a positive relationship between laissez-faire leadership and discharge complications for lower complexity surgeries,
providing partial support for Hypothesis 1b and Hypothesis 5. Finally, there was no support for the other distal outcome of complications at follow-up.

In summary, there were conflicting findings regarding the effects of laissez-faire leadership on both proximal and distal surgical outcomes: In some instances it negatively impacted surgical outcomes (e.g., blood loss in highly complex cases, complications at discharge), and in others it positively predicted the outcomes (e.g., reduced blood loss in lower complexity cases).

**Over-controlling leadership and proximal and distal surgical outcomes.** Hypothesis 1c stated that over-controlling leadership would negatively predict surgical outcomes. The results are reported in Table 3-5. No support was found for the effect of over-controlling leadership on the proximal outcome of errors. Over-controlling leadership predicted the proximal outcome of blood loss ($\beta = 13.70, p < .05; PRV = 3.7\%$) and linear complexity moderated the relationship between over-controlling leadership and blood loss ($\beta = 5.23, p < .05; \text{total } PRV = 4.1\%$). This total relationship is illustrated in Figure 3-5. Probing shows that there is a positive relationship between over-controlling leadership and blood loss for all but the lowest levels of complexity (Table 3-10). The relationship gets stronger as complexity increases, providing partial support for Hypothesis 1c and Hypothesis 4.

Over-controlling leadership also predicted the distal outcome of number of complications upon discharge. However, this effect was in the opposite direction than was predicted ($\beta = -4.42, p < .05$), and was not moderated by complexity. Specifically, over-controlling leadership was associated with a decreased incidence of complications while patients were still hospitalized.

Finally, over-controlling leadership predicted the distal outcome of complications at follow-up, and linear complexity moderated this relationship ($\beta = 4.77, p < .05$; see Figure 3-6.
and Table 3-10). There was only a significant positive relationship between over-controlling leadership and complications for high complexity cases.

In summary, Hypothesis 1c and Hypothesis 4 that over-controlling leadership predicts surgical outcomes moderated by complexity were supported, as over-controlling leadership predicts blood loss and complications at follow-up, at higher levels of complexity. Contrary to predictions, however, over-controlling leadership was related to a lower incidence of complications at discharge, and this relationship was not moderated.

**Mediation.** Hypothesis 2 predicted that psychological safety would mediate the relationship between leadership and surgical outcomes. Hypothesis 5 predicted that any such mediation effects would be moderated by the complexity of the surgery. These are illustrated on the a- and b-path in Figure 3-1. Like Hypothesis 1 and 4, these hypotheses will be examined simultaneously, as any main effects cannot be interpreted without consideration of moderating effects.

Hypotheses 2a-c stated that transformational leadership, laissez-faire leadership, and over-controlling leadership, respectively would predict psychological safety, and Hypothesis 5a-c predicted that these relationships would be moderated by complexity. Hypothesis 2a predicted that transformational leadership would predict psychological safety, but this hypothesis was not supported (see Table 3-3). Thus Hypothesis 2a and Hypothesis 5a were not supported. Hypothesis 2b predicted that laissez-faire leadership would be related to psychological safety, and Hypothesis 5b predicted that this relationship would be moderated by complexity. Table 3-4 shows the results of the laissez-faire leadership models, which yielded no significant predictors for psychological safety.
Finally, Hypotheses 2c and 5c predicted that over-controlling leadership would predict psychological safety, and that the relationship between over-controlling leadership and psychological safety would be moderated by complexity. Results (see Table 3-5) show that over-controlling leadership did predict psychological safety, as the relationship between over-controlling leadership and psychological safety was moderated by curvilinear complexity (main models: $\beta = -0.66, p < .01$; follow-up model: $\beta = -0.56, p < .01$; total $PRV = 14.8\%$). This interaction is illustrated in Figure 3-7, and probed in Table 3-11. Results show a negative relationship between over-controlling leadership and psychological safety for both low and high complexity surgeries, providing some support for Hypotheses 2c and 5c.

Hypothesis 2d stated that psychological safety would predict proximal and distal surgical outcomes, and hypothesis 5d predicted that these relationships would be moderated by complexity. Since neither transformational leadership nor laissez-faire leadership predicted psychological safety, it is not necessary to examine the results of psychological safety in these models, as psychological safety could not mediate these relationships (Hypotheses 2d-e). Thus, examination of results is limited to the results of over-controlling leadership (Table 3-5). There was a direct effect of psychological safety on the proximal outcome of errors ($\beta = 0.24, p < .05$), and complexity also moderated the relationship between psychological safety and errors in a curvilinear manner ($\beta = -0.18, p < .05$; total $PRV = 9.4\%$). Interpretation of this moderated effect is withheld until the overall mediating effects are tested below.

Included at the bottom of Table 3-5 is a test of the mediating effect for Hypothesis 2g. Results show that psychological safety did not mediate the relationship between over-controlling leadership and errors, at any level of complexity (Table 3-5), failing to support Hypothesis 2g. While both paths of the mediation relationship were significant (i.e. path a: over-controlling
leadership predicted psychological safety, and path b: psychological safety predicted errors), psychological safety did not mediate the relationship between over-controlling leadership and errors. Given this, the impact of psychological safety on errors is not explored further.

In summary, neither transformational leadership nor laissez-faire leadership predicted psychological safety, thus psychological safety cannot mediate the relationship between transformational leadership or laissez faire leadership and any surgical outcome. Over-controlling leadership predicted psychological safety, but statistical tests show that psychological safety was not a mediator of the relationships between over-controlling leadership and surgical outcomes.

Supplementary Analyses

A number of supplementary analyses were undertaken and are presented here.

Leadership and blood loss. Since the variable “blood loss” is composed of both surgeons’ expected blood loss and actual blood loss, one possible reason for the findings obtained is that reports of expected blood loss could be related to leadership. Specifically, reports of expected blood loss may be differentially related to different types of leadership. For example, over-controlling leaders could be more confident in their abilities, and not expect as much blood loss, which when compared to actual blood loss makes it appear that they are performing worse. To test this possibility, expected blood loss was compared across the types of leadership using a correlation analysis. No significant correlations emerged between blood loss and transformational ($r = .10, p > .05$), laissez-faire ($r = -.04, p > .05$), or over-controlling leadership ($r = -.07, p > .05$).

Over-controlling leadership and complications at discharge. The negative relationship between over-controlling leadership and complications at discharge was opposite to
predictions. In order to further understand this, the regression was re-run controlling for blood loss versus expected blood loss (the variable used in the main analysis), which was positively predicted by over-controlling leadership in the regression analysis. Controlling for blood loss could explain the variance in complications at discharge, and help explain the results. After controlling for blood loss, complexity moderated the relationship between over-controlling leadership and complications at discharge in a curvilinear manner ($\beta = 2.80, p < .05$). Further probing showed that over-controlling leadership was negatively related to complications for low to medium levels of complexity, but positively related to complications in high-complexity cases (Figure 3-8; Table 3-13).

3.5 Discussion

My study explored the effects of leadership in the extreme context of an operating room. The operating room provided an interesting setting in which to study leadership, as it is a setting that is highly-structured, and makes use of transient teams. I expanded the conceptualization of leadership to focus on the measure of consistency of leadership behaviours. The direct effects of the consistency of transformational, laissez-faire, and over-controlling leadership on proximal and distal surgical outcomes were tested. Mediation effects of psychological safety and boredom were also explored, as were the moderating effects of surgical complexity. The model was tested using a combination of observational, survey, and objective archival measurements of study variables. Results demonstrate that leadership does indeed play a role in the achievement of positive surgical outcomes, and that complexity is important to consider. However, none of the mediational hypotheses were supported. The results are discussed below, followed by a consideration of the study strengths and limitations, and areas of future research and practical implications.
The Effect of Leadership in the Operating Room

My study was the first to use an observed measured of leadership, and was also one of the first to examine the consistency of different leadership behaviours. This conceptualization of leadership provided a unique perspective on leadership that is closer to a true experience of leadership in any organization. Results of the analysis demonstrated that the consistency of each leadership behaviour examined (i.e. transformational, laissez-faire, over-controlling) yielded significant effects on surgical outcomes, and are discussed below.

Transformational leadership. Transformational leadership was hypothesized to positively impact both proximal (i.e. errors, blood loss), and distal (e.g., complications at discharge, complications at follow-up) surgical outcomes. However, transformational leadership only predicted one surgical outcome, follow-up complications, and only at a low level of complexity. Moreover, the result was not in the predicted direction: Instead, transformational leadership was related to a greater chance of complications occurring after lower complexity surgeries. However, the protracted time between the actual leadership behaviour during surgery and the outcome (i.e. at least 30 days post surgery) could allow for patient behaviour to play a role. In particular, I suggest that the positive relationship between transformational leadership and follow-up complications could be capturing a patient effect of transformational leadership: The same behaviours a surgeon uses in the operating room are likely also used during interactions with patients which could elicit greater trust in the physician, and a greater willingness to report any experienced complications.

Findings from different areas support this interpretation. First, empathy, a part of the individualized consideration component within transformational leadership, is related to patients being more forthcoming with their symptoms and concerns (Halpern, 2001; Kim et al., 2004).
Second, findings from the study of unions also support my interpretation of the effects of transformational leadership on reported complications. Specifically, research of coal mining has shown that following unionization, and the increased trust by employees that they are no longer as vulnerable to reprisals by management, reports of injuries increase (e.g., Morantz, 2013). Thus, transformational surgeons who foster trust in their patients may see an increase in the “reporting” of complications.\(^{20}\)

Transformational leadership did not predict any other outcomes (i.e. errors, blood loss, complications at discharge). This was unexpected given the pervasiveness of positive and significant results of transformational leadership on various individual and organizational outcomes (Barling et al., 2010). However, publication bias, social science’s “aversion to the null” (Ferguson & Heene, 2012) and study authors tendency only to present the most significant or compelling results of their research, leaving outcomes without significant results by the wayside, has been noted, and null findings may not be surprising. By including null results, a more realistic understanding of the effects of leadership may be gained.

Another possible reason for the null findings is that transformational leadership may be the behavioural norm in this hospital. If positive behaviours are expected by staff, the effects of that leadership may fail to have further significant effects. This explanation is also consistent with a potential ceiling effect, where a restriction of range attenuates correlations, leading to insignificant results. Examining the data, it can be seen that on average there is a large proportion of transformational leadership behaviours ($M = .87, SD = 0.26$, range=0.0 to 1.0). This suggests

---

\(^{20}\) A discussion with a practicing surgeon confirms the possibility of this theory.
that it is perhaps the incidence of the negative leadership behaviours (i.e. laissez-faire and over-controlling leadership) that have negative impacts on surgical outcomes.\(^{21}\)

**Laissez-faire leadership.** The proportion of laissez-faire leadership predicted one proximal surgical outcome (i.e. blood loss), and one distal surgical outcome (i.e. complications at discharge). The effect of laissez-faire leadership on blood loss was moderated by complexity. As predicted, for higher complexity surgeries, there was a positive relationship between laissez-faire leadership and blood loss. Passive leadership behaviours allow an environment where team members are less cautious, and perform poorly with no consequences. This extends findings from traditional organizations that laissez-faire leadership has a detrimental impact on performance (e.g., Hinkin & Schriesheim, 2008), and on workplace injuries and accidents (e.g., Mullen et al., 2011), and supports the notion that passive leadership has a detrimental impact in extreme contexts.

However, contrary to predictions, there was a negative effect of laissez-faire leadership on blood loss for lower to mean complexity surgeries. This does not support the evidence on laissez-faire leadership’s effects on performance and safety outcomes noted above, and instead raises questions about the impact of laissez-faire leadership under lower complexity conditions. One possible explanation is that under conditions of low to medium complexity, the conditions inherent in the situation (i.e. clear roles, clear procedures, standardized behaviours) may be

\(^{21}\) A final possibility for the null results of transformational leadership is that this study used observed consistency of transformational leadership, opposed to survey measures of transformational leadership. Transformational leadership is predominantly measured using survey measures, with observational data comprising less than 4% of all leadership research (Hiller, DeChurch, Murase, & Doty, 2011). This difference in method could help explain the null findings, and to eliminate this possibility, observers were asked to rate surgeons using Beauchamp et al.’s (2010) survey after each surgery, and this measure was correlated with the observed proportional measure at 0.38 (\(p < .01\)), representing a moderate effect size (Cohen, 1988).
sufficient by themselves to ensure adequate performance, rendering leadership influence less important. Low to medium complexity conditions might be considered strong situations (e.g., Cooper & Withey, 2009), explaining why any effects of leadership could be different within extreme contexts. This pattern emerged in Katz-Navon, Naveh and Stern’s (2007) study, which showed that across a sample of hospitals, safety-self-efficacy was related with increased patient safety, only when standardization was low. In the case of leadership, strong situations and standardization could be considered a substitute for leadership (e.g., Kerr & Jermier, 1978); the substitutes for leadership model posits that situational variables can substitute, neutralize, or enhance leadership behaviours, which have been demonstrated to explain more variance in traditional organizational outcomes (e.g., job attitudes, performance) than leadership behaviours (Podsakoff, MacKenzie, & Bommer, 1996).

In contrast, laissez-faire leadership positively predicted complications at discharge for low and mean levels of complications. Laissez-faire leadership was related to more complications while still under hospital care, but this effect was weaker as complexity increased, and became insignificant for higher levels of complexity. The positive relationship between laissez-faire leadership and complications at discharge extends findings on the negative effects of laissez-faire on performance and safety (noted above), showing that laissez-faire leadership has negative effects on longer-term performance outcomes in the extreme context of the operating room. However, this positive effect between laissez-faire leadership and complications is opposite to the negative relationship between laissez-faire leadership and blood loss found for similarly lower complexity surgeries. This raises another potential scenario for the favourable relationship found for laissez-faire leadership and blood loss under lower complexity: Taken together, these two findings might suggest that laissez-faire leadership encourages taking short-
cuts or risks during surgery that result in less blood loss, and in the longer-term, symptoms of underlying issues emerge in the form of greater complications during recovery.

In summary, results were mixed on the effects of laissez-faire leadership on proximal surgical outcomes, as laissez-faire leadership was positively related to blood loss in higher complexity surgeries, but negatively related to blood loss in low to medium complexity surgeries. However despite this favourable result for blood loss in low to medium complexity surgeries, laissez-faire leadership was positively related the distal surgical outcome of complications at discharge, also for lower complexity surgeries.

**Over-controlling leadership.** Of the three leadership behaviors, over-controlling leadership had the most widespread significant results on surgical outcomes, predicting the proximal outcome of blood loss, and both distal outcomes of complications at discharge and follow-up.

Over-controlling leadership was positively related to blood loss. This finding was in line with the hypothesized relationship, extending the limited literature suggesting that over-controlling leadership has undesirable organizational outcomes (e.g., Dupre & Barling, 2006), and the wider literature showing a lack of workplace control is detrimental to performance (e.g., Dupre et al., 2004). However, this relationship was moderated by complexity of the surgery: This relationship was higher as complexity increased. Thus, the moderation was not in the predicted direction. I predicted that the relationship between over-controlling leadership and blood loss would be stronger for both low and high levels of complexity (and weaker for mean level complexity). Instead, over-controlling leadership only predicted blood loss as complexity (and presumably ambiguity) increased. Like laissez-faire leadership, I suggest that surgical
complexity may act as a proxy for situation strength, such that lower complexity is characteristic of stronger situations, in which leadership behaviours may be less important to outcomes.

Over-controlling leadership was related to fewer complications at discharge, which is again contrary to what I predicted. This relationship was investigated in a post-hoc analysis, controlling for blood loss because patients who have unexpected amounts of blood loss (i.e. blood loss in excess of expectations) during surgery may receive more attention by hospital staff while in recovery, and therefore experience fewer complications. Because over-controlling leadership was positively correlated with blood loss, patients of over-controlling leaders may need or receive extra care during recovery, and again experience fewer complications while still hospitalized. Results of this analysis showed that controlling for blood loss, over-controlling leadership predicted complications at discharge, and complexity moderated this relationship in a curvilinear manner: In line with my original hypothesis, over-controlling leadership predicted more complications at discharge in higher complexity surgeries after controlling for blood loss. However, a negative relationship emerged between over-controlling leadership and complications at mean levels of complexity. This result remains puzzling, as I had expected that while the effect might be weaker, it would still be positive. Padilla, Hogan and Kaiser (2007) suggested that destructive leadership does not always have destructive outcomes, especially in systems with strong checks and balances that parallel strong situations. These authors noted that destructive leadership could even be associated with positive outcomes, especially in the shorter-term, and that it is the long-term ramifications that make the leadership behaviours “destructive”. This could be an example of such a situation (i.e. a strong situation, average length of stay of 3.7 days), but further research is needed to understand this relationship.
Last, over-controlling leadership was positively related to long-term complications, for higher levels of surgical complexity. This is consistent with the hypothesized relationship between over-controlling leadership and surgical outcomes, again supporting the detrimental effects of over-controlling leadership behaviours, and extending earlier findings on the negative effects of a lack of control on performance and other organizational outcomes. However, as was the case for over-controlling leadership and blood loss, the moderating effects of complexity was contrary to what was predicted. I had predicted that the relationship between over-controlling leadership and complications would be greater in both lower and higher complexity surgeries, but results show this relationship only at higher levels of complexity. Thus, as was the case regarding blood loss, over-controlling leadership matters most in higher complexity surgeries.

In summary, the impact of over-controlling leadership on proximal and distal surgical outcomes is mixed. The complexity of the surgery is a crucial factor in the relationship between over-controlling leadership and surgical outcomes. In higher complexity surgeries, over-controlling leadership predicted more blood loss, more complications at discharge (when controlling for blood loss), and more complications at follow-up. Further research is needed to understand the relationship between over-controlling leadership and complications at discharge for lower complexity surgeries, which was a negative relationship in this sample.

**Consistency of leadership.** Finally, this study also aimed to incorporate consistency of leadership. In addition to the findings discussed above, the descriptive statistics of the leadership variables support Mullen et al.’s (2011) findings that leadership is not consistent, and the same leader can exhibit multiple behaviours. The findings of my study suggest that transformational leadership behaviours are the most common ($M = 0.87, SD = 0.26$), but this also suggests that on average surgeons are also using some laissez-faire and over-controlling leadership behaviours.
The proportion of these negative behaviours are much lower (i.e. laissez-faire: $M = 0.03, SD = 0.12$; over-controlling: $M = 0.05, SD = 0.11$), but as the findings discussed above suggest, as the proportion of these behaviours increases (and thus the proportion of positive behaviours decreases), overall the more negative patient outcomes that can be expected. This is consistent with Hannah et al.’s (2010) supposition that consistency of positive leadership is important, as leaders only have minimal leeway to behave contrary before negative outcomes emerge.

**Mediating and Moderating Effects**

While over-controlling leadership was negatively related to psychological safety in low and high complexity cases, overall psychological safety did not mediate the relationship between leadership and surgical outcomes. This is surprising given its role as a mediator in prior research (e.g., Edmondson, 1999; Nembhard & Edmondson, 2006). To try to understand why this is the case, I will look separately at the leadership behaviour -> psychological safety path, and psychological safety -> surgical outcomes path.

Neither transformational leadership nor laissez-faire leadership predicted psychological safety, contributing to the non-mediation in these cases. Only over-controlling leadership predicted psychological safety. As discussed earlier, transformational leadership was not a predictor of the majority of study variables, and this could be because transformational leadership is the normative leadership style ($M = 0.88, SD = 0.26$) in this particular hospital. If staff members are accustomed to transformational leadership behaviours, it may only be deviations from these behaviours that negatively impact psychological safety, rather than positive leadership behaviours enhancing psychological safety. In this case it appears that when transformational leadership behaviours are replaced by over-controlling leadership behaviours psychological safety suffers. Laissez-faire leadership behaviours may have no effect, as in such
strong situations the absence of leadership does not make teams feel less psychologically safe. Perhaps it is only the active, over-controlling leadership behaviours that lead team members to question whether the team is safe for interpersonal risk-taking.

For the second part of the mediation, psychological safety only predicted one surgical outcome, namely errors. While the investigation of psychological safety predicting errors was not a focus of the analysis, it warrants discussion. Psychological safety only predicted errors, and the relationship was opposite to what was expected (i.e. psychological safety was positively related to errors, moderated by complexity such that the relationship was weaker as complexity increased). While this outcome was not hypothesized, it is not unprecedented: Psychological safety leads to increased reporting of errors (e.g., Edmondson, 1996). Beyond errors, psychological safety did not predict any surgical outcomes. While psychological safety is related to performance (e.g., Baer & Frese, 2003), the surgical outcomes were not performance measures.

Finally, psychological safety may not be an important variable, given that all nurses in my study are unionized, enhancing feelings of psychological safety. As noted earlier, unionization is associated with increased reporting of safety issues, as employees feel less susceptible to management reprisal (e.g., Morantz, 2013). So the organizational-level context could supersede the need for psychological safety.

The way in which boredom was measured resulted in a lack of variance and range restriction; as a result, predictions involving boredom as a mediator were not investigated further.

Linear and curvilinear complexity was a consistent moderator of the relationships between leadership behaviours and surgical outcomes. The initial hypothesis predicted that
leadership would matter more in lower complexity surgeries (to alleviate complacency; Hannah et al. 2009), and higher complexity surgeries (to deal with ambiguity). This would have manifested in results showing stronger relationships between leadership and surgical outcomes for lower and higher complexity surgeries. However, this hypothesis was not supported, as none of the leadership behaviours (i.e. transformational, laissez-faire, over-controlling) yielded similar effect sizes under low and high complexity surgeries.\(^{22}\)

Instead, different patterns of moderation emerged for different leadership behaviours. For transformational leadership, there was a positive relationship between transformational leadership and follow-up complications at low levels of complexity. For laissez-faire, there was a negative relationship between laissez-faire leadership and blood loss for low to mean levels of complexity, and a positive relationship under high complexity surgeries. For laissez-faire again, there was a positive relationship between laissez-faire leadership and complications at discharge for lower levels of complexity. For over-controlling leadership, over-controlling leadership was positively related to blood loss, complications at discharge, and complications at follow-up for high complexity surgeries, but over-controlling leadership was negatively related to complications at discharge for lower complexity surgeries. Possible explanations of each of these patterns were discussed above. None of the theorized explanations (e.g., patient effect, strong situations), however, applied consistently to each of the observed patterns. The findings therefore show that complexity is an important moderating influence in the case of surgery, but

\(^{22}\) While mediation was not significant, the predicted moderation pattern described here is illustrated in the relationship between over-controlling leadership and psychological safety (Figure 3-7): A significantly negative relationship between over-controlling leadership and psychological safety for low and high complexity, but an insignificant relationship at mean complexity.
how it moderates the effects of different types of leadership on different patient outcomes remains to be understood.

**Study Strengths**

This study benefitted from a number of conceptual and methodological strengths.

First, the moderating role of complexity advances our understanding of the effects of leadership in extreme contexts. As Johns (2006) noted, context is crucial to truly understanding leadership. Including complexity allowed a significant contextual variable to be examined in the relationship of leadership and performance.

Second, considering complexity also extended thinking beyond traditional linear relationships. Historically, organizational researchers have considered linear moderation almost exclusively (e.g., Ng & Feldmen, 2011). Linear moderation suggests the relationship between two variables gets either increasingly positive or increasingly negative in the presence of a moderator. In the extreme case of an “X”-shaped interaction, the relationship between two variables is negative under one moderator condition, zero at the mean, and positive at the other end. Evidence from my research shows that there is more to be gained by considering other moderating patterns. As described earlier, including a quadratic term allows for patterns where the effect size of a relationship does not steadily increase or decrease, and can reach an inflection point, where the effect size starts to change in the other direction. This allows a pattern where at the extreme, the relationship between two variables is the same at both extremes of the moderator, and zero at the mean. While the results of my study did not consistently conform to this pattern, examining many of the simple slopes demonstrates the futility of continuing to conceptualize solely in linear ways. For example, (see Table 3-7), the relationship between laissez-faire leadership and blood loss is positive at low complexity, becomes increasingly
negative until complexity is just below the mean, and the relationship inflects and becomes increasingly positive. If these values were plotted, an approximate “U”-shaped effect would emerge. By expanding thinking beyond linear effects, a more accurate picture of the moderating effect of complexity was gained.

Third, my research took a temporal view, examining both proximal and distal surgical outcomes. By using this temporal view, I was able to draw conclusions about the immediate and long-term effects of the behaviours that occur in the operating room.

Fourth, I introduced the concept of over-controlling leadership. While over-controlling leadership has been recognized by practitioners and scholars as a common leadership behaviour (e.g., micromanaging; White, 2010), my study was among the first to incorporate it into an empirical study. The mean proportion of over-controlling leadership of 0.05 (SD = 0.11) shows that, while it appears low in contrast to transformational leadership, it is a behaviour that does occur in the operating room, and the consistent significant results demonstrate that it is a leadership behaviour that matters in the operating room with negative effects on surgical outcomes.

Fifth, my study incorporated a new way to conceptualize leadership: consistency. Hannah et al., (2010) highlighted the need to incorporate consistency into leadership research, and Mullen et al., (2011) empirically demonstrated the advantages of doing so. Leaders do not behave consistently, and different leadership behaviours can be exhibited by a single leader. This study incorporated consistency by looking at the proportion of observed leadership behaviours of surgeons that were transformational, laissez-faire, and over-controlling. The predictor variables did not reflect transformational behaviors on a scale of 1 to 5, but instead represented how consistently leaders behaved transformational, in relation to other leadership
behaviours. By incorporating the notion of consistency, a more accurate representation of the true experience of leadership was captured.

Finally, this research study was completed in an extreme context, specifically the unique extreme context of the operating room, where transient teams are used, and work takes place in what can be described as a strong situation (where team members have highly prescribed roles and procedures). Given these factors, the model tested could be considered a robust test of the theory of leadership in extreme contexts, and finding significant relationships could have been difficult. But results demonstrate that leadership (i.e. transformational, laissez-faire, over-controlling) was still important, even in this extreme context.

Methodologically, this study benefitted from a number of strengths. First, data were collected from multiple sources, using multiple measurement techniques, reducing and eliminating the common sources of biases such as mono-method and mono-source data. Second, observations of leadership offered a more objective view of leadership, and demonstrated that observers can be trained to reliably collect leadership data. Third, outcome variables were collected across multiple time periods (i.e. errors and blood loss were collected immediately, complications were measured at discharge, and long-term complications were measured at least thirty days following surgery). Finally, the outcomes examined were chosen by practitioners (e.g., anesthesiologists, surgeons). This enhances the ecological validity of the results, as the outcomes impact the entire healthcare system.

Limitations

The findings of the current study should be considered given a number of limitations. First, my study was conducted in a teaching hospital (i.e. medical students and residents were often present in the room, and possibly performed a portion of the surgery). Teaching hospitals
differ from non-teaching hospitals in terms of patient outcomes (e.g. more complications, higher morbidity rates; Khuri et al., 2001), and anecdotally, the current study observers noted that behaviourally on the rare occasion where no students were present, interaction and discussion between team members was limited. Whether the current results extend to non-teaching hospitals remains to be assessed. Second, data were collected at a single hospital. This potentially weakens the generalizability of study findings. Replication is needed across different hospitals to account for possible cultural, political, social and other macro-level differences (e.g. Kellogg, Breen, Ferzoco, Zinner, & Ashley, 2006). Third, while the observation of errors and leadership behaviours provided a unique perspective, some behaviours and errors may have been missed by the observers given their physical location in the operating room (i.e., standing behind staff/away from the operating table so as not to interfere), potentially introducing measurement error. This could have resulted in an underestimation of the true number of surgical errors per hour. When outcome variables have measurement errors, model estimates may be inconsistent with true population values (Fuller, 2009). For the leadership variables, the observational data may have deviated from the true proportion of transformational, laissez-faire, and over-controlling leadership behaviours, and this is important because measurement errors in multilevel data create underestimated parameters (Kromrey et al., 2006). While I did not use multilevel modeling, the “complex” modeling technique used is mathematically comparable to multilevel modeling, and inferences about parameter sizes should be made with caution.

Fourth, for ethical reasons, observers and participants (both staff and patients) were informed of the nature of the study. Specific leadership behaviours were not disclosed to participants, and no hypotheses were disclosed to either observers or participants, but it remains possible that staff modified their behaviour. This could have resulted in a positively skewed
proportion of positive behaviours (i.e. transformational leadership, psychological safety-related
behaviours) observed. Additionally, the observers may have attempted to guess hypotheses, but,
with the exception of errors, all outcome variables were objective, collected after the surgery,
often on different days, and often by different observers. This minimizes the threat of this threat
to construct validity. Furthermore, the interrater reliabilities of the observers suggest that
observers were objective, and any potential bias in the observations of leadership as they were
being collected was minimal. Finally, despite the many methodological strengths of the study, I
used a non-experimental design as a result of which true causality cannot be inferred, and
reverse-causality remains a possible alternative (e.g., blood loss predicting over-controlling
leadership).

**Future Directions**

Findings from this study offer interesting avenues of future research, in terms of both
contectual and methodological development. First, it is clear that a consideration of complexity
in future research could yield interesting and applicable results for future research. Future
research in the medical environment should incorporate complexity, but beyond healthcare,
complexity in other extreme contexts (e.g., incorporate magnitude and probability of
consequences into law enforcement based on area crime rates) and traditional contexts (e.g.,
in incorporate measures of how interdependent different teams or organizations are with each other)
may have similar effects with regards to leadership.

Second, isolating effective mediators could be conceptually beneficial. For example,
trust might mediate any effects of each leadership behaviour on the surgical outcomes. Third,
there were a number of unsupported hypotheses about the effects of leadership on surgical
outcomes. Some reasons for these findings were discussed earlier, but future research might
benefit from expanding the types of surgical outcomes studied. Two ways in which this might be useful would be to ensure outcomes are more specific to the actual surgeries (e.g., outcomes specific to bypass surgery; Desai, 2015), or relying more on psychological or attitudinal measures (e.g., meaningfulness of work, empathy, subjective patient satisfaction and well-being).

Fourth, expanding the study of the effects of leadership within surgeries to the patient-doctor relationship could offer new insights. For example, the puzzling effects of transformational leadership on follow-up complications in this study might be explained by an understanding of the patient-physician relationship, as physician empathy affects the subsequent behaviours of those patients (e.g., being forthcoming with symptoms, being more likely to follow physician prescriptions and recommendations; Halpern, 2001; Kim et al., 2004).

Fifth, future research should further examine the measurement of consistency in the study of leadership. My research incorporated a focus on consistency by using a proportional measure of observed behaviours, but as Mullen et al. (2011) show, the use of interactions between scale measures of leadership also yielded unique results and insights. Future research should continue the use of either of these methods, and even consider comparing their effects with traditional measures (i.e. absolute scale scores) to determine the unique effects that each conceptualization captures.

Sixth, the need to expand beyond self-report data in organizational research has been noted for decades (e.g., Podsakoff & Organ, 1986), as self-reports are vulnerable to such problems as distortion, respondent affect, and social desirability. Observational measures such as those used here address these problems, as they reduce subjectivity and potential biases that arise in survey measures. Future research should consider the incorporation of observational
measures of leadership to broaden our understanding of leadership. This will also create helpful normative data with which to understand the true distribution of observed leadership behaviours.

Seventh, the need to untangle causal direction remains. For example, does blood loss occur as a result of over-controlling leadership, does over-controlling leadership occur as a result of blood loss, or is there a reciprocal, spiraling effect? Longitudinal research, even longitudinal data within a surgery, could help untangle such relationships, combined with advanced statistical techniques, such as cross-lagged models.

Eighth, the observers noted that on the rare occasion that there were no students or residents in the operating room, there were far fewer observed leadership behaviours and interactions between team members. As a result, it would be beneficial to conduct similar research in a non-teaching hospital.

Ninth, I have relied on the theory of situation strength throughout this study, by making inferences about the setting being studied using logic (e.g., roles, hierarchy, standard procedures), and by applying the research of others who have also suggested this context is strong (e.g., Healy & McKay, 2000; Katz-Navon et al., 2007). However, no data were available to directly test the role of situational strength. Future research would benefit from explicitly testing the effect of situation strength in this context, or other extreme contexts, as it could be a valuable construct to explain behaviour and effect sizes.

Finally, while this study was one of the first to examine leadership in extreme contexts, especially in situ during a surgery, it is unclear whether or not the team members performing the surgery experienced their situation as “extreme”. The average professional tenure of a surgeon in this study was over 23 years, and it is possible that they subjectively experienced many of the procedures they engaged in as routine, and no different than sitting at a desk entering numbers.
into Excel. If this is the case, these surgeries would not constitute a strong test of the role of extreme contexts. Future research could incorporate the subjective experiences of the context, by first conducting qualitative research with surgeons about what constitutes an extreme context, and then by testing only a subset of surgeries that have been deemed “extreme”.

**Implications**

In the month following my study, Canada’s blood supply hit critically low levels, prompting a campaign to increase donations lest patient treatment be delayed (Teotonia, 2014). One way to avoid this from re-occurring is to reduce the demand for blood: If surgeons are using less blood, the strain on the system will be reduced. The results of my study suggest that this might be achieved through leadership training for surgeons. By reducing levels of laissez-faire and over-controlling leadership styles, overages on the amount of blood used during surgery could be lessened. Furthermore, the entire hospital system may be less strained as high-risk patients encounter fewer complications during their recovery. The efficacy of leadership training has been demonstrated in organizations by researchers (e.g., Barling, Weber & Kelloway, 1996), and the healthcare environment is one that could benefit from such training.

**Conclusion**

In this examination of different types of leadership in the extreme context of the operating room, leadership influenced the medically-relevant performance outcomes of blood loss and complications. Negative leadership behaviours (i.e. laissez-faire and over-controlling leadership) in the operating room have significant detrimental impacts to patient health. Furthermore, context was important, as surgical complexity moderated many of the relationships between leadership behaviours and surgical outcomes, showing that leadership has different effects as the conditions of the extreme context change. While these findings do need to be
replicated, the results of this research of the effects of leadership in this extreme context could reduce some of the risks of anyone facing surgery.
3.6 References

http://www.asahq.org/Home/For-Members/Clinical-Information/ASA-Physical-Status-Classification-System.


Steele, J. P. (2011). *Antecedents and consequences of toxic leadership in the US Army: A two year review and recommended solutions*. Center for Army Leadership, Fort Leavenworth, KS.


Sweeney, P. J. (2010). Do soldiers reevaluate trust in their leaders prior to combat operations? *Military Psychology, 22* (Suppl. 1), S70-S88.


Table 3-1: Descriptive statistics, reliability, and inter-rater agreement

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Scale</th>
<th>IRA</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergency case</td>
<td>150</td>
<td>0.16</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urology dept</td>
<td>150</td>
<td>0.13</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OB-Gyn dept</td>
<td>150</td>
<td>0.27</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>150</td>
<td>0.53</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YOB</td>
<td>150</td>
<td>1963.15</td>
<td>22.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASA</td>
<td>149</td>
<td>2.45</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anesthesia</td>
<td>150</td>
<td>0.19</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>150</td>
<td>4.52</td>
<td>1.47</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFL</td>
<td>150</td>
<td>0.87</td>
<td>0.26</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF</td>
<td>150</td>
<td>0.03</td>
<td>0.12</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC</td>
<td>150</td>
<td>0.05</td>
<td>0.11</td>
<td>0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psych. Safety</td>
<td>150</td>
<td>5.65</td>
<td>0.53</td>
<td>0.84</td>
<td>0.74</td>
<td>0.63</td>
</tr>
<tr>
<td>Boredom</td>
<td>150</td>
<td>1.59</td>
<td>0.29</td>
<td>0.91</td>
<td>0.92</td>
<td>0.60</td>
</tr>
<tr>
<td>Errors/hour</td>
<td>150</td>
<td>0.37</td>
<td>0.62</td>
<td>0.78</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Bloodloss/hour vs. expected</td>
<td>143</td>
<td>-1.93</td>
<td>7.46</td>
<td></td>
<td></td>
<td>0.09</td>
</tr>
<tr>
<td>Length of stay</td>
<td>150</td>
<td>3.65</td>
<td>7.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complications - discharge</td>
<td>150</td>
<td>0.17</td>
<td>0.41</td>
<td></td>
<td></td>
<td>0.36</td>
</tr>
<tr>
<td>Follow-up days</td>
<td>102</td>
<td>55.11</td>
<td>49.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complications - follow-up</td>
<td>102</td>
<td>0.47</td>
<td>0.50</td>
<td></td>
<td></td>
<td>NPD</td>
</tr>
</tbody>
</table>

Note: IRA for leadership and errors variables calculated using weighted Cohen's Kappa
Note: IRA for complexity calculated using Brown and Hauenstein’s (2005) \( \alpha_{WG} \)
Note: IRA for psych. safety, and boredom calculated using James, Demaree & Wolf’s (1984) \( r_{WG(J)} \)
Table 3.2: Correlation table

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergency case</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urology dept</td>
<td>-17*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OB-Gyn dept</td>
<td>-22**</td>
<td>-22**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-14</td>
<td>-25**</td>
<td>.56**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YOB</td>
<td>.05</td>
<td>-26**</td>
<td>.18*</td>
<td>.02</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASA</td>
<td>.05</td>
<td>.03</td>
<td>-25**</td>
<td>-.17*</td>
<td>-.53**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anesthesia</td>
<td>.02</td>
<td>.18*</td>
<td>-.17*</td>
<td>-.03</td>
<td>-.15</td>
<td>.04</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>.12</td>
<td>-.09</td>
<td>-.19*</td>
<td>-.19*</td>
<td>-.26**</td>
<td>.39**</td>
<td>-.05</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity^2</td>
<td>.20*</td>
<td>-.09</td>
<td>-.08</td>
<td>-.01</td>
<td>.03</td>
<td>.10</td>
<td>.03</td>
<td>.20*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFL</td>
<td>-.06</td>
<td>.04</td>
<td>.04</td>
<td>-.04</td>
<td>.04</td>
<td>.05</td>
<td>-.03</td>
<td>.03</td>
<td>.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC</td>
<td>.11</td>
<td>.22**</td>
<td>-.06</td>
<td>-.04</td>
<td>-.01</td>
<td>-.02</td>
<td>.07</td>
<td>-.02</td>
<td>-.37**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF</td>
<td>-.08</td>
<td>-.13</td>
<td>-.02</td>
<td>-.03</td>
<td>-.03</td>
<td>.09</td>
<td>-.07</td>
<td>.13</td>
<td>-.04</td>
<td>-.36**</td>
<td>-.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psych. Safety</td>
<td>-.15</td>
<td>.19*</td>
<td>-.04</td>
<td>.00</td>
<td>-.07</td>
<td>.06</td>
<td>-.05</td>
<td>-.12</td>
<td>.15</td>
<td>-.02</td>
<td>.27**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom</td>
<td>.04</td>
<td>-.28**</td>
<td>-.05</td>
<td>.04</td>
<td>.04</td>
<td>.06</td>
<td>-.24**</td>
<td>.01</td>
<td>.04</td>
<td>-.01</td>
<td>.07</td>
<td>-.02</td>
<td>-.17*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Errors/hour</td>
<td>-.09</td>
<td>-.02</td>
<td>.09</td>
<td>.18*</td>
<td>.04</td>
<td>-.05</td>
<td>-.07</td>
<td>-.04</td>
<td>.12</td>
<td>-.06</td>
<td>-.02</td>
<td>.24**</td>
<td>.06</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood loss/hour vs. expected</td>
<td>.13</td>
<td>.04</td>
<td>-.10</td>
<td>.12</td>
<td>-.11</td>
<td>-.04</td>
<td>.01</td>
<td>.07</td>
<td>-.14</td>
<td>.06</td>
<td>.12</td>
<td>.04</td>
<td>-.10</td>
<td>.02</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of stay</td>
<td>.13</td>
<td>.04</td>
<td>-.17*</td>
<td>-.14</td>
<td>-.20*</td>
<td>.16</td>
<td>.01</td>
<td>.20*</td>
<td>.10</td>
<td>-.06</td>
<td>.06</td>
<td>-.02</td>
<td>.04</td>
<td>.02</td>
<td>-.05</td>
<td>-.10</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complications - discharge</td>
<td>.31**</td>
<td>-.02</td>
<td>-.20*</td>
<td>-.22**</td>
<td>-.24**</td>
<td>.39**</td>
<td>-.06</td>
<td>.30**</td>
<td>.22**</td>
<td>.00</td>
<td>.06</td>
<td>.02</td>
<td>-.11</td>
<td>-.03</td>
<td>-.10</td>
<td>.00</td>
<td>.38**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Follow-up days</td>
<td>-.02</td>
<td>.09</td>
<td>-.18</td>
<td>-.04</td>
<td>-.10</td>
<td>.03</td>
<td>.07</td>
<td>.19</td>
<td>-.17</td>
<td>-.20*</td>
<td>-.37**</td>
<td>-.07</td>
<td>-.02</td>
<td>.06</td>
<td>.12</td>
<td>-.23*</td>
<td>.08</td>
<td>-.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Complications - follow-up</td>
<td>.08</td>
<td>.05</td>
<td>.02</td>
<td>-.01</td>
<td>-.25*</td>
<td>-.01</td>
<td>.18</td>
<td>.11</td>
<td>.21*</td>
<td>.04</td>
<td>-.11</td>
<td>.02</td>
<td>-.03</td>
<td>.06</td>
<td>-.10</td>
<td>.02</td>
<td>.28**</td>
<td>.13</td>
<td>.00</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Table 3-3: Model results of transformational leadership

<table>
<thead>
<tr>
<th>Mediator:</th>
<th>Proximal Surgical Outcomes</th>
<th>Distal Surgical Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Errors</td>
<td>Blood loss</td>
</tr>
<tr>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
</tr>
<tr>
<td><strong>Mediator:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psych Safety</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.03 (0.06)</td>
<td>0.03 (0.06)</td>
</tr>
<tr>
<td>Emergency</td>
<td>-0.16 (0.12)</td>
<td>-0.16 (0.12)</td>
</tr>
<tr>
<td>Urology</td>
<td>0.25 (0.22)</td>
<td>0.25 (0.22)</td>
</tr>
<tr>
<td>OB-Gyn</td>
<td>-0.09 (0.19)</td>
<td>-0.09 (0.19)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.05 (0.09)</td>
<td>0.05 (0.09)</td>
</tr>
<tr>
<td>YOB</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>ASA</td>
<td>-0.06 (0.05)</td>
<td>-0.06 (0.05)</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>0.02 (0.11)</td>
<td>0.02 (0.11)</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.00 (0.04)</td>
<td>0.00 (0.04)</td>
</tr>
<tr>
<td>Complexity^2</td>
<td>-0.02 (0.02)</td>
<td>-0.02 (0.02)</td>
</tr>
<tr>
<td>TFL</td>
<td>0.20 (0.18)</td>
<td>0.20 (0.18)</td>
</tr>
<tr>
<td>TFL*Cmplx</td>
<td>0.04 (0.23)</td>
<td>0.04 (0.23)</td>
</tr>
<tr>
<td>TFL*Cmplx^2</td>
<td>0.04 (0.09)</td>
<td>0.04 (0.09)</td>
</tr>
<tr>
<td><strong>Full Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.11 (0.09)</td>
<td>-0.31 (1.02)</td>
</tr>
<tr>
<td>Emergency</td>
<td>-0.07 (0.11)</td>
<td>2.59 (2.28)</td>
</tr>
<tr>
<td>Urology</td>
<td>-0.06 (0.15)</td>
<td>1.59 (1.67)</td>
</tr>
<tr>
<td>OB-Gyn</td>
<td>-0.04 (0.19)</td>
<td>-0.81 (1.87)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.29 (0.15)</td>
<td>-0.66 (1.58)</td>
</tr>
<tr>
<td>YOB</td>
<td>0.00 (0.00)</td>
<td>0.03 (0.03)</td>
</tr>
<tr>
<td>ASA</td>
<td>0.03 (0.07)</td>
<td>-0.85 (0.92)</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>-0.11 (0.10)</td>
<td>-1.36 (1.78)</td>
</tr>
<tr>
<td>LOS/FollDays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disch. comp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>-0.02 (0.03)</td>
<td>0.19 (0.51)</td>
</tr>
<tr>
<td>Complexity^2</td>
<td>0.00 (0.02)</td>
<td>0.25 (0.25)</td>
</tr>
<tr>
<td>TFL</td>
<td>0.01 (0.19)</td>
<td>-3.21 (1.85)</td>
</tr>
<tr>
<td>Psych. Safety</td>
<td>0.21* (0.10)</td>
<td>0.11 (1.57)</td>
</tr>
<tr>
<td>TFL*Cmplx</td>
<td>-0.09 (0.05)</td>
<td>-0.09 (1.33)</td>
</tr>
<tr>
<td>TFL*Cmplx^2</td>
<td>0.05 (0.03)</td>
<td>-0.31 (0.60)</td>
</tr>
<tr>
<td>PsySaf*Cmplx</td>
<td>-0.15* (0.07)</td>
<td>-0.25 (0.57)</td>
</tr>
<tr>
<td>PsySaf*Cmplx^2</td>
<td>0.03 (0.03)</td>
<td>0.36 (0.39)</td>
</tr>
</tbody>
</table>

**Indirect effect**

| Low | NA | NA | NA | NA |
| Medium | | | | |
| High | | | | |

**PRV**

| TFL | 0.3% | 1.9% | NA | NA |
| TFL moderated effect | 1.2% | 2.0% | NA | NA |

* p<.05, ** p<.01

PRV = proportion reduction in variance
### Table 3-4: Model results of laissez-faire leadership

<table>
<thead>
<tr>
<th></th>
<th>Proximal Surgical Outcomes</th>
<th>Distal Surgical Outcomes</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Errors</td>
<td>Blood loss</td>
<td>Discharge complications</td>
<td>Follow-up complications</td>
</tr>
<tr>
<td></td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
</tr>
<tr>
<td><strong>Mediator:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psych Safety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.03 (0.06)</td>
<td>0.03 (0.06)</td>
<td>0.03 (0.06)</td>
<td>-0.02 (0.07)</td>
</tr>
<tr>
<td>Emergency</td>
<td>-0.16 (0.12)</td>
<td>-0.16 (0.12)</td>
<td>-0.16 (0.12)</td>
<td>-0.15 (0.16)</td>
</tr>
<tr>
<td>Urology</td>
<td>0.31 (0.24)</td>
<td>0.31 (0.24)</td>
<td>0.31 (0.24)</td>
<td>0.35 (0.22)</td>
</tr>
<tr>
<td>OB-Gyn</td>
<td>-0.08 (0.20)</td>
<td>-0.08 (0.20)</td>
<td>-0.08 (0.20)</td>
<td>-0.02 (0.21)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.04 (0.09)</td>
<td>0.04 (0.09)</td>
<td>0.04 (0.09)</td>
<td>0.09 (0.11)</td>
</tr>
<tr>
<td>YOB</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>ASA</td>
<td>-0.06 (0.06)</td>
<td>-0.06 (0.06)</td>
<td>-0.06 (0.06)</td>
<td>0.01 (0.06)</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>0.02 (0.12)</td>
<td>0.02 (0.12)</td>
<td>0.02 (0.12)</td>
<td>0.09 (0.15)</td>
</tr>
<tr>
<td>Complexity</td>
<td>-0.01 (0.04)</td>
<td>-0.01 (0.04)</td>
<td>-0.01 (0.04)</td>
<td>0.00 (0.04)</td>
</tr>
<tr>
<td>Complexity^2</td>
<td>-0.03 (0.02)</td>
<td>-0.03 (0.02)</td>
<td>-0.03 (0.02)</td>
<td>-0.04 (0.02)</td>
</tr>
<tr>
<td>LF</td>
<td>0.12 (0.36)</td>
<td>0.12 (0.36)</td>
<td>0.12 (0.36)</td>
<td>-0.44 (0.45)</td>
</tr>
<tr>
<td>LF*Cmplx</td>
<td>-0.72 (0.44)</td>
<td>-0.72 (0.44)</td>
<td>-0.72 (0.44)</td>
<td>-0.01 (0.22)</td>
</tr>
<tr>
<td>LF*Cmplx^2</td>
<td>-0.33 (0.20)</td>
<td>-0.33 (0.20)</td>
<td>-0.33 (0.20)</td>
<td>0.07 (0.15)</td>
</tr>
<tr>
<td><strong>Full Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.10 (0.09)</td>
<td>-0.29 (0.98)</td>
<td>-1.74 (0.36)</td>
<td>NA</td>
</tr>
<tr>
<td>Emergency</td>
<td>-0.06 (0.10)</td>
<td>2.02 (2.37)</td>
<td>0.18 (0.51)</td>
<td>0.66 (0.82)</td>
</tr>
<tr>
<td>Urology</td>
<td>-0.04 (0.17)</td>
<td>1.06 (1.55)</td>
<td>0.00 (0.53)</td>
<td>-0.10 (0.87)</td>
</tr>
<tr>
<td>OB-Gyn</td>
<td>-0.02 (0.20)</td>
<td>-2.02 (1.79)</td>
<td>-0.88 (0.81)</td>
<td>1.22 (0.66)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.27 (0.15)</td>
<td>-0.31 (1.29)</td>
<td>-0.12 (0.57)</td>
<td>-0.78 (0.64)</td>
</tr>
<tr>
<td>YOB</td>
<td>0.00 (0.00)</td>
<td>0.03 (0.03)</td>
<td>-0.01 (0.01)</td>
<td>-0.05** (0.01)</td>
</tr>
<tr>
<td>ASA</td>
<td>0.04 (0.07)</td>
<td>-1.05 (0.94)</td>
<td>-0.03 (0.29)</td>
<td>-1.07** (0.37)</td>
</tr>
<tr>
<td>LOS/FollDays</td>
<td>-0.11 (0.09)</td>
<td>-0.96 (1.66)</td>
<td>0.00 (0.63)</td>
<td>1.43* (0.67)</td>
</tr>
<tr>
<td>Disch. Comp</td>
<td></td>
<td></td>
<td>0.05** (0.02)</td>
<td>-0.01 (0.01)</td>
</tr>
<tr>
<td>Complexity</td>
<td>-0.02 (0.03)</td>
<td>0.54 (0.49)</td>
<td>0.11 (0.23)</td>
<td>0.18 (0.17)</td>
</tr>
<tr>
<td>Complexity^2</td>
<td>0.00 (0.00)</td>
<td>0.38 (0.22)</td>
<td>-0.12 (0.09)</td>
<td>0.25 (0.13)</td>
</tr>
<tr>
<td>LF</td>
<td>0.23 (0.55)</td>
<td>-17.87** (3.25)</td>
<td>2.02 (1.31)</td>
<td>-0.50 (3.28)</td>
</tr>
<tr>
<td>Psych. Safety</td>
<td>0.21* (0.10)</td>
<td>0.43 (1.57)</td>
<td>0.62 (0.61)</td>
<td>0.64 (0.66)</td>
</tr>
<tr>
<td>LF*Cmplx</td>
<td>-0.24 (0.30)</td>
<td>11.81* (5.39)</td>
<td>-3.16* (1.45)</td>
<td>-0.97 (1.89)</td>
</tr>
<tr>
<td>LF*Cmplx^2</td>
<td>-0.20 (0.19)</td>
<td>11.11** (2.64)</td>
<td>-1.61 (1.17)</td>
<td>-1.79 (1.29)</td>
</tr>
<tr>
<td>PsySaf*Cmplx</td>
<td>-0.15* (0.07)</td>
<td>-0.21 (0.53)</td>
<td>0.12 (0.55)</td>
<td>-0.22 (0.38)</td>
</tr>
<tr>
<td>PsySaf*Cmplx^2</td>
<td>0.04 (0.03)</td>
<td>0.23 (0.40)</td>
<td>-0.06 (0.22)</td>
<td>-0.35 (0.25)</td>
</tr>
<tr>
<td><strong>Indirect effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Medium</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>High</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>PRV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LF</td>
<td>0.0%</td>
<td>1.6%</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>LF moderated</td>
<td>0.3%</td>
<td>7.4%</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01

PRV = proportion reduction in variance
Table 3-5: Model results of over-controlling leadership

<table>
<thead>
<tr>
<th>Mediator:</th>
<th>Proximal Outcomes</th>
<th>Distal Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Errors</td>
<td>Blood loss</td>
</tr>
<tr>
<td></td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
</tr>
<tr>
<td><strong>Mediator:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psych Safety</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.02 (0.08)</td>
<td>0.02 (0.08)</td>
</tr>
<tr>
<td>Emergency</td>
<td>-0.16 (0.12)</td>
<td>-0.16 (0.12)</td>
</tr>
<tr>
<td>Urology</td>
<td>0.24 (0.23)</td>
<td>0.24 (0.23)</td>
</tr>
<tr>
<td>OB-Gyn</td>
<td>-0.07 (0.19)</td>
<td>-0.07 (0.19)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.09 (0.09)</td>
<td>0.09 (0.09)</td>
</tr>
<tr>
<td>YOB</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>ASA</td>
<td>-0.02 (0.05)</td>
<td>-0.02 (0.05)</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>-0.01 (0.10)</td>
<td>-0.01 (0.10)</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.04 (0.04)</td>
<td>0.04 (0.04)</td>
</tr>
<tr>
<td>Complexity^2</td>
<td>-0.03*(0.01)</td>
<td>-0.03*(0.01)</td>
</tr>
<tr>
<td>OC</td>
<td>-0.07 (0.49)</td>
<td>-0.07 (0.49)</td>
</tr>
<tr>
<td>OC*Cmplx</td>
<td>0.34 (0.42)</td>
<td>0.34 (0.42)</td>
</tr>
<tr>
<td>OC*Cmplx^2</td>
<td>-0.66**(0.16)</td>
<td>-0.66**(0.16)</td>
</tr>
<tr>
<td><strong>Full Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.11 (0.09)</td>
<td>-0.79 (1.07)</td>
</tr>
<tr>
<td>Emergency</td>
<td>-0.07 (0.11)</td>
<td>3.58 (2.38)</td>
</tr>
<tr>
<td>Urology</td>
<td>-0.03 (0.15)</td>
<td>2.39 (1.58)</td>
</tr>
<tr>
<td>OB-Gyn</td>
<td>-0.02 (0.19)</td>
<td>-0.64 (1.83)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.29 (0.15)</td>
<td>-0.49 (1.59)</td>
</tr>
<tr>
<td>YOB</td>
<td>0.00 (0.00)</td>
<td>0.03 (0.03)</td>
</tr>
<tr>
<td>ASA</td>
<td>0.04 (0.07)</td>
<td>-0.91 (0.94)</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>-0.10 (0.09)</td>
<td>-1.2 (1.71)</td>
</tr>
<tr>
<td>LOS/FollDays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disch. Comp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>0.00 (0.04)</td>
<td>0.18 (0.47)</td>
</tr>
<tr>
<td>Complexity^2</td>
<td>-0.01 (0.02)</td>
<td>0.25 (0.31)</td>
</tr>
<tr>
<td>OC</td>
<td>0.58 (0.58)</td>
<td>13.70**(5.25)</td>
</tr>
<tr>
<td>Psych. Safety</td>
<td>0.24* (0.10)</td>
<td>0.45 (1.50)</td>
</tr>
<tr>
<td>OC*Cmplx</td>
<td>0.17 (0.36)</td>
<td>5.23**(2.48)</td>
</tr>
<tr>
<td>OC*Cmplx^2</td>
<td>-0.31 (0.25)</td>
<td>-0.79 (3.34)</td>
</tr>
<tr>
<td>PsySaf*Cmplx</td>
<td>-0.18* (0.08)</td>
<td>0.28 (0.70)</td>
</tr>
<tr>
<td>PsySaf*Cmplx^2</td>
<td>0.02 (0.03)</td>
<td>0.44 (0.49)</td>
</tr>
</tbody>
</table>

**Indirect effect**

| Low       | -0.47 (0.25) | NA | NA | NA |
| Medium    | -0.02 (0.12) | 3.4% | NA | NA |
| High      | -0.03 (0.07) | 4.1% | NA | NA |

**PRV**

| OC         | 0.0% | 3.4% | NA | NA |
| OC moderated effect | 0.6% | 4.1% | NA | NA |

*p<.05, **p<.01
PRV = proportion reduction in variance
Table 3-6: Model results of leadership predicting boredom

<table>
<thead>
<tr>
<th>Leadership behaviour</th>
<th>Transformational leadership</th>
<th>Laissez-faire leadership</th>
<th>Over-controlling leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.07 (0.05)</td>
<td>0.07 (0.05)</td>
<td>0.08 (0.05)</td>
</tr>
<tr>
<td>Emergency</td>
<td>-0.02 (0.09)</td>
<td>-0.01 (0.09)</td>
<td>-0.03 (0.08)</td>
</tr>
<tr>
<td>Urology</td>
<td>-0.27*(0.13)</td>
<td>-0.28*(0.13)</td>
<td>-0.28*(0.13)</td>
</tr>
<tr>
<td>OB-Gyn</td>
<td>-0.12 (0.09)</td>
<td>-0.11 (0.10)</td>
<td>-0.13 (0.09)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.04 (0.07)</td>
<td>0.03 (0.07)</td>
<td>0.04 (0.07)</td>
</tr>
<tr>
<td>YOB</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>ASA</td>
<td>0.02 (0.03)</td>
<td>0.02 (0.03)</td>
<td>0.02 (0.03)</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>-0.16**(0.06)</td>
<td>-0.16*(0.06)</td>
<td>-0.16**(0.06)</td>
</tr>
<tr>
<td>Complexity</td>
<td>-0.01 (0.02)</td>
<td>-0.02 (0.02)</td>
<td>-0.01 (0.02)</td>
</tr>
<tr>
<td>Complexity^2</td>
<td>0.00 (0.01)</td>
<td>0.01 (0.01)</td>
<td>0.00 (0.01)</td>
</tr>
<tr>
<td>Leadership behaviour</td>
<td>0.04 (0.09)</td>
<td>0.19 (0.26)</td>
<td>-0.18 (0.22)</td>
</tr>
<tr>
<td>Leadership*Cmplx</td>
<td>-0.04 (0.04)</td>
<td>0.13 (0.18)</td>
<td>0.04 (0.16)</td>
</tr>
<tr>
<td>Leadership*Cmplx^2</td>
<td>-0.02 (0.02)</td>
<td>-0.04 (0.09)</td>
<td>0.02 (0.09)</td>
</tr>
</tbody>
</table>
Table 3-7: Simple slopes analysis of transformational leadership and follow-up complications

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>7.66**</td>
<td>2.23</td>
</tr>
<tr>
<td>-1.5</td>
<td>5.05**</td>
<td>1.59</td>
</tr>
<tr>
<td>-1</td>
<td>2.75</td>
<td>1.56</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.74</td>
<td>1.70</td>
</tr>
<tr>
<td>0</td>
<td>-0.97</td>
<td>1.79</td>
</tr>
<tr>
<td>0.5</td>
<td>-2.37</td>
<td>1.88</td>
</tr>
<tr>
<td>1</td>
<td>-3.47</td>
<td>2.29</td>
</tr>
<tr>
<td>1.5</td>
<td>-4.27</td>
<td>3.27</td>
</tr>
<tr>
<td>2</td>
<td>-4.77</td>
<td>4.80</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Table 3-8: Simple slopes analysis of laissez-faire leadership and blood loss per hour vs. expected complexity

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>2.96</td>
<td>3.35</td>
</tr>
<tr>
<td>-1.5</td>
<td>-10.58*</td>
<td>4.57</td>
</tr>
<tr>
<td>-1</td>
<td>-18.57**</td>
<td>5.12</td>
</tr>
<tr>
<td>-0.5</td>
<td>-21.00**</td>
<td>4.61</td>
</tr>
<tr>
<td>0</td>
<td>-17.87**</td>
<td>3.25</td>
</tr>
<tr>
<td>0.5</td>
<td>-9.19**</td>
<td>3.03</td>
</tr>
<tr>
<td>1</td>
<td>5.05</td>
<td>6.52</td>
</tr>
<tr>
<td>1.5</td>
<td>24.84*</td>
<td>12.10</td>
</tr>
<tr>
<td>2</td>
<td>50.19**</td>
<td>19.16</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Table 3-9: Simple slopes analysis of laissez-faire leadership and complications upon discharge

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1.91</td>
<td>2.36</td>
</tr>
<tr>
<td>-1.5</td>
<td>3.15*</td>
<td>1.44</td>
</tr>
<tr>
<td>-1</td>
<td>3.58*</td>
<td>1.39</td>
</tr>
<tr>
<td>-0.5</td>
<td>3.20*</td>
<td>1.46</td>
</tr>
<tr>
<td>0</td>
<td>2.02</td>
<td>1.31</td>
</tr>
<tr>
<td>0.5</td>
<td>0.04</td>
<td>1.34</td>
</tr>
<tr>
<td>1</td>
<td>-2.75</td>
<td>2.35</td>
</tr>
<tr>
<td>1.5</td>
<td>-6.34</td>
<td>4.24</td>
</tr>
<tr>
<td>2</td>
<td>-10.74</td>
<td>6.82</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Table 3-10: Simple slopes analysis of over-controlling leadership and blood loss per hour vs. expected

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>0.07</td>
<td>12.26</td>
</tr>
<tr>
<td>-1.5</td>
<td>4.07</td>
<td>6.20</td>
</tr>
<tr>
<td>-1</td>
<td>7.68*</td>
<td>3.11</td>
</tr>
<tr>
<td>-0.5</td>
<td>10.89**</td>
<td>4.05</td>
</tr>
<tr>
<td>0</td>
<td>13.70**</td>
<td>5.25</td>
</tr>
<tr>
<td>0.5</td>
<td>16.12**</td>
<td>5.39</td>
</tr>
<tr>
<td>1</td>
<td>18.13**</td>
<td>4.82</td>
</tr>
<tr>
<td>1.5</td>
<td>19.76**</td>
<td>5.40</td>
</tr>
<tr>
<td>2</td>
<td>20.98*</td>
<td>9.10</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Table 3-11: Simple slopes analysis of over-controlling leadership and complications at follow-up

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>-0.20</td>
<td>8.43</td>
</tr>
<tr>
<td>-1.5</td>
<td>-2.14</td>
<td>4.86</td>
</tr>
<tr>
<td>-1</td>
<td>-2.82</td>
<td>2.81</td>
</tr>
<tr>
<td>-0.5</td>
<td>-2.24</td>
<td>2.58</td>
</tr>
<tr>
<td>0</td>
<td>-0.40</td>
<td>2.95</td>
</tr>
<tr>
<td>0.5</td>
<td>2.70</td>
<td>2.99</td>
</tr>
<tr>
<td>1</td>
<td>7.05*</td>
<td>2.85</td>
</tr>
<tr>
<td>1.5</td>
<td>12.67**</td>
<td>3.64</td>
</tr>
<tr>
<td>2</td>
<td>19.54**</td>
<td>6.10</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Table 3-12: Simple slopes analysis of over-controlling leadership and psychological safety

(a) Main model

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>-3.40**</td>
<td>1.05</td>
</tr>
<tr>
<td>-1.5</td>
<td>-2.07**</td>
<td>0.64</td>
</tr>
<tr>
<td>-1</td>
<td>-1.07**</td>
<td>0.38</td>
</tr>
<tr>
<td>-0.5</td>
<td>-0.41</td>
<td>0.36</td>
</tr>
<tr>
<td>0</td>
<td>-0.07</td>
<td>0.49</td>
</tr>
<tr>
<td>0.5</td>
<td>-0.07</td>
<td>0.62</td>
</tr>
<tr>
<td>1</td>
<td>-0.40</td>
<td>0.71</td>
</tr>
<tr>
<td>1.5</td>
<td>-1.06</td>
<td>0.76</td>
</tr>
<tr>
<td>2</td>
<td>-2.06**</td>
<td>0.78</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01

(b) Follow-up model

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>-2.93**</td>
<td>0.97</td>
</tr>
<tr>
<td>-1.5</td>
<td>-1.72**</td>
<td>0.62</td>
</tr>
<tr>
<td>-1</td>
<td>-0.79*</td>
<td>0.40</td>
</tr>
<tr>
<td>-0.5</td>
<td>-0.14</td>
<td>0.38</td>
</tr>
<tr>
<td>0</td>
<td>0.22</td>
<td>0.49</td>
</tr>
<tr>
<td>0.5</td>
<td>0.31</td>
<td>0.61</td>
</tr>
<tr>
<td>1</td>
<td>0.11</td>
<td>0.70</td>
</tr>
<tr>
<td>1.5</td>
<td>-0.36</td>
<td>0.77</td>
</tr>
<tr>
<td>2</td>
<td>-1.12</td>
<td>0.84</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Table 3-13: Simple slopes analysis of over-controlling leadership and complications at discharge (post-hoc analysis)

<table>
<thead>
<tr>
<th>Value of complexity</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>0.85</td>
<td>5.96</td>
</tr>
<tr>
<td>-1.5</td>
<td>-2.78</td>
<td>3.73</td>
</tr>
<tr>
<td>-1</td>
<td>-4.98*</td>
<td>2.35</td>
</tr>
<tr>
<td>-0.5</td>
<td>-5.74**</td>
<td>1.92</td>
</tr>
<tr>
<td>0</td>
<td>-5.06*</td>
<td>2.02</td>
</tr>
<tr>
<td>0.5</td>
<td>-2.95</td>
<td>2.09</td>
</tr>
<tr>
<td>1</td>
<td>0.61</td>
<td>2.06</td>
</tr>
<tr>
<td>1.5</td>
<td>5.60*</td>
<td>2.35</td>
</tr>
<tr>
<td>2</td>
<td>12.03*</td>
<td>3.49</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Figure 3-1: Illustration of full model
Figure 3-2: Transformational leadership and complexity interaction on follow-up complications
Figure 3-3: Laissez-faire leadership and complexity interaction on blood loss per hour vs. expected
Figure 3-4: Laissez-faire leadership and complexity interaction on complications at discharge
Figure 3-5: Over-controlling leadership and complexity interaction on blood loss per hour vs. expected
Figure 3-6: Over-controlling leadership and complexity interaction on follow-up complications
Figure 3-7: Over-controlling leadership and complexity interaction on psychological safety

(a) Main model:

(b) Follow-up model:
Figure 3-8: Over-controlling leadership and complexity interaction on complications at discharge
Chapter 4

General Discussion

The goal of this dissertation was to examine the physical and psychological consequences faced by people working in the extreme context of healthcare, and how leadership and contextual elements affect those outcomes. Study 1 examined the development of the psychological outcomes of empathy, self-efficacy, and daytime sleepiness of newcomers to the practice of medicine. Leadership and contextual-specific experiences were found to significantly affect changes in the outcomes. Study 2 investigated the physical consequences of extreme contexts, looking at the effect of leadership on proximal and distal surgical outcomes. Surgical complexity was shown to moderate many of the relationships between leadership and surgical outcomes.

Together, these studies represent two of the first empirical studies of leadership in extreme contexts. The studies examined physical and psychological outcomes, and help us understand the experience of working in this extreme environment. Results suggest that leadership does matter in extreme contexts, and has effects on the physical consequences faced by patients, and psychological consequences faced by employees. Furthermore, incorporating contextual elements (e.g., unique factors such as morbidity and mortality and beneficiary/patient contact, surgical complexity) tells us more about working in extreme contexts. Contextual elements directly influence psychological consequences, and moderate the effects of leadership on physical outcomes.

Theoretical Contributions

Leadership. This dissertation contributes to the leadership literature in many ways. First, study findings suggest that the same leadership behaviours that have been widely studied in
traditional contexts also have effects in extreme contexts. Hannah, Uhl-Bien, Avolio, and Cavarretta (2009) noted the need to identify the different behaviours and skills required by leaders in extreme contexts, but findings from my two studies suggest that with adjustments (e.g., incorporating contextual variables such as complexity), the same leadership behaviours that have been the focus of leadership research can be used to study leadership in extreme contexts.

Second, to the long list of organizational variables influenced by transformational leadership, this dissertation adds self-efficacy (which albeit in general is not new) and daytime sleepiness of medical and nursing students (Study 1), and post-discharge complications of individuals who have had surgery (Study 2). However, unlike prior research which tends to highlight the positive consequences of transformational leadership (e.g., Lowe, Kroeck, & Sivasubramaniam, 1996; Wang, Oh, Courwright, & Colbert, 2011), the findings suggest that perhaps in extreme contexts transformational leadership does not always have the intended consequences: Transformational leadership positively influenced changes in self-efficacy (consistent with traditional organization findings; e.g., Pillai & Williams, 2004), but may exert negative effects on female students’ daytime sleepiness. Transformational leadership also increased the chances that a low-complexity patient would experience a post-discharge complication. While I expect that this relationship is due to the trust fostered between surgeons and patients (as discussed in Chapter 3), it does suggest that one of the beneficial effects of transformational leadership may exert a potential strain on the medical system (e.g., complications at follow-up included such complications as patients seeking out ER care or appointments with medical professionals).

Third, the findings concerning laissez-faire leadership add to the ever-expanding research that shows the negative impacts of non-leadership. Laissez-faire leadership negatively impacted
medical trainees’ empathy and daytime sleepiness (Study 1), and proximal (i.e. blood loss) and distal (i.e. complications at discharge) surgical outcomes (Study 2). Laissez-faire leadership in traditional contexts does not always exert significant effects (e.g., Corrigan, Diwan, Campion & Rashid, 2002), but the findings of my dissertation suggest that non-leadership can be harmful in extreme contexts. This is consistent with Hannah et al.’s (2009) suggestion that leadership in extreme contexts is necessary for reducing ambiguity and fostering sense making.

Fourth, I also studied two active forms of negative leadership: Abusive supervision in Study 1, and over-controlling leadership in Study 2. No significant results emerged for the impact of abusive supervision in Study 1, likely due to the infrequency of the behavior and resulting range restriction. Range restriction within abusive supervision is not uncommon problem when studying abusive supervision (e.g., on a 5-point scale: $M = 1.34$, $SD = 0.57$, Byrne et al., 2014; $M = 1.39$, $SD = 0.60$, Tepper, Henle, Schurer Lambert, & Giacalone, 2008; $M = 1.27$, $SD = 0.47$, Tepper, Moss, & Duffy, 2011). Over-controlling leadership, defined in this dissertation as leadership behaviours that constrain or limit subordinates’ discretion in performing their job, significantly impacted blood loss, complications at discharge, and complications at follow-up. Over-controlling leadership is a significant contribution to leadership theory, capturing what are anecdotally common leadership behaviours (e.g., micromanaging), that evidently impact surgical outcomes. Over-controlling leadership may be a more interesting, relevant, and promising negative leadership behaviour to study than abusive supervision.

---

23 As discussed in Chapter 3, there was a favourable relationship between laissez-faire and blood loss for low complexity surgeries, but in this case laissez-faire leadership may augment the strength of the situation by not interfering.
Finally, Study 2 also highlighted the notion of consistency of leadership. Researchers have noted the need to address consistency in leadership research (e.g., Hannah, Sumanth, Lester, & Cavarretta, 2014): Simply stated, it is unlikely that any one leader would perform only one leadership behavior (e.g., transformational leadership and never enact any others, such as over-controlling leadership and laissez-faire leadership). Leaders can (and do) display a variety of leadership behaviours; capturing the tendency to do this more accurately captures the true experience of leadership, improving the applicability of research findings to real organizations. Mullen, Kelloway and Teed (2011) empirically illustrated the utility of doing so, and this dissertation demonstrated an additional way of doing this.

**Extreme contexts.** This dissertation also contributes to the literature on extreme contexts. First, Study 1 showed that the unique experiences of working in an extreme context impact psychological outcomes. Exposure to morbidity and mortality (i.e. death and suffering of patients) negatively impacts self-efficacy and daytime sleepiness. On the other hand, interactions that highlight the positive, prosocial effects of helping patients positively influence self-efficacy and daytime sleepiness. These findings highlight some of the consequences faced by those working in the medical environment.

Second, boredom was added to the extreme context framework in Chapter 1 of this dissertation. Hannah et al., (2009) briefly mention boredom in their framework as well, noting the need for leaders to overcome such feelings in extreme contexts. However, the results in this dissertation suggest that boredom was neither a significant predictor (Study 1) nor mediating (Study 2) variable. Results suggested that boredom was not a common experience in the operating room: Thus, it may not be that boredom does not matter in extreme contexts, but this could not be studied because of range restriction.
Finally, Study 2 examined the contextual variable of surgical complexity as a moderator of the relationship between leadership and surgical outcomes. Results showed that surgical complexity, representing the probability and magnitude of potential surgical consequences, significantly affected whether and how leadership influenced surgical outcomes. In their theoretical piece on leadership in extreme contexts, Hannah et al., (2009) discuss the potential impact of different magnitudes and probabilities on employees in extreme contexts, and the findings of this study suggest that they impact the effectiveness of leadership as well.

**Future Research**

The findings in this dissertation, combined with some limitations of my research, suggest a number of avenues of future research. First, the significant findings of the impact of leadership behaviours and contextual factors highlight the need to continue research within extreme contexts, and to begin to answer some of the questions I posed at the outset of this dissertation. Organizational researchers should test the boundaries of organizational theories using extreme contexts to help us understand which do and do not apply to those working under such conditions.

Second, researchers should directly compare relationships (e.g., transformational leadership predicting daytime sleepiness) in traditional contexts to those in extreme contexts, using contextual moderators. Researchers could collect specific information on extremity in a workplace (e.g., probability of extreme events in respondents’ jobs), to further test the impact of leadership behaviours on organizational outcomes. Research that takes this higher perspective will help us understand the specific differences and similarities of working in extreme contexts.

Third, I offered a number of explanations for some of the unexpected negative consequences of transformational leadership (e.g., sleepiness of female trainees, follow-up
complications of low complexity surgical patients). Transformational leadership is widely studied as a form of positive leadership, given the breadth of positive outcomes with which it is related. However, I wonder if this is because leadership researchers choose positive outcomes to study, and fail to investigate potentially negative outcomes. As Hiller, DeChurch, Murase and Doty (2011) discussed, ignoring potentially relevant leadership outcomes will prevent us from fully understanding the effects of leadership and advance our theories. A few researchers have identified the need to study the potential negative effects of transformational leadership (e.g., Yukl, 1999), and have found a few such outcomes, such as the negative effects of performance expectations on sleep (Syrek & Antoni, 2014)\textsuperscript{24}, the negative effects of inspirational motivation on anxiety and organizational commitment for employees with autism spectrum disorder (Parr, Hunter, & Scott Ligon, 2013), and curvilinear effects of transformational leadership on research and development innovation (Eisenbeiß & Boerner, 2010). Perhaps the unfavourable effects of transformational leadership on female trainees’ sleepiness and on low complexity patients can be added to the list of the downsides of transformational leadership. A full appreciation of transformational leadership must include its risks and side effects.

Fourth, as Hannah et al., (2009) suggested, different behaviours may be required given the temporal state of the organization. Johns (2006) mentioned the need to understand entire events. Studying an entire cycle of an extreme event would yield practical information, and let us answer the questions on which behaviours are more effective, and when. Trust and credibility-building behaviours may be important in preparatory stages, but more transactional behaviours may be required in reactive phases, followed by transformational leadership behaviours to foster learning and development, and preparedness for the next cycle. Fifth, taking

\textsuperscript{24} Though it is important to note that the authors only studied performance expectations, and not the true comprehensive theory of transformational leadership.
an even longer-term approach, future research may benefit from incorporating a career-long approach. Study 1 captured the early experiences of medical trainees, when the majority of their experiences were novel. Study 2 captured the operating room, where the average professional tenure of a surgeon was over 23 years. It would be interesting to understand when (and if) experiences cease to feel novel, and become part of “just another day”. If employees in extreme contexts undergo psychological change where they no longer experience their contexts as extreme, it may be even easier to apply findings from traditional contexts. This consideration of “when” may be an important moderator that helps us understand the organizational psychology of extreme contexts.

Sixth, as noted above, boredom was not a significant predictor or mediator in extreme contexts, despite my rationale that it should be included as part of an extreme context. But before boredom is completely discounted, researchers should consider studying it in different ways. Extreme contexts with longer times between extreme events (e.g., firefighting) may yield more pronounced effects, as employees may experience longer, more pronounced feelings of boredom, compared to the studies in this dissertation, especially surgery where the incidence of boredom was very low. Researchers could also consider looking at boredom in even smaller time periods, by means of diary studies or studies that prompt participants to rate their emotions whenever they receive a text message (e.g., Uy, Foo, & Aguinis, 2010), to properly capture the true experience of boredom, and its consequences. Recent theoretical and empirical advances in the study of boredom also suggest that boredom is multidimensional (Baratta & Spence, 2015), with seemingly opposite factors in some cases (e.g., low arousal, and high arousal). The multidimensional nature of boredom might explain the lack of findings in my research, and offers an opportunity to future research to uncover the consequences of different dimensions.
Seventh, gender skewness was a characteristic of both samples in this dissertation (i.e. high proportion of female trainees in Study 1, high proportion of male surgeons in Study 2). This may be a limitation in terms of generalizability of the findings to other populations, though a consideration of other extreme contexts such as the military, firefighting, law enforcement, and space exploration suggest a similar male domination. Even in the upper echelons of traditional organizations there is a disproportionate lack of females (e.g., Swanson, 2015). So the findings of Study 2 may in fact be applicable to these populations as well. However, more and more women are entering these more-extreme professions (e.g., Hutchins, 2014; Long, 2014; “Women in the Canadian Armed Forces”, 2014), and future research should focus on the specific effects of this trend on those women, as research from traditional organizations suggest that women face a series of obstacles and challenges when navigating the workplace (e.g., Eagley & Carli, 2007). Findings from Study 1 already suggest that females react differently to transformational leadership in terms of daytime sleepiness, and other research suggests that female medical students (e.g., Nora et al., 2002), female military personnel (e.g., Matthews, Ender, & Laurence, 2009), and female police officers (e.g., Garcia, 2003) face unique obstacles. Understanding how to effectively incorporate women into extreme contexts will be necessary.

Eighth, I noted in the introductory chapter, and Hannah et al., (2009) acknowledged, that leadership is expected to have effects in extreme contexts, but that extreme contexts are also expected to have an effect on leadership: It is hypothetically a reciprocal relationship. The research presented in this dissertation focuses mostly on the effects of leadership in extreme contexts, but did not examine elements of extreme contexts as potential antecedents to leadership behaviour. Given that leadership did in many cases affect physical and psychological outcomes faced by those working in and affected by extreme contexts, it is important to understand what
causes leaders to behave in certain ways. For example, as discussed in Study 2, does over-controlling leadership cause blood loss, or does blood loss cause a surgeon to behave more over-controlling? Does the amount of trauma that leaders have experienced affect the way they lead and develop medical trainees? These dynamic relationships will be important to understand.

Ninth, future research considering extreme contexts may also benefit from considering the literature on high reliability organizations (e.g., Roberts, 1990; Roberts, Bea & Bartles, 2001): A high reliability organization is defined as an organization that faces the risk of catastrophic failures, yet rarely commits errors (Christianson, Sutcliffe, Miller, & Iwashyna, 2011). Among behaviours and principals of high reliability organizations, are learning from, discussing, and celebrating errors, fostering resiliency, and using procedures such as checklists to ensure diligence (Christianson et al, 2011). Many extreme context organizations can already be considered high reliability organizations, and others can learn from and apply these principals. A final consideration to be made with respect to high reliability organizations is the potential range restriction that they present, as was perhaps evident in this dissertation.

Tenth, future research should replicate the research presented in this dissertation. Researchers should also replicate in different geographies (e.g., in other countries), and other extreme contexts (though generalizability was never a central aim of my research). Finally, causation cannot be confirmed for either of the studies presented in this dissertation, and so future research should use experimental manipulation to further understand the effects of leadership and extreme contexts on those working in them.

**Organizational Implications**

One of the benefits of studying a specific context is that any findings may have greater implications for that context, in this case to medical settings. First, results show that leadership
behaviours in the medical context impact both physical and psychological outcomes. But the majority of physicians and nurses do not receive leadership training, unless they undertake more managerial roles (e.g., Stoller, 2009; Sullivan, Bretschneider, & McCausland, 2003). There is an opportunity in healthcare to provide more widespread leadership training that could potentially improve the psychological states of junior staff (i.e. empathy, self-efficacy, daytime sleepiness as studied in Study 1), and surgical outcomes that directly impact patients (i.e. blood loss, complications as studied in Study 2). Second, results suggest that context-relevant factors also affect physical and psychological outcomes. Experiences with patient morbidity and mortality, and beneficiary contact, affect the psychological development of trainees. Ways to attenuate the impact of morbidity and mortality (e.g., resilience training, counseling) should be explored, as should ways to encourage and increase the depth of beneficiary contact. The complexity of a surgery influences how much surgeons’ leadership behaviours influence the outcomes of a surgery. This information may help hone in on the types of behaviours that should be trained based on the complexity of a surgery.

Conclusion

The aim of this dissertation was to investigate the effects of leadership and contextual elements in the extreme context of healthcare. With two studies examining the psychological and physical outcomes faced by those working in such contexts, I learned that (a) leadership matters in extreme contexts, that (b) context shapes its effects, and (c) it affects the development of medical trainees. Leadership behaviours affect the empathy, self-efficacy, and daytime sleepiness of medical trainees, as well as blood loss and complications from surgery. Contextual factors also affect the self-efficacy and daytime sleepiness of medical trainees, and the relationship between leadership behaviours and surgical outcomes.
This dissertation represents one of the first empirical investigations of leadership in extreme contexts. Future research should aim to replicate study findings, and further investigate leadership in extreme contexts, and contrast its effects with those in traditional contexts. The results of my research suggest the potential impact of leadership behaviours in extreme contexts, which could improve the psychological and physical outcomes of employees working in and constituents affected by these extreme contexts.
4.1 References


Appendix A – Study 1: Full questionnaire

1. Please enter your full name so your responses can be matched to those in other periods

**Demographics**
1. Age
2. Gender

**Empathy**

**Interpersonal Reactivity Index**

*Davis (1980)*

5-point scale running from 0 (does not describe me well), to 4 (describes me very well)

**Perspective-taking**
1. Before criticizing somebody, I try to imagine how I would feel if I were in their place.
2. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments. (-)
3. I sometimes try to understand my friends better by imagining how things look from their perspective.
4. I believe that there are two sides to every question and try to look at them both.
5. I sometimes find it difficult to see things from the "other guy's" point of view. (-)
6. I try to look at everybody's side of a disagreement before I make a decision.
7. When I'm upset at someone, I usually try to "put myself in his shoes" for a while.

**Empathic concern**
8. When I see someone being taken advantage of, I feel kind of protective toward them.
9. When I see someone being treated unfairly, I sometimes don't feel very much pity for them. (-)
10. I often have tender, concerned feelings for people less fortunate than me.
11. I would describe myself as a pretty soft-hearted person.
12. Sometimes I don't feel sorry for other people when they are having problems. (-)
13. Other people's misfortunes do not usually disturb me a great deal. (-)
14. I am often quite touched by things that I see happen.

**Self-Efficacy**

**Medical student-specific self-efficacy**

*Artino, 2012*

The following items address your confidence in relation to your medical knowledge and skills. For each item, select the response that best reflects your level of confidence. At this point in your medical training, how confident are you that you can. . . 5-point scale (not at all confident, slightly confident, moderately confident, quite confident, and extremely confident)
1. Apply knowledge of normal function to each of the major organ systems?
2. Effectively manage the uncertainty associated with patient care, such as when the patient has multiple treatment options, each with its own risks and benefits?
3. Apply knowledge of epidemiology of common diseases, such as heart disease, to reduce disease incidence?
4. Use effective listening skills when interacting with a patient?
5. Demonstrate caring when counseling a patient?
6. Accurately gather essential information from a patient?
7. Perform a thorough physical exam?
8. Develop an appropriate differential diagnosis?
9. Generate a patient-specific treatment plan?
10. Use information technology to support patient-care decisions?
11. Work effectively with other health care professionals to provide high-quality patient care?
12. Improve clinical practice using a systematic approach?
13. Demonstrate sensitivity to patients’ cultural differences?
14. Balance professional responsibilities with personal responsibilities?

**Sleep Quality**

*On average, how many hours do you sleep each night? _____*

**Epworth Sleepiness scale**
*Johns, 1991*

How likely are you to doze off or fall asleep in the following situations, in contrast to just feeling tired? This refers to your usual way of life in recent times. Even if you have not done some of these things recently try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation: 0=would never doze, 1=slight chance of dozing, 2=moderate chance of dozing, 3=high chance of dozing.

1. Sitting and reading
2. Watching TV
3. Sitting, inactive in a public place (e.g., theater or a meeting)
4. As a passenger in a car for an hour without a break
5. Lying down to rest in the afternoon with circumstances permit
6. Sitting and talking to someone
7. Sitting quietly after a lunch without alcohol
8. In a car, while stopped for a few minutes in traffic

**Leadership Behaviours**

**Transformational Leadership**
*Beauchamp (2010)*

On a scale of 1= strongly disagree to 5=strongly agree
My leader…

1. Shows that s/he cares about me
2. Acts as a person that I look up to
3. Is enthusiastic about what I am capable of achieving
4. Provides me with tasks and challenges that get me to think in different ways
5. Motivates me to try my hardest
6. Gets me to question my own and others’ ideas
7. Is optimistic about what I can accomplish
8. Behaves as someone I can trust

Laissez-faire leadership
Hinkin and Schriesheim (2008) – 8 items on a 7-point Likert scale

1. I often perform well and still receive no praise from my leader
2. When I perform well my leader usually does nothing
3. My good performance often goes unacknowledged by my leader
4. I don’t often get praised by my leader when I perform well
5. My leader gives me no feedback when I perform poorly
6. When I perform poorly in my job I receive no criticism from my leader
7. When I perform poorly my leader does nothing
8. My poor performance often gets no response from my leader

Abusive Supervision
Mitchell & Ambrose (2007)
My leader…
1. Ridicules me
2. Tells me my thoughts or feelings are stupid
3. Puts me down in front of others
4. Makes negative comments about me to others
5. Tells me I’m incompetent
6. Invades my privacy
7. Doesn’t give me credit for jobs requiring a lot of effort
8. Blames me to save himself/herself from embarrassment
9. Breaks promises he/she makes
10. Lies to me

Morbidity and Mortality
Elements of context – developed by author
Please rate the frequency of which you experienced the following during your past rotation. 5-point scale, 1=never, 2=less than once per week, 3=once per week, 4=many times per week, 5=daily:

1. Felt hopeless than you could improve a patient’s quality of life
2. Felt hopeless that you could save a patient’s life
3. Had a patient die
4. Had a patient get worse instead of better
5. Were unable to help a patient
6. Were unable to convince a patient to follow instructions
7. Witnessed a patient die
8. Witnessed a patient in extreme pain
9. Had to tell a patient’s family that the patient died
10. Had to perform a severe procedure in-the-moment without being able to tell the patient first
11. Had a patient unexpectedly die
12. Had a patient unexpectedly get worse

**Beneficiary Contact**


Thinking about your most recent rotation, please rate the following on a scale of 1=strongly disagree to 7=strongly agree

**Frequency**
1. My rotation allowed frequent communication with the people who benefit from my work.
2. My rotation often gave me the opportunity to meet the people who benefit from my work.
3. My rotation enabled me to interact regularly with the people who benefit from my work.

**Breadth**
1. My rotation provided me with contact with different groups of people who benefit from my work.
2. My rotation allowed me to interact with a variety of people who benefit from my work.
3. My rotation enabled me to meet diverse groups of people who benefit from my work.

**Depth**
1. My rotation enabled me to build close relationships with the people affected by my work.
2. My rotation allowed me to form emotional connections with the people who benefit from my work.
3. My rotation gave me the chance to have meaningful communications with the people who benefit from my work.

**Boredom**

*Drory (1982)*

What percent of the time during the past rotation did you experience:
1. Feeling bored
2. Feeling that you wish to do something else now
3. Feeling of monotony
4. Feeling that time goes very slowly
5. Feeling that nothing happens
6. Feeling that you wish to be done your shift
Appendix B – Study 2: Training Materials

Leadership in the OR: Observer Training
June 11, 2014

Agenda

• Introduction & Overview

• Leadership
  – Theory
  – Methodology

• Medical Outcomes
  – Patient Safety & Errors
  – Other Data

• Patient Consent

• Summary & Questions
Why Study Leadership in the OR?

- Lack of research on leadership in the OR, but...
- Different forms of leadership have been shown to have significant effects on a wide variety of behavioral outcomes
  - Team & individual performance
  - Safety & errors
- Do these behaviours translate into the OR, and do they have the same impacts?
  - Principles of psychology, organizational behavior and leadership are now being applied in the broad health care sector (eg patient safety)
  - Strong indications that leadership might also affect surgeon’s behavior, surgical teams and patient outcomes
- If so, can we spark a change and incorporate leadership training into medical and hospital training?

Study Overview

- Sampling
- Pre-surgery
  - Patient consent
  - Patient & surgery information
  - Expected complexity
- During-surgery
  - Timing
  - Checklists
  - Leadership behaviours
  - Errors & complications
- Immediately post-surgery
  - Post-hoc complexity
  - Team surveys
  - Observer surveys
- Upon discharge
  - Patient records
- 30-day follow-up
  - Phone call
**Agenda**

- Introduction & Overview

- Leadership
  - Theory
  - Methodology

- Medical Outcomes
  - Patient Safety & Errors
  - Other Data

- Patient Consent

- Summary & Questions

---

**Very broad “definition” of leadership**

- Leadership is not about one person does by herself or himself; instead, leadership is all about the kinds of behaviors that put other people in a position to that it makes it more likely that they will excel (or fail)

- Leadership is about moments
Our study

• Most leadership research focuses only on one theory
• Three different “types” of leadership
  – Transformational leadership
  – Laissez faire
  – Over control

Transformational Leadership

• Most widely studied leadership theory
• Comprises four different behaviors
  – Idealized influence
  – Inspirational motivation
  – Intellectual stimulation
  – Individualized consideration
These are very "big" leadership ideas, but what does it all mean in reality?

**Idealized influence**

- *Idealized influence* reflects the ethical component of transformational leadership. Leaders who behave in a manner consistent with idealized influence ... avoid self-interest, and instead engage in behaviors that are good for the organization and its members.

- [They] ... are motivated by their moral commitment to the collective good, rather than what is good for themselves. Leaders who behave in a manner consistent with the values of idealized influence would opt to do what is right, rather than what is easy or expedient, and resist temptations to maximize their own or their organizations’ short-term interests.

- These leaders would most often be described by others, perhaps especially by their followers, as role models who act with integrity, and show a deep respect for others.
The surgeon...  

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acts as a person that team members look up to</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Treats team members in ways that build their respect</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Talks about his/her personal values</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Behaves as someone team members can trust.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>


**Inspirational motivation**

- *Inspirational motivation* involves those behaviors that help employees perform “beyond expectations”—both beyond the expectations employees hold for themselves, and that others hold for them.
- These leaders have a vision of what needs to be accomplished ..., and they convey to their followers in many different ways: They tell stories, use symbols and metaphors, and interact with their followers.
- They inspire by setting high but realistic goals for their followers; this demonstrates their belief in their followers’ abilities and a trust in their integrity, essentially establishing a self-fulfilling prophecy.
**Inspirational motivation**

- Through these behaviors and interactions, inspirationally motivating leaders nurture a deep sense of self-efficacy and resilience in their followers, helping them to believe in themselves so that they will confront the internal psychological hurdles and external barriers that inhibit high levels of performance, and persist in their efforts.

---

<table>
<thead>
<tr>
<th>The surgeon...</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstrates that s/he believes in team members</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Is enthusiastic about what team members are capable of achieving</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Motivates team members to try their hardest</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Is optimistic about what the team can accomplish</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 1: The Surgeon's Role in Intellectual Stimulation

<table>
<thead>
<tr>
<th>The surgeon...</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provides members with tasks and challenges that get them to think in different ways</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Gets members to question their own and others’ ideas</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Encourages members to look at issues from different sides</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Creates lessons that really encourage team members to think</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

**Individualized consideration**

- Leaders high in *individualized consideration* provide needed emotional and instrumental support to their employees, thereby developing their personal and work-related development in the longer-term.

- In doing so, these leaders’ interactions with their employees, however brief and in whatever form (e.g., face-to-face, email, videoconferencing), would be characterized by their active listening, caring and focus on the other person.

---

**Individualized consideration**

- [These] behaviors ... define the quality of the leader-follower relationship, and underlie the extent to which leaders’ idealized influence, inspirational motivation and intellectual stimulation would be accepted and enacted.

- Individualized consideration makes the personal and organizational development of employees a leadership priority.

- Leaders who display individualized consideration would most likely be perceived by their followers as development-oriented, empathic and compassionate.
The surgeon...

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shows that s/he cares about team members</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Tries to know every member of the team</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Tries to help team members who might be struggling</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Recognizes the needs and abilities of each member of the team</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>


---

**Over-control**

- Research shows that
  - feeling in control is associated with physical and psychological well-being, as well as higher levels of job performance
  - feeling that one lacks control is associated with lower physical and psychological well-being, as well as lower levels of job performance
Research also shows that feeling over controlled by another results in adverse consequences

- In marriages
- At work

In particular, over control refers to behaviors that aimed at restricting, constraining or controlling other people, thoughts, emotions and behaviors

Case example

Laissez faire

- Non-leadership, or poor leadership
- Very little research; initially believed to be inconsequential but we now know that this has negative outcomes (e.g., bullying, ambiguity, more accidents and injuries)
- Reward omission—these leaders don’t acknowledge, praise, recognize when they should
- Punishment omission—these leaders don’t punish, correct when they should


<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th></th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Team members often perform well and still receive no praise from the surgeon</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. When team members perform well, the surgeon usually does nothing</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. Team members’ good performance often goes unacknowledged by the surgeon</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4. Team members don’t often get praised by the surgeon when they perform well</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. The surgeon gives team members no feedback when they perform poorly</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. When team members perform poorly they receive no criticism from the surgeon</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. When team members perform poorly the surgeon does nothing</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Team members' poor performance often gets no response from the surgeon</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thank you!
Leadership Methodology – Live Observations (pg. 8)

- Form to fill out during surgery (page 8 of forms) beginning when the patient enters the room and ending when the patient leaves the room
  - Time (hh:mm)
    - Note the time you see the behaviour, or the behaviour begins
  - Error/Leadership (circle)
  - Team member performing (see page 5 for team list)
    - Any team member
    - Only note 1 person, use additional lines if a behaviour is followed by another
  - Leadership behaviour category (circle)
  - Additional notes
    - Not necessary, but if you want to explain a choice or provide context, include it here
- This pack only has one sheet, but you will be given multiple sheets to be safe

Leadership Methodology – Surveys (pg. 10-11)

- Using this to confirm the reliability of the observations, and for additional data
- When completing, think about the surgeon only
- Include your name at the top, and include in the envelope
Leadership Examples

- Watch these few clips and use the observation sheet to note any behaviours
- Following each clip you’ll also be asked to complete the survey

Agenda

- Introduction & Overview
- Leadership
  - Theory
  - Methodology
- Medical Outcomes
  - Patient Safety & Errors
  - Other Data
- Patient Consent
- Summary & Questions
Patient Safety & Complications

- What is patient safety?
- What is an error?
- How to spot an error?
- Examples
  - Technical
    - Sponges and materials left inside patients
    - Incorrect blood type given
    - Improper incision/incorrect incision site
    - Multiple attempts, esp. resulting in delays
    - Puncture of an organ, artery or other tissue
  - Procedural
    - Failure to check patient information
    - Failure to complete checklists

Observation Form (pg. 8)

- Recorded simultaneously on the same form as leadership behaviours
  - Time (hh:mm)
  - Error/Leadership (circle)
  - Team member(s) performing action
    - Could be one or more people, list all
  - Describe error, rate severity (refer to bottom on page, details next slide)
  - Additional notes
    - Likely not necessary, but if a participant wants to explain something after, include it here
We will now go through each item on the observation sheet, and discuss where necessary.

For your reference, note where to find each piece of information.

Observation Sheet (pg. 4-7)

- Page 4:
  - Logistics:
    - Your name
    - Observation #
    - Date of surgery
    - Scheduled time of surgery
  - Patient Information:
    - Gender
    - CR#
    - YOB
    - ASA
    - Date admitted to hospital
Observation Sheet (pg. 4-7)

• Page 4:
  – Surgery information
    • Department
    • Type of surgery (elective/emergency)
    • Name of operation(s)
    • Surgeon rated complexity
    • Type of anesthesia
    • Anesthetist-rated complexity
    • Nurse-rated complexity
    • Expected blood loss

Observation Sheet (pg. 4-7)

• Page 5:
  – Team roster
    • Name & Position
    • Space to check once you’ve received their completed survey
    • If you have time, write their names on the surveys to be distributed
**Observation Sheet (pg. 4-7)**

- Page 6:
  - During Surgery
    - Time you enter the room
    - Time patient enters
    - Time surgeon enters
    - Time anesthesiologist begins anesthetic
    - Time surgery begins
    - Notes on what is the “start”

---

**Team Member Survey**

- Page 9
  - Distribute to all team members (bring some extra pens)
  - Make sure to collect

  - If they want to include specific notes, ask them to write on the reverse side
  - If they have any specific questions, direct them to Amy
Discharge (pg. 12)

• Complete at the 30-day follow-up, or if you have downtime and patient has been discharged
  – Your name
  – Date of discharge
  – Length of stay in PACU
  – Deviations
  – Post-op complication & the Clavien Dindo (see next slide)
  – Did patient:
    • Receive antibiotics
    • Visit the ICU
    • Receive blood transfusions

30-day follow-up (pg. 12)

• Phone call follow-up to patient
• Amy will keep track of when follow-up is necessary, and together we will all make these calls
• Information:
  – Completed by
  – Date
  – Unplanned visits to medical provider
  – Antibiotics
  – Other complications
  – List any deviations
  – List any complications (and Clavien-Dindo)
Agenda

• Introduction & Overview

• Leadership
  – Theory
  – Methodology

• Medical Outcomes
  – Patient Safety & Errors
  – Other Data

• Patient Consent

• Summary & Questions

Patient Consent (pg. 2-3)

• Explain project to them
• Explain why we’re doing it
• Explain exactly what is required of them (nothing right now, brief phone call in 30 days)
• Ask if they’d be willing to help us out
• Give sheets to them, give them time to look it over
• Ask if they have any questions
• Thank them profusely!
Agenda

• Introduction & Overview

• Leadership
  — Theory
  — Methodology

• Medical Outcomes
  — Patient Safety & Errors
  — Other Data

• Patient Consent

Summary & Questions

Study Overview

• Sampling

• Pre-surgery
  — Patient consent
  — Patient & surgery information
  — Expected complexity

• During-surgery
  — Timing
  — Checklists
  — Leadership behaviours
  — Errors & complications

• Immediately post-surgery
  — Post-hoc complexity
  — Team surveys
  — Observer surveys

• Upon discharge
  — Patient records

• 30-day follow-up
  — Phone call
Appendix C – Study 2: Observation Sheet

Observation Sheet

Your name: ____________________  Observation #: ______
Date of surgery: ______________________

Patient Information
Gender:  Male___ Female ___  CR#: ___________________  Year of Birth: ______
ASA: __________
Date admitted to hospital: ______________

Surgery Information
Department: __________
Type of surgery:  Elective___  Emergency ____
Surgeon-rated complexity (scale of 1-10): ______  Comments: ________________________________

Type of anesthesia:  □ General  □ Spinal  □ Neurolept  □ Local

Anesthetist-rated complexity (scale of 1-10): ___  Comments: 
Nurse-rated complexity (scale of 10): ___  Comments: 

Expected blood loss:___________
**Team members (list everyone in the room and their role):**

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please also write names on team members’ surveys.
**During Surgery**

Time observer enters the room: ________
Try to enter with the first team members, but if not possible, please list team members who are already in the room:

Time patient enters the room: ________
Time surgeon enters the room: ________

Time anesthesiologist begins anesthetic: ________

Time surgery begins: ________ (refer to list of procedures to determine what constitutes the beginning of the surgery)

Time surgeon finishes: ________
Time begin waking up patient: ________
Time patient wakes: ________

Leadership behaviours, errors & complication events – see attached sheet

**Post Surgery**

Distribute questionnaires to team members, and ensure you collect them by the end of the shift.
### Leadership behaviours & errors/complications

<table>
<thead>
<tr>
<th>Time</th>
<th>Error/Leadership (circle)</th>
<th>If error/complication, describe</th>
<th>If leadership, indicate which type (circle)</th>
<th>Additional Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E L</td>
<td></td>
<td>TFL LF OC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E L</td>
<td></td>
<td>TFL LF OC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E L</td>
<td></td>
<td>TFL LF OC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E L</td>
<td></td>
<td>TFL LF OC</td>
<td></td>
</tr>
</tbody>
</table>

(cont)

**General Comments:**
Appendix D – Study 2: Team member survey

Name ____________________________

**Psychological Safety**

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th></th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. If you make a mistake on this team, it is often held against you.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Members of this team are able to bring up problems and tough issues</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. People on this team sometimes reject others for being different.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. It is safe to take a risk on this team.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. It is difficult to ask other members of this team for help.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. No one on this team would deliberately act in a way that undermines my efforts.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Working with team members of this team, my unique skills and talents are valued and utilized.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Boredom**

*How much of the time during the surgery did you experience the following:*  

<table>
<thead>
<tr>
<th></th>
<th>Never/almost never</th>
<th>Infrequently</th>
<th>Some of the time</th>
<th>Most of the time</th>
<th>Almost all the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feeling bored</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Feeling that I wish to do something else now</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Feelings of monotony</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Feeling that time goes very slowly</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Feeling that nothing happens</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Feeling that I wish to be at the end of the surgery now</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Demographics (on a separate page collected at time of consent form signing)**

- Age ______
- Profession (surgeon, resident, nurse, etc.) ___________
- Tenure (years working in medical field, excluding university training, medical school/university, but including residency) ________ in years
- Tenure at this hospital (including residency) ________ in years
- Gender: M ____ F ____
Appendix E – Study 2: Postoperative outcomes to be collected

Complete this sheet 30 days following the surgery, using patient records for Upon Discharge information, and by phone call for the 30-day follow-up.

Upon discharge: Completed by: ______________________
Date of discharge: ______________

Please list any post-operative complications:

Did the patient receive any antibiotics? _______
Did the patient spend time in the ICU? _______ If yes, how long?
Did the patient receive any blood transfusions? _______ If yes, how many units? _______

30-day follow-up: Completed by: ______________________
Date: _______________
Did the patient have any unplanned visits to: The emergency room? _______
Their primary physician? ________
A walk-in clinic? ________
The OR? _______
Other? _______

Did the patient receive any antibiotics since discharge? _______

Did the patient have any other post-operative complications? ___________
If yes, please describe: