EXPLORING A NEW METHODOLOGICAL APPROACH FOR CAPTURING THE ‘SLOWING DOWN’ MOMENTS OF OPERATIVE PRACTICE

by

Laurent Eskandar St-Martin

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

Institute of Medical Science
University of Toronto

© Copyright by Laurent St-Martin 2014
Exploring a New Methodological Approach for Capturing the ‘Slowing Down’ Moments of Operative Practice

Laurent Eskandar St-Martin

Master of Science
Institute of Medical Science
University of Toronto

2014

Abstract

The phenomenon of ‘slowing down’ in response to important cues in the operative field is proposed as a hallmark of expert surgical judgment. As part of a larger program of research, the purpose of this study was to explore a methodology for capturing ‘slowing down’ moments using a standardised task. Edited videos of 6 laparoscopic cholecystectomies were shown to 10 expert surgeons (>250 laparoscopic cholecystectomies completed). Participants were asked to think aloud while watching them as if observing each procedure in the operating room. Each session was audiotaped and transcribed. Many examples of ‘slowing down’ moments were identified in the transcripts, including several categories that were previously uncharacterised or undescribed. A subset of ‘slowing down’ moments was compared between participants. Many appeared to be inconsistent between expert surgeons, suggesting that with this methodology alone, formal teaching and assessment of the ‘slowing down’ phenomenon will be challenging.
Acknowledgments

These last 2 years have been quite an adventure and I couldn’t have done it without the help and support of many great individuals.

First and foremost, thank you to my supervisor Carol-anne Moulton. I came to you without much of a CV (I think I called it a resume back then) and you took me in anyway. This wasn’t an easy project for me but you were patient and encouraging as I worked my way through it. Your love for what you do is contagious and made me want to be as good a researcher as I could be. I have learned a great deal from you. It’s been fun and a privilege working with you. Thank you for everything.

To my committee members – Steve Gallinger, Helen MacRae and Ranil Sonnadara – and also to Glenn Regehr, thank you for your time and patience. Your feedback and ideas have without a doubt made this thesis a better piece of work.

A big thank you goes to Nathan Zilbert, who has been an important part of this research project from the moment he landed on the graduate scene last summer. I’m grateful for your patience reading through each transcript and answering my questions, your ideas and feedback, and your friendship.

To Priyanka Patel and Carween Mui, thanks for the laughs. It’s been fun sharing an office with you.

To the Wilson Centre in general, thank you for making me feel at home. It’s been a real pleasure working here. Without the many great people who make up this place, this experience would not have been the same.
To my family, thank you for your continued moral support. I couldn’t ask for more encouraging people. The financial support was appreciated too.

And to Jess, thanks for being by my side. I know this thesis (and the occasional grant) has taken up a lot of my time over the last 2 years. Thanks for your support, in particular in these last few months as I scrambled to complete this thesis. I love you.
Table of Contents

Chapter 1: Literature Review p. 1
Chapter 2: Research Aims p. 27
Chapter 3: Methods p. 28
Chapter 4: Results p. 39
Chapter 5: Discussion p. 72
Chapter 6: Conclusions p. 83
Chapter 7: Future Directions p. 84
References p. 86
Appendices p. 96
List of Tables

Table 1: Brief description of the 6 edited laparoscopic cholecystectomy videos presented to a sample of expert surgeons

Table 2: Demographic statistics for the 10 surgeon participants

Table 3: Level of consistency between participants in the expression of individual moments of concern

Table 4: Participants’ expressions of concern in response to the operating surgeon’s decision to remove a clip that had been placed on the cystic duct in Case 1
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Framework for understanding surgical judgment</td>
<td>2</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Categories of ‘slowing down’ moments</td>
<td>9</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Manifestations of the 'slowing down' phenomenon observed in the operating room</td>
<td>11</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Endsley model of situation awareness in dynamic decision making</td>
<td>16</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Moments of concern (MCs) expressed by participants during Case 1</td>
<td>59</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Moments of concern (MCs) expressed by participants during Case 2</td>
<td>60</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Moments of concern (MCs) expressed by participants during Case 3</td>
<td>61</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Moments of concern (MCs) expressed by participants during Case 4</td>
<td>62</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Moments of concern (MCs) expressed by participants during Case 5</td>
<td>63</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Moments of concern (MCs) expressed by participants during Case 6</td>
<td>64</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Refinement of the framework of categories of ‘slowing down’ moments.</td>
<td>73</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Conceptual drawing showing continuous fluctuations in the surgeon’s cognitive effort over the course of a surgical procedure</td>
<td>75</td>
</tr>
</tbody>
</table>
List of Appendices

Appendix A: Think Aloud Training Exercises

p. 96
Chapter 1
Literature Review

1.1 Introduction

The traditional approach to surgical training is based on an apprenticeship model dating back to William Halsted (1904). It is presumed that, given enough time under the direct supervision of experienced surgeons, the trainee will learn all the knowledge and skills needed to become a competent independent practitioner. However, as a result of recent changes such as resident work-hour limitations and reduced trainee independence, this unstructured model is no longer acceptable (Grantcharov & Reznick, 2009). Surgical educators have been working towards a more efficient training process, prompting a shift towards a more structured model of training built around specific competencies (Grantcharov & Reznick, 2009; Reznick & MacRae, 2006). Important progress has been made in the area of technical competencies, including increased use of simulation (Gauger et al., 2010; Grantcharov et al., 2004) and the proliferation of objective assessment tools (Martin et al., 1997; Vassiliou et al., 2005; Vassiliou et al., 2010). Judgment, however, remains an elusive construct that is challenging to teach and assess. For the most part, trainees learn it through ‘osmosis’ and assessment is often informal and subjective.

In order to make surgical judgment more explicit, a framework was developed in recent work (Figure 1). Much of this work focused on a particular aspect of the framework referred to as ‘slowing down when you should’ (Moulton, 2010). In essence, ‘slowing down’ marks the transition from a routine mode of thinking to a more effortful one in the critical or uncertain moments of practice. This transition relies on the perception of critical cues; for example, the suspicion of a tumour during a routine procedure will prompt a surgeon to ‘slow down’ and
deliberate the best course of action. Moulton et al. (2007) propose that the ability to perceive the cue and ‘slow down’ appropriately is a hallmark of expert judgment.

Figure 1: Framework for understanding surgical judgment (from “Exploring the role of ‘slowing down when you should’ in operative surgical judgment” by Moulton [2010])
The ‘slowing down’ framework could serve as a focal point for more formal and explicit teaching and assessment of surgical judgment. Still, how to capture this aspect of judgment is challenging because, while other competencies such as technical skills and teamwork are amenable to direct observation, judgment is situated at the level of thought. In order to make tangible the internal processes of judgment, new methods are required. To this end, the purpose of this thesis was to explore a novel methodology for capturing the ‘slowing down’ moments of intraoperative practice.

The following literature review introduces the concepts which guided this research. To begin, a brief overview is provided of the ‘slowing down’ framework and the literature upon which it is based. Next, perception and situation awareness are discussed to provide an understanding of how surgeons make sense of intraoperative cues. In the final section, some of the current methodologies used to capture surgical judgment are explored, and an approach for capturing the ‘slowing down’ phenomenon in a standardised setting is proposed.

1.2 ‘Slowing Down When You Should’

Moulton et al. conducted several qualitative studies that described and categorised the ‘slowing down’ phenomenon in a surgical context (Moulton, Regehr, Lingard, Merritt, & Macrae, 2010a, 2010b, 2010c). This section examines the literature underpinning the ‘slowing down’ framework and then reviews some of the findings of the original ‘slowing down’ studies.
1.2.1 The role of intuition and reflection in professional practice

Whether it be driving a car or diagnosing a patient’s condition, skills that were once effortful become intuitive over time (although ‘intuition’ can have several meanings, the term is used here to refer to judgment that is fast, unconscious and based on past experience). Based on their observations of learners and interviews with instructors, Fitts and Posner (1967) propose three phases of skill acquisition. In the cognitive phase, the learner seeks to “understand” the task, learning rules and concepts. Performance in this phase is effortful, as the learner must “attend to cues, events, and responses that later go unnoticed” (p.12). In the associative phase, errors become less frequent as individual skill components learned in the first phase become integrated. In the final or autonomous phase, the task becomes more automated and less subject to conscious control. Intuitive performance is thus described as the outcome of learning.

Other authors go further to propose that intuition is a (if not the) defining feature of expertise (Chase & Simon, 1973; Dreyfus, Dreyfus, & Athanasiou, 1986; Gobet & Simon, 1996). For example, a popular model of skill acquisition describes the progression from novice to expert as a transition from rule-based reasoning to experience-based intuition (Dreyfus et al., 1986). According to this model, the novice has no intuition. She recognises an objective cue and applies a rule based on that cue. However, as she progresses through the intermediate stages of learning (advanced beginner, competent, and proficient), she learns to understand meaningful patterns of cues and later entire situations. A database of situations and suitable responses develops over time such that, at the expert level, there is an intuitive response for most situations. The expert’s performance – from the interpretation of cues to subsequent decision making and action – is fluid and unconscious (Dreyfus et al., 1986).
This notion of intuitive expertise finds empirical support in the naturalistic decision making (NDM) literature, which explores how people perform amid the pressures and uncertainties of a real-life setting (Zsambok & Klein, 1997). Experts in action seldom compare options when making a decision, but instead make a quick assessment of the situation and relate it to past experiences. The recognition of a situation as familiar then brings to mind appropriate goals, cues, expectations, and courses of action (Klein, 1993). Intuition therefore relies on both vast experience (to build a database of scenarios) and good situation awareness (to detect that the scenario at hand is similar to one that has been encountered before). NDM studies of various experts (e.g., neonatal nurses [Crandall & Getchell-Reiter, 1993], battle command teams [Thordsen, Galushka, Klein, Young, & Brezovic, 1990], and naval officers [Kaempf, Wolf, & Miller, 1993]) suggest that up to 96% of decisions are made intuitively. Because these studies often focus on challenging and unusual cases, it is conceivable that the role of intuition is even greater in routine activities (Klein, 1998).

There are important benefits to an intuitive approach to practice. It allows for the quick and effective resolution of most problems. Moreover, as the thought processes responsible for task performance become unconscious, valuable cognitive resources are liberated, which can then be invested in other tasks (Kahneman, 1973). To illustrate this, Hsu et al. (2008) asked novice and experienced laparoscopic surgeons to perform a basic laparoscopic task (peg transfer) and a cognitive task (mathematical addition). Participants were first asked to perform each task in isolation, then both at the same time. Dual-tasking had little effect on technical or cognitive task performance in the experienced group, but impaired performance on the cognitive task in the novice group. As efficient as intuitive processes are, however, they are likely to be inadequate in complex or uncertain situations that fall outside the scope of routine practice. These situations
are more difficult to match to prior experiences and therefore demand a more deliberate mode of thinking (Klein, 1998).

In a model of expertise based on this deliberate mode of thinking, Bereiter and Scardamalia (1993) make the distinction between the expert and the experienced non-expert to highlight the importance of reflection in professional practice. The experienced non-expert tries to eliminate problems, reframing them to conform to learned routines. The expert, on the other hand, simplifies problems as little as possible and, in so doing, develops a deep knowledge of her field of work. While the expert’s approach is more demanding, Bereiter and Scardamalia (1993) point out that it need not be used all the time. As certain situations become familiar with experience, dealing with them becomes less effortful. As a result, both the expert and the experienced non-expert can handle most situations with minimal problem-solving. What sets the expert apart, then, is her approach to the unfamiliar. Faced with a novel or complex problem, she reflects on its nuances in order to adapt an appropriate solution, while the experienced non-expert insists on using an existing solution that might not meet the particular needs specific to the problem (Bereiter & Scardamalia, 1993).

Schön (1983, 1987) makes a similar argument. Routine tasks can be performed intuitively, without awareness of the mediating thought processes. A surfer, for example, might struggle to explain how he remains balanced in turbulent waters. Schön (1983) refers to this tacit know-how as knowing-in-action. He acknowledges its importance in practice but also stresses the importance of reflection-in-action, a deliberate process that is concurrent with the performance. This involves thinking about actions and their outcomes and adjusting performance in the moment, like a musician improvising in response to the audience or other band members (Schön, 1983). Like Bereiter and Scardamalia (1993), Schön (1983, 1987) proposes that reflection is
central to dealing with novel, unstable, or uncertain situations. He also notes that a given situation can have both routine and non-routine components, and that unexpected outcomes (pleasant or not) can evoke the transition from knowing-in-action to reflection-in-action. A general surgeon who encounters dense adhesions at the start of a ‘routine’ laparoscopic procedure, for example, might be prompted to slow down and consider alternative approaches to improve visualisation (e.g., using an additional port, converting to laparotomy).

Considered together, the expertise literature seems to suggest that both intuition and reflection are essential. Intuition is fast and often effective but is also prone to error in non-routine situations. In contrast, reflection provides for solutions that are adapted to the specific needs of particular problems but is also more resource-intensive. Perhaps, then, expert judgment lies in the proper coordination of these two modes of thinking, increasing the use of reflection when the situation requires it. The appropriate transition to a more reflective mode of thinking (referred to as ‘slowing down when you should’) relies on the perception of pertinent cues and appropriate situation awareness; this transition has been suggested to be a hallmark of expert judgment (Louis & Sutton, 1991; Moulton et al., 2007).

1.2.2 Limited attentional resources and the ‘slowing down’ transition

The mechanism of the ‘slowing down’ transition can be understood through the attention literature. There is general agreement that our attentional resources are limited (although there is some debate over the nature of these limitations [Kahneman, 1973; Wickens, 2008]). We can allocate attentional resources with relative freedom, allowing us to multi-task. However, performance suffers when the attentional demands of concurrent tasks exceed the resources available, as Hsu et al.’s (2008) experiment suggests. In these situations, attention must be
withdrawn from extraneous stimuli to focus on the task(s) considered to be most important. In a complex environment such as the operating room, numerous stimuli are competing for the surgeon’s attention – the patient, the monitors, a conversation with a colleague, etc. Confronted with a challenging event or an unexpected cue (e.g., sudden bleeding), the surgeon might withdraw from conversation (for example) in order to remain in control of the situation. In light of this literature, ‘slowing down’ can be thought of as a reallocation of attention resources, increasing focus on the procedure in response to important cues or events in order to ensure a safe patient outcome.

1.2.3 The ‘slowing down’ framework

Moulton et al. (2010a, 2010b, 2010c) explored the ‘slowing down’ phenomenon in a surgical setting in a series of qualitative studies. These studies involved semi-structured interviews with 28 academic surgeons from various specialties who had a reputation for having good surgical judgment, as well as observations of 29 operative cases performed by 5 hepato-pancreatobiliary (HPB) surgeons (the observation phase was focused on HPB cases because this was the principal investigator’s area of clinical practice). Data were collected and analysed in a manner consistent with the grounded theory methodology, the purpose of which is to generate theoretical frameworks that explain a particular social phenomenon (Kennedy & Lingard, 2006). In this section, we review some of the main findings of the original ‘slowing down’ studies.
1.2.3.1 Categorisation of ‘slowing down’ moments

As surgeons discussed their experiences with the ‘slowing down’ phenomenon, two broad categories of ‘slowing down’ moments became apparent (Moulton et al., 2010c): proactively planned moments that are anticipated prior to the operation and situationally responsive moments that occur in response to emergent intraoperative information (Figure 2).

Proactively planned ‘slowing down’ moments are critical intraoperative steps that are identified while planning for the case (Moulton et al., 2010c). As these anticipated critical steps arise during the procedure, the surgeon can make a conscious and intentional transition to a more effortful mode, minimising the risk of getting into trouble. Proactively planned ‘slowing down’ moments can be further categorised as either procedural-specific or patient-specific (Moulton et al., 2010c). Procedural-specific moments are challenging or critical technical steps that are encountered each time a particular procedure is performed. Patient-specific moments, on the

![Diagram of categories of 'slowing down' moments]

**Figure 2:** Categories of ‘slowing down’ moments (from “‘Slowing down when you should’: initiators and influences of the transition from the routine to the effortful” by Moulton et al. [2010c])
other hand, reflect the subtle nuances of an individual case that deviate from the routine. Factors such as anatomical variations and comorbidities might call for a different approach to a procedure, or perhaps a different procedural plan altogether.

Situationally responsive ‘slowing down’ moments are prompted by unexpected intraoperative findings (Moulton et al., 2010c). Some of these ‘surprises’ might be difficult to anticipate, such as the incidental finding of a tumour during a ‘routine’ procedure. However, others might be the result of insufficient preoperative planning. In contrast to planned ‘slowing down’ moments, which are proactive, the reactive and sometimes urgent nature of situationally responsive moments appears to require more complex and effortful problem-solving in order to remain in control of the operation (Moulton et al., 2010c). Some surgeons commented that when responding to something unexpected, their own transition to a more effortful mode sometimes went unnoticed until after the problem had been dealt with, suggesting that the ‘slowing down’ transition can be more gradual than sudden and that its initiation can be unconscious (Moulton et al., 2010c).

1.2.3.2 Manifestations of the ‘slowing down’ phenomenon

During the interviews, participants described observable behaviours that were sometimes associated with their ‘slowing down’ moments in the operating room. Expanding on these descriptions with observational data, Moulton et al. (2010b) characterised 4 observable manifestations of the ‘slowing down’ phenomenon in operative practice (from most extreme to most subtle): stopping, removing distractions, focusing more intently and fine-tuning (Figure 3).
**Figure 3:** Manifestations of the 'slowing down' phenomenon observed in the operating room (from "Slowing down to stay out of trouble in the operating room: remaining attentive in automaticity" by Moulton et al. [2010b])

*Stopping* was the most pronounced manifestation and involved a momentary pause in the progress of the procedure (Moulton et al., 2010b). Some stops were unexpected and were associated with a recognised need to gather more information before continuing on with the procedure. For example, upon encountering unexpected disease, a surgeon might have stopped for a moment to consult with a colleague. Other stops were planned ahead of time and were intended to prepare for a critical step of the procedure (a proactively planned ‘slowing down’ moment), making sure that the operating room personnel was in position and that the required materials were available (Moulton et al., 2010b).

A less extreme manifestation of the ‘slowing down’ phenomenon was *removing distractions* (Moulton et al., 2010b). During these ‘slowing down’ moments, the surgeon continued to operate but became distracted by a particular stimulus (e.g., the background music, a conversation taking place) and requested that the distraction be removed. This is consistent with the concept of limited cognitive resources described earlier; the removal of the distraction liberated cognitive
resources that the surgeon could re-invest into the surgical problem. The surgeon did not appear to be distracted by the stimulus in question earlier in the case; it became bothersome during the ‘slowing down’ moment due to the increased cognitive demands of the emergent situation.

Sometimes, a surgeon who had been engaged in extraneous activities, such as a conversation with the operating room team, withdrew from these activities to focus on the surgical task at hand (Moulton et al., 2010b). This manifestation was referred to as focusing more intently. In contrast with the previous manifestation (removing distractions), the surgeon did not appear to be bothered by the surrounding noise or conversations.

The most subtle of the four observable manifestations was fine-tuning (Moulton et al., 2010b). Fine-tuning moments were minor ‘slowing down’ moments that might occur numerous times over the course of a surgical procedure. These moments occurred in response to emergent cues and involved a technical readjustment for the purpose of keeping the procedure safe and on track. Moulton et al. (2010b) describe one particular example of fine-tuning in an observational field note:

“The surgeon was encircling the portal vein at the hilum of the liver. He recognized the vein was bigger than anticipated and pulled the dissecting instrument out to then encircle the much larger vein. He was conversing throughout this maneuver with the fellow about the prognosis of the case and “paused” momentarily as he corrected the angle of the right-angled instrument.” (p. 1574)

As this example suggests, a surgeon might remain engaged in extraneous activities while fine-tuning. In fact, fine-tuning moments are so subtle that they are unlikely to be noticed by someone without considerable knowledge of the procedure (Moulton et al., 2010b).
The four observable manifestations of ‘slowing down’ are representative of a spectrum of cognitive requirements (Moulton et al., 2010b). More demanding moments of the procedure will generally be associated with more pronounced manifestations (to liberate enough cognitive resources to adequately manage the situation) and vice versa. However, these four manifestations should not be thought of as a linear sequence. In other words, it is possible for a surgeon to progress from fine-tuning to stopping (for example) without exhibiting the intermediate manifestations (focusing more intently and removing distractions) (Moulton et al., 2010b).

1.2.4 Implications of the ‘slowing down’ framework for research in surgical education

Through the ‘slowing down’ studies, a framework was created that made explicit a component of expert judgment that had until then been implicit in surgical practice. As part of this framework, it was established that appropriate intraoperative ‘slowing down’ relies on the preoperative anticipation of challenging moments through the perception and interpretation of relevant cues (e.g., tumour invasion of adjacent organs as seen on CT imaging) as well as the intraoperative detection of challenging moments through the perception and interpretation of emergent cues (e.g., a gush of blood) (see Section 1.2.3.1). This provided a basis for further, more focused research on surgical judgment. The present study focused on capturing ‘slowing down’ moments and the perception and interpretation of the emergent intraoperative cues that initiate them.

1.3 Perception and Situation Awareness

By definition, ‘slowing down’ moments are initiated by a cue that causes a transition from the routine to the effortful. In the field of surgery, this cue is often perceptual in nature. In this
section, the importance of perception in professional practice is explained using Schön’s (1983) concept of naming and framing, and the construct of situation awareness is introduced to describe how individuals manage perceptual information in complex environments.

1.3.1 The role of perception in professional practice

Situations in professional practice are often uncertain and complex. To solve a problem, the practitioner must first define the problem, figuring out “what is wrong and in what direction the situation needs to be changed” (Schön, 1983, p. 40). Schön (1983) suggests this happens through a process of ‘naming and framing,’ where *naming* is the identification of relevant cues and *framing* is the interpretation of these cues to define the problem at hand. For example, a surgeon might notice a dilated common bile duct (naming) and conclude that the bile duct is obstructed (framing). Once the problem has been named and framed, the practitioner can move on to problem solving and decision making. Continuing with our example, the surgeon might think about possible causes of obstruction (e.g., an impacted stone, a tumour) and how to make the diagnosis. Perception is therefore a fundamental aspect of clinical judgment because failing to perceive (or name) a cue could lead to a course of action (e.g., a treatment plan, an operative manoeuvre) that is ineffective or even harmful.

Superior perceptual abilities develop with expertise. Experts have the “power to see the invisible” (Klein, 1998), perceiving meaningful information that others cannot. Compared to novices, experts are better able to recognise patterns, detect anomalies, and make subtle discriminations. For instance, de Groot (1965) asked chess masters and non-masters to recall 10 mid-game positions, taken from actual chess games, after 2-5 seconds of viewing. Masters were able to recall more than 90% of the pieces while non-masters recalled about 50-70%.
Retrospective reports revealed that, despite the time pressure, the masters were noticing meaningful configurations of pieces (e.g., a castled position) that could be encoded as large “chunks” of data. However, when Chase and Simon (1973) repeated the experiment using random positions, a master did no better than weaker subjects. This suggests that experts’ perceptual abilities (e.g., detecting patterns in chess pieces) are not innate and generalisable but are instead specific to their domain of experience.

In surgical practice, certain decisions can be made based on static perceptual information. As an example, a surgeon might use CT imaging to evaluate whether a tumour is resectable. The operating room, however, is a complicated environment where available information is dynamic, abundant, and sometimes incomplete. Achieving a good surgical outcome requires being able to perceive and interpret emergent information throughout the procedure.

### 1.3.2 Situation awareness

Situation awareness (SA) is “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (Endsley, 1988, p. 792). In simpler terms, it is an ongoing understanding of what is happening in a dynamic environment. One prominent model describes three hierarchical levels of SA (Figure 4) (Endsley, 1995b). In Level 1 (perception), task-relevant cues are extracted from the environment. Then, in Level 2 (comprehension), these cues are interpreted to generate a holistic understanding of the current situation. Based on this understanding, we can then anticipate what might happen in the near future in Level 3 (projection). The information upon which SA is built can be acquired through direct observation, system interfaces, and other people (Endsley, 2000).
Figure 4: Endsley model of situation awareness in dynamic decision making (adapted from “Toward a theory of situation awareness in dynamic systems” by Endsley [1995])

Good SA requires the appropriate allocation of attentional resources. This allocation is mediated by both bottom-up and top-down processes (Endsley, 1995b). Bottom-up processes are perception-driven and spontaneous, directing attention to salient cues in the environment, such as sudden movements or bright objects (Connor, Egeth, & Yantis, 2004). In a surgical setting, a bottom-up cue such as a stream of blood might prompt a situationally responsive ‘slowing down’ moment. In contrast, top-down processes are cognition-driven and deliberate, directing attention to expected cues (Connor et al., 2004). Proactively planned ‘slowing down’ moments, for example, encourage the surgeon to seek out particular structures to avoid getting into trouble.

Due to the limitations of attention, information sampling has to be strategic, taking into consideration the relative importance of cues (Endsley, 1995b).

Goals are also important, dictating the requirements of SA. In a surgical team, for instance, the surgeon, nurse and anaesthetist each have distinct roles and responsibilities. It follows that what each team member must know about the same environment (the operating room) will be quite different. Goals direct the selection of schemas, the allocation of attention (through top-down processing), and the interpretation of perceived information (Endsley, 1995b). At the same time,
bottom-up cues can prompt the selection of a new goal (Endsley, 1995b). As an example, a person driving along a familiar route (initial goal) might notice an animal that has run onto the road, prompting him to brake to avoid a collision (new goal). SA is therefore the product of a continuous interaction between top-down and bottom-up processes, with new information altering SA which then guides the search for subsequent information (Endsley, 2000). This iterative process of information gathering and interpretation is known as situation assessment, and is distinct from SA which is the knowledge that results from this process (Endsley, 2000).

There is an important link between SA and performance outcomes because our understanding of a situation informs the decisions we make and the actions we take. Poor SA has contributed to accidents in aviation (Endsley, 1995a), offshore drilling (Sneddon, Mearns, & Flin, 2006) and nuclear power plants (Kemeny, 1979). Poor SA has also been associated with medical error. In a review of 252 laparoscopic bile duct injuries, misperception was the main cause of error in 97% of incidents (Way et al., 2003). Some of these injuries involved the deliberate cutting of a misidentified duct, while others occurred because the duct was obscured and the surgeon believed he or she was dissecting a safe distance from it. In contrast, 3% of bile duct injuries were attributed to poor technical skill (Way et al., 2003). Therefore, the importance of SA in clinical practice cannot be understated.

Recent research has attempted to capture SA in a clinical context. Hogan et al. (2006) used the Situation Awareness Global Assessment Technique (SAGAT) to assess trauma management skills in 3 simulated scenarios. SAGAT was first developed to measure pilot SA in a flight simulator and involves pausing a simulation at random moments and asking a series of questions assessing the participant’s SA at all 3 levels (perception, comprehension, projection) (Endsley, 1988). Hogan et al. (2006) found that SAGAT scores increased with level of training and
correlated with traditional checklist scoring, suggesting that SAGAT is valid in a simulated clinical setting. SA is also gaining recognition in surgical assessment (Sevdalis et al., 2008; Yule et al., 2008). For instance, the Non-Technical Skills for Surgeons (NOTSS) rating tool features SA as one of its four main categories of non-technical skills, alongside decision making, communication and teamwork, and leadership (Yule et al., 2008). However, because this tool evaluates observable behaviours, ratings of SA are based on the surgeon’s interactions with team members and objects outside of the operative field (e.g., consulting a CT scan, warning the anaesthetist about anticipated blood loss). To date, little research has been done to capture the surgeon’s SA with respect to the operative field, perhaps because this aspect of SA is more difficult to observe. For example, the surgeon might not exhibit an observable behaviour when she notices an important anatomical landmark. SA of the operative field is in this sense ‘hidden’ in the surgeon’s mind and therefore capturing and exploring it requires alternative methodological approaches, described below.

1.4 Methodologies for Capturing Surgical Judgment

In this section, an overview of how surgical judgment has been captured in previous work is provided. Some methods that might be used to capture ‘slowing down’ moments are considered, and the importance of standardising tasks for the purposes of teaching and assessment is discussed.

1.4.1 Previous efforts to capture surgical judgment

The current gold standard for capturing surgical judgment is the oral examination (Meterissian, Zabolotny, Gagnon, & Charlin, 2007). Participants are presented with a clinical scenario and are
asked to work through the diagnosis and/or management while thinking aloud. This approach has also been used to compare the judgment of expert and novice surgeons in a research setting (Abernathy & Hamm, 1994). The oral examination is useful for capturing problem solving and decision making abilities. However, perception is overlooked since the relevant cues are provided for the participant (e.g., “on exploring the abdomen you feel a 4-5-cm vague mass in the head of the pancreas” [Abernathy & Hamm, 1994, p. 55]). In other words, it cannot be known whether the participant would have noticed these cues and identified them as important without prompting.

A recent alternative approach to capturing surgical judgment is the Script Concordance Test (SCT) (Charlin, Roy, Brailovsky, Goulet, & van der Vleuten, 2000; Meterissian et al., 2007). In this test, participants are presented with written clinical scenarios, and are asked to rate the influence of a new piece of clinical information (e.g., a symptom, a test finding) on a possible diagnosis, investigative action, or treatment option using a Likert scale (Charlin et al., 2000). The SCT acknowledges that clinical problems are often ambiguous, and therefore no single correct answer is expected. Rather, each answer is assigned a value based on the degree of agreement among a panel of experts (i.e. the greater the proportion of panel members who choose a particular answer, the more that answer is worth). According to Meterissian et al. (2007), “challenging authentic clinical scenarios, incomplete clinical data, and multiple possible correct answers…are what distinguish [the SCT] from multiple choice questions and oral examinations” (p. 249). The SCT has several appealing features, such as its ease of administration and objective scoring. Perhaps most important, it has been shown to discriminate between levels of training in a number of specialties (Charlin et al., 2006; Meterissian et al., 2007; Sibert et al., 2002); for example, an SCT for intraoperative decision making found a significant difference between the scores of junior and senior general surgical residents (Meterissian et al., 2007). However, like the
oral examination, the SCT still provides the participant with the clinical findings that require interpretation.

Few studies on surgical judgment have incorporated a perceptual component into their design (Chatterjee, Ng, Kwan, & Matsumoto, 2009; Dominguez, 1997; Ghaderi, Ericsson, Farrell, & Harris, 2012; Samuelson, Cadeddu, & Matsumoto, 2006). For example, Ghaderi et al. (2012) aimed to capture and describe the cognitive processes that occur while performing a laparoscopic cholecystectomy. To do this, the authors asked experienced surgeons to think aloud while watching videos of their own operative cases. Ghaderi et al. (2012) found the surgeons were monitoring the operative field throughout the procedure, using perceptual information to plan and anticipate future events. For example, surgeons looked for cues (e.g., abnormal structures) that might require a deviation from the standard course of action. In other work, Dominguez (1997) explored factors involved in the decision to convert a laparoscopic case to an open procedure. She conducted interviews with surgeons and residents who were asked to think aloud while watching an edited video of a challenging laparoscopic cholecystectomy, imagining themselves performing the operation. About half of the participants felt at one point or another that the case should be opened. Among other findings, Dominguez (1997) noted that fewer ‘openers’ mentioned the pulsation of a particular structure that had just been divided (which would suggest a blood vessel); failing to notice this important cue would have meant poorer SA, which might in turn have contributed to the discomfort associated with the decision to open.

These last studies suggested to the research team that the think aloud method could capture a surgeon’s continuous perception and interpretation of emergent cues, making this method a potential candidate for capturing the ‘slowing down’ moments of intraoperative practice. It was
unclear just how ‘slowing down’ moments would be manifested in a think aloud transcript, but Dominguez’s data suggested that a surgeon would vocalise concern in response to critical cues:

“I am kind of worried…I can’t see where the common bile duct and the cystic duct join together. There is all this fat and inflammation there and if it were me, I am digging into things that I can’t identify.” (Dominguez, 1997, p. 146)

1.4.2 The think aloud method

The ‘think aloud’ method has been used to capture expert judgment in domains such as chess (de Groot, 1965), education (Swanson, O’Connor, & Cooney, 1990) and medicine (Boshuizen & Schmidt, 1992; Ghaderi et al., 2012). It was first described by the behaviourist John Watson in 1920 as a superior alternative to introspection (self-observation), which was popular at the time among researchers interested in cognition (Watson, 1920). Watson argued that introspection could not be trusted, using his experience with golfers as evidence:

“No one since objective studies upon golf have been made trusts the [introspective] report of a golf player. He will tell you that he never takes his eyes off the ball when making a stroke. The cinema shows that he is a prevaricator. I have never been able to get one valuable scientific statement from a golf player. He does not know how he holds his hands, he cannot tell how he stands, nor the arc he makes with his club, nor whether the arc can vary within wide limits and not affect his stroke. He knows practically nothing about the condition his body is in.” (Watson, 1920, pp. 100-101)

Instead, Watson proposed that asking a participant to think aloud while working on a specific problem would be more reliable. Watson (1920) believed that thinking was nothing more than
“silent talking” and therefore that it engaged the same neural processes that are involved in normal speech; for this reason, he reasoned that the instruction to think aloud would not distort the content of a participant’s thoughts.

Several decades later, Ericsson and Simon (1980) proposed that the verbalisations generated during a think aloud experiment represent the information stored in short-term memory (STM) at a given time. Some information is stored in STM in verbal code and can be directly verbalised without any additional processing. Information can also be stored in STM in nonverbal code (e.g., visual imagery) and verbalising this information requires some intermediate processing to convert it to verbal code (Ericsson & Simon, 1980). This additional processing is unlikely to alter the contents of STM (and therefore performance) in a significant way (as long as the conversion from nonverbal to verbal code is not too complicated for the subject), but may prolong the time required to perform the task (Ericsson & Simon, 1980). STM has a limited capacity for information, which means that its contents can easily be replaced by new incoming information. However, some of the outgoing information in STM may be stored in long-term memory for later retrieval before it is entirely lost (Ericsson & Simon, 1980).

In a think aloud study, participants are asked to verbalise the inner thoughts involved in performing a specific task, without providing explanations or justifications (Ericsson & Simon, 1993). The collection of verbal reports can be either concurrent with the task (i.e., verbalising thoughts as they occur) or retrospective (i.e., verbalising thoughts that occurred at an earlier time). The retrospective think aloud method is best suited for short tasks (2-10 seconds) and for perceptual-motor tasks in which timing is critical and a slowing down of thinking could have severe consequences, such as slalom skiing (Ericsson & Simon, 1993). However, in most situations the concurrent approach is preferable because it avoids some of the concerns
associated with retrospection, such as hindsight bias and the fallibility of memory (Ericsson & Simon, 1993).

Once collected, the think aloud data can be analysed using several methods. For example, protocol analysis involves breaking up the data into small segments (e.g., individual sentences), which are then coded in isolation of one another (Ericsson & Simon, 1993). The goal of this method is to develop a deep understanding of the thought process (i.e., the specific sequence of thoughts involved in completing a task). In contrast, content analysis involves the identification of recurrent themes in the data (Hsieh & Shannon, 2005). The purpose of this method is to develop a more general conceptual understanding of the phenomenon of interest.

There is considerable evidence to suggest that think aloud data are reliable (for an extensive review of the literature, see Ericsson and Simon [1993]). For example, computer models generated from verbal reports can approximate human performance on tasks such as move selection in Chinese checkers (Anzai, 1977). Moreover, a recent meta-analysis of 94 studies found no difference in accuracy of performance while thinking aloud in comparison to a silent control condition irrespective of whether the task was visual or nonvisual, which suggests that thinking aloud does not interfere with the natural sequence of thoughts (Fox, Ericsson, & Best, 2011).

However, the think aloud method also has certain limitations. Tasks become more automatic with practice and thoughts become more difficult to articulate; some of the intermediate steps between stimulus and response might therefore be inaccessible (Ericsson & Simon, 1993). Thoughts can occur faster than speech (Ericsson & Simon, 1993), which can make thinking aloud more challenging during a perceptual task that involves fast-changing stimuli. The ability to think aloud may also be diminished when participants are under high cognitive load (Ericsson
A verbal report would therefore represent an incomplete account of a surgeon’s thoughts. However, since this study was interested in the ‘slowing down’ moments of a surgeon, capturing his or her response to a cue (for example, concern or surprise) was more important than capturing the exact sequence of thoughts that led to that response. For this reason, the think aloud method seemed an appropriate choice to begin to explore the ‘slowing down’ phenomenon.

1.4.3 Other potential methods for capturing the perception of ‘slowing down’ moments

There were additional methods of data collection that may also have been appropriate. ‘Slowing down’ is a transition to a more effortful mode of thinking. A direct investigation of ‘slowing down’ moments would therefore have required a continuous measure of cognitive effort. One such measure is pupil diameter, which shows small increases with cognitive effort (Hess & Polt, 1964). This measure is reliable, sensitive and consistent in tasks that involve simple visual stimuli (e.g., Porter, Troscianko, & Gilchrist, 2007). However, variations in luminance and colour can introduce error into the pupil measurements, making this approach problematic for use with complex visual stimuli (e.g., the operative field) (Porter et al., 2007). Another direct approach would have been to measure event-related brain potentials (ERPs), which are electroencephalogram (EEG) components associated with stimulus recognition and processing (Kappenman & Luck, 2012). Some components such as the P300 are thought to index the level of attention engaged in processing a stimulus (Fabiani, Gratton, & Federmeier, 2007), so ERPs could have alerted the research team to the recognition of critical cues, thus marking ‘slowing down’ moments. However, interpreting the data would have been complex and time-consuming. It was also unclear whether it would have been possible to separate artifacts introduced by eye
movement from the EEG signals of interest (Fabiani et al., 2007), further limiting the value of this measure.

Although these methods could have been used to detect increases in cognitive effort (i.e., ‘slowing down’ moments), the limitations described would have made their implementation challenging. Moreover, these methods would not have provided access to how surgeons interpret perceived cues (i.e., what is it about the situation that prompts a ‘slowing down’ moment?), which is important information from an educational standpoint. For this reason, this study focused on the think aloud method in an initial effort to capture ‘slowing down’ moments.

1.4.4 The importance of task standardisation

The choice of task was another important consideration in developing an approach to capturing ‘slowing down’ moments. According to Ericsson and Smith’s (1991) guidelines for research on expertise, when designing a task that captures a component of expert performance (e.g., the ‘slowing down’ phenomenon), it is important to choose a task that can be standardised to allow comparisons between individuals. From a teaching perspective, a standardised task would make it possible to compare the ‘slowing down’ moments of experts and novices; differences between the two groups could then be used to inform the design of learning interventions (Williams & Ericsson, 2005). From an assessment perspective, a standardised task would help ensure a fair and objective evaluation of performance for all individuals.

Asking surgeons to think aloud while performing an actual surgical procedure might have provided the most ‘authentic’ data for capturing the ‘slowing down’ moments of surgeons but would not have allowed standardised comparisons between surgeons due to the confounding factors of each case (e.g., extent of disease, available personnel). In addition, this task would
have been unacceptable from an ethical point of view given the potential for interference with patient care. An alternative would have been to ask surgeon participants to think aloud while performing a procedure in a standardised simulated environment. While this would have allowed participants to control their own environment and therefore led to more natural ‘slowing down’ moments, current simulators lack the perceptual complexity and nuances that real cases provide. Alternatively, pre-recorded intraoperative videos of surgical procedures would allow standardised comparisons between participants while preserving some of the subtle and complex nuances of real-life cases. This study explored the ‘slowing down’ moments elicited using this approach with the surgeon participants viewing operative videos of other surgeons.

1.5 Summary

The ongoing transition towards a more structured model of surgical training relies on formal and explicit teaching and assessment of specific competencies. While progress has been made in certain areas, how to address abstract competencies such as judgment remains challenging. In recent work, the phenomenon of ‘slowing down’ in response to critical intraoperative cues has been identified as a hallmark of expert surgical judgment. As a first step towards teaching and assessment of judgment, this study explored a methodology intended to capture ‘slowing down’ moments in a standardised setting using videotaped surgical procedures in combination with the think aloud method.
Chapter 2
Research Aims

The aim of this thesis was to explore a new methodology for capturing ‘slowing down’ moments under standardised conditions. The methodology consisted of asking surgeons to think aloud while watching videotaped surgical procedures. This study aimed to answer 2 specific research questions:

1. **How are ‘slowing down’ moments manifested in the context of thinking aloud while watching videos of surgical procedures?**
   Moulton et al. (2010c) developed a framework characterising the ‘slowing down’ moments of surgeons in actual practice. The aim of this research question was to characterise the ‘slowing down’ moments that are captured in a standardised setting using videos of operations of other surgeons.

2. **Can ‘slowing down’ moments be compared between individual surgeons?**
   The aim of this research question was to determine whether it is possible to compare precise ‘slowing down’ moments between individuals (i.e., in response to event X, surgeon A ‘slowed down’ but surgeon B did not). Having an approach that would permit comparisons between individuals was important for future teaching and assessment purposes. For example, in comparing the ‘slowing down’ moments of experts and trainees it might be possible to find specific differences that could inform the design of learning interventions for surgical judgment.
Chapter 3
Methods

3.1 Overview
Ten expert general surgeons were asked to think aloud while watching edited videos of laparoscopic cholecystectomy (LC). The session transcripts formed the data set. ‘Slowing down’ moments were then identified in the transcripts. In an effort to make comparisons between participants, a subset of these ‘slowing down’ moments (which were termed ‘moments of concern’) were counted and reported using descriptive statistics. This chapter provides a description of the research team and a detailed description of the study design and methods.

3.2 Research Team
The primary researcher was a graduate student with an undergraduate background in basic science. He had never been involved in qualitative or education research prior to the start of this thesis. There were six other members of the research team. The project supervisor was an HPB surgeon and education scientist with a long-standing interest in surgical judgment that began with her PhD dissertation exploring the ‘slowing down’ phenomenon. She had expertise in qualitative methodologies, as well as clinical expertise in the LC procedure. The third member of the team was a senior general surgical resident and Master of Education candidate with a research interest in surgical judgment. The fourth team member was a colorectal surgeon with extensive experience in surgical education and was on the original team that studied the role of ‘slowing down’ in surgical judgment. The fifth team member was an HPB surgeon with clinical expertise in the LC procedure. The sixth member of the team was a scientist whose research interests include surgical skill acquisition and evaluation of performance. The final team member
was a cognitive psychologist and health professions education researcher who contributed expertise in research design.

3.3 Ethics Approval

This study was approved by the University Health Network research ethics board and the University of Toronto office of research ethics.

3.4 Participants

This research was conducted with a sample of expert surgeons; it was aimed to recruit a homogeneous group of participants within which minimal variation in ‘slowing down’ moments would be expected. For the purpose of participant selection, an expert surgeon was defined as a trained general surgeon with extensive experience in LC (>250 cases completed) and who was performing LCs on a regular basis in her or his current practice. There is no standard definition of expert judgment in the surgical literature. However, it is common practice in studies of technical skill to define expertise based on surgical case volumes and/or reputation among peers. In these studies, experts are often defined as having completed more than 100 cases of the procedure of interest (e.g., Aggarwal et al., 2007). It was decided to use 250 cases as a more conservative inclusion criterion. A list of “some of the best” LC surgeons in the Toronto area was obtained from a senior general surgeon to further guide the recruitment process. This particular surgeon is recognised as a pioneer in advanced laparoscopic techniques and has extensive knowledge about other general surgeons in the region through his role as site director for the University of Toronto’s Minimally Invasive Surgery fellowship training program.
Potential participants were contacted via e-mail and their volunteer participation was requested. They were informed that the purpose of this study was to explore surgeons’ responses to intraoperative perceptual information and that participation would require about 60 minutes of their time. Ten expert surgeons were recruited and participated. Saturation, the point at which no new themes emerged from the data set, guided the sample size (Morse, 1995).

### 3.5 Videos

De-identified videos of LC surgeries were obtained from a general surgeon affiliated with the University of Toronto who had collected videos from various local surgeons. LC was an appropriate procedure for this study for several reasons. First, LC is one of the most common general surgical procedures; the prevalence of this procedure provided access to a large expert population. Second, using a laparoscopic procedure enabled participants to see exactly what the operating surgeon saw during the actual procedure. And third, haptic feedback is limited in laparoscopic surgeries, making the correct interpretation of visual information even more important; misperception of visual cues during LC is the leading cause of common bile duct injuries, which are the main serious, and well-recognised, iatrogenic complication associated with this procedure (Way et al., 2003). The LC videos were presented without sound because it was thought that audio of the conversation between members of the operating room team might be distracting to the participants and might influence or distort their interpretations of the cases.

#### 3.5.1 Case selection

Six LC cases were chosen to represent a range of difficulty and operator skill as determined by the two HPB surgeons on the research team. The heterogeneous sampling of cases was intended
to determine whether certain types of cases might be more suitable than others for use with this methodology.

It appeared as though residents performed, either in part or in whole, some of the cases selected. This would likely have implications for the data generated. For example, due to a reduced level of technical and decision making skill, it was possible that residents would elicit more ‘slowing down’ moments from the participants than would more experienced operating surgeons. In contrast, it was possible that participants would be less interested and engaged in the task when watching a resident’s performance than when watching the performance of a surgeon at a level of experience closer to their own. Again, the intent was to understand which types of cases might work best with this methodology.

3.5.2 Video editing

The six videos were edited in Windows Live Movie Maker Version 2011 (Microsoft Corporation, Redmond, Washington), focusing on the more important steps of the procedure (e.g., exposing Calot’s triangle, clipping and dividing structures) as well as notable case-specific events (e.g., perforation of the gallbladder, management of severe bleeding). It was decided to edit the cases because ‘slowing down’ is known to occur during the critical or uncertain moments of a case; removing some of the quiet or uneventful segments of each case would therefore maximise opportunities to capture the phenomenon of interest while also minimising participant fatigue. An effort was made to edit each case in a way that would provide participants with enough information to make informed judgments and that would make it easy for them to follow the flow of the operation. For the sake of time and also to keep participants interested and engaged, some cases were edited to stop in the middle of the operation once the events of interest
had passed. Decisions about which segments of a case to retain and which to remove were made with the help of the project supervisor.

Each edited video averaged 7 minutes and 45 seconds in length (range: 4.6 min - 9.6 min) for a total of 46 minutes. A brief description of the salient features of each video is provided in Table 1 (not provided to the participants). The videos were stored on a laptop computer, allowing data collection to take place at a location convenient for each participant.
<table>
<thead>
<tr>
<th>Case</th>
<th>Duration</th>
<th>Important features</th>
</tr>
</thead>
</table>
| 1    | 8 min 3 sec   | • The cystic duct is divided before the cystic artery  
                     • A clip is placed proximally on the cystic duct and then removed         |
| 2    | 8 min 52 sec  | • The gallbladder is punctured early in the operation  
                     • A blood vessel is injured, leading to significant bleeding              |
| 3    | 4 min 36 sec  | • Quick operation (less than 20 min from beginning to end)  
                     • Only one structure is isolated, clipped and divided                   |
| 4    | 7 min 43 sec  | • Blood, bile, and severe inflammation make the anatomy difficult to interpret  
                     • The operating surgeon spends time dissecting near a large tubular structure that could be the common bile duct  
                     • A large blood vessel is injured, leading to severe bleeding         |
| 5    | 7 min 35 sec  | • A small cystic artery is divided before the cystic duct has been isolated      
                     • The cystic duct is clipped above what appears to be a gallstone; the operating surgeon then places a second set of clips below the presumed gallstone |
| 6    | 9 min 38 sec  | • The cystic artery is divided before the cystic duct has been isolated          
                     • A fast retrograde dissection causes bleeding and spillage of stones, and forces a subtotal cholecystectomy  
                     • An unusually large cystic duct is clipped and divided                  |

**Table 1:** Brief description of the 6 edited laparoscopic cholecystectomy videos presented to a sample of expert surgeons.
3.6 Think Aloud Training and Instructions

To familiarise participants with the think aloud method, a training script was prepared (see Appendix A). First, an explanation of the think aloud method was provided. Then, a series of simple practice problems were presented (e.g., what is the sixth letter after B?), providing feedback to the participants after each one. Training continued until it seemed participants were comfortable with thinking aloud and understood what was expected of them. The training script was adapted from Ericsson and Simon (1993).

For the main task, the following think aloud instructions were prepared:

I am now going to show you 6 video clips of laparoscopic cholecystectomies. I would like you to imagine that you are observing each procedure in the operating room but unscrubbed. I would like you to think aloud while you watch these procedures. The idea is the same as with the practice problems. You do not have to describe or explain anything to me. Just say whatever you happen to be thinking. If you are in the middle of a thought and notice something else, do not worry about completing sentences and move on to the next thought. We want the most accurate record of what is in your mind at the time. Remember to keep thinking aloud as best as you can throughout. If you remain silent for a while, I will remind you to keep talking.

These instructions were developed through pilot testing conducted with members of the research team and a surgical fellow. The open-ended nature of the instructions gave participants the freedom to focus their attention on whichever cues interested them and therefore required an ability to recognise critical cues without any prompting on the researcher’s part.
It was felt that asking participants to imagine themselves in the role of a third-party observer (not actually operating) would be best for capturing their responses to emergent perceptual cues. An alternative would have been asking participants to imagine themselves in the role of the supervising surgeon observing as their trainee performed the operation. However, it was thought this might encourage participants to focus on the technical aspect of the performance and how it could have been improved at the expense of other cues (e.g., anatomical). Similarly, it was thought that asking participants to role-play and imagine *themselves* performing the operation would prompt a discussion of the operator’s performance and how they (the participants) might do things differently. These decisions about the specific nature of the think aloud instructions were important as a result of the use of standardised cases, which placed participants in the passive role of an observer.

### 3.7 Data Collection

Data collection was performed between August and November 2012. All sessions were conducted by the primary researcher in a quiet environment at the workplace of the participants. At the start of each session, participants were reminded of the purpose of the research and informed consent was obtained. Following 5 minutes of think aloud training, participants were read the think aloud instructions shown above. Participants were then shown the 6 edited LC videos in a fixed order to eliminate case order as a confounding variable. To ensure consistency in the material viewed by each participant, participants were unable to control playback of the videos (i.e., pause, rewind or fast-forward); this made it possible to compare participants’ responses to a particular cue or event. Notes were taken when participants pointed to specific
cues on the screen while thinking aloud; this information was used later to disambiguate vague statements (“I’m not sure what that structure is back there”) when reviewing the data.

To help ensure verbal reports that reflect the participant’s actual inner thoughts, the primary researcher sat next to the participant rather than across from her or him whenever possible to minimise socialisation. He avoided conversation with the participants while the videos were playing, aside from providing the occasional reminder to keep thinking aloud if the participants remained silent for more than a few seconds.

Despite instructions not to teach or commentate, it was noted early on that participants would sometimes engage in these activities (e.g., defining medical terms, discussing general rules of thumb for the procedure) instead of thinking aloud. This was not surprising as the participants work in teaching hospitals and are therefore used to teaching while operating. However, participants deviating from the think aloud instructions could result in lost opportunities to capture valuable data, in particular during the critical events of the case. To mitigate this, the instructions were modified to more clearly explain how thinking aloud differs from teaching.

As much as possible, the cases were presented to participants without interruption. Where this was not possible (e.g., a participant answering a page), the case was paused and the interruption was kept as brief as possible; this occurred three times over the course of the study.

Demographic data were collected from the participants at the end of each session.

Each session was audiotaped and transcribed, with one exception: one participant recording for Case 4 was not transcribed (and was therefore excluded from the analysis) because the participant was familiar with the case. All transcription was performed by trained transcribers. The transcription was performed verbatim, including nonverbal sounds such as ‘umms’ and
‘aahs’ as well as laughter, pauses in speech, etc.; this additional information sometimes helped to interpret the meaning of a statement. A total of 125 pages of data were generated.

3.8 Data Analysis

The transcripts were analysed using the method of qualitative content analysis. Qualitative content analysis involves the systematic identification (‘coding’) of conceptual categories in textual data (e.g., interview transcripts, observation notes, books) for the purpose of understanding a particular phenomenon (in the case of this study, the ‘slowing down’ transition as manifested in a think aloud transcript) (Hsieh & Shannon, 2005). The coding process can be either inductive or deductive (Elo & Kyngäs, 2008). In an inductive approach, the coding categories are not determined \textit{a priori} but rather are allowed to emerge from the data. In a deductive approach, the coding categories are created \textit{a priori} based on existing literature. Because it was unclear how ‘slowing down’ moments would manifest themselves in the context of watching videos of surgical procedures, an inductive approach to coding was initially adopted. To integrate findings into the pre-existing ‘slowing down’ framework described in Moulton’s (2010) work, a more deductive approach to the data was used in the later stages of analysis. In this way, it can be said that data analysis was both inductive and deductive. Throughout the entire analysis process, the ‘slowing down’ framework and the situation awareness literature acted as sensitising concepts.

All transcripts were read and analysed by the primary researcher and a senior general surgical resident. Having surgical investigators involved in the analysis was essential for understanding technical terms as well as surgical jargon (e.g., “the Bulgarian technique”). Each transcript was
analysed while watching the videos and listening to the participant audio. This provided context which facilitated interpretation of the transcript.

Before the coding began, several transcripts were read to get a general impression of the data. Then, each transcript was read and statements representing ‘slowing down’ moments were highlighted. Each statement had to suggest an increase in the participant’s level of cognitive effort (e.g., looking for phrases such as “I would be careful here” or “we’re in trouble”). Volume and tone of voice on the audiotape were used to help with interpretation when the meaning of certain statements was unclear. A search was also made for statements that resembled Moulton’s descriptions of ‘slowing down’ moments in previous work (Moulton et al., 2010a, 2010b, 2010c; Moulton et al., 2007). Through a process of constant comparison between new examples of ‘slowing down’ moments and previous examples, ‘slowing down’ moments sharing similar characteristics were grouped into categories. When a new example did not fit into one of the existing categories, a new category was created. Emergent findings were discussed at regular meetings with the project supervisor. Transcripts were re-coded based on these discussions as the coding structure evolved. NVivo Version 8.0 (QSR International, Doncaster, Victoria, Australia) was used to facilitate management of the data set.
Chapter 4
Results

The Results chapter is presented in five sections. First, demographic data for the participants are provided. Second, the qualitative categories of ‘slowing down’ moments that emerged from the data set are described. Third, the challenges in isolating ‘slowing down’ moments using this methodology are discussed and findings that challenge the notion of discrete ‘slowing down’ moments are presented. Fourth, the various other categories of statements that were generated using the think aloud method are outlined. In the final section, a comparison of ‘slowing down’ moments between individual surgeons is presented.

4.1 Demographics

Demographic statistics for the participants can be found in Table 2. Nine of 10 participants had completed their general surgical training in Toronto.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of independent practice</td>
<td>Range = 2-25</td>
</tr>
<tr>
<td></td>
<td>Mean = 9.8</td>
</tr>
<tr>
<td></td>
<td>SD = 8.2</td>
</tr>
<tr>
<td>Current setting of practice</td>
<td>Academic (n=4)</td>
</tr>
<tr>
<td></td>
<td>Community (n=6)</td>
</tr>
<tr>
<td>Estimated number of LCs performed in average month</td>
<td>Range = 2-18</td>
</tr>
<tr>
<td></td>
<td>Mean = 8.1</td>
</tr>
<tr>
<td></td>
<td>SD = 4.45</td>
</tr>
<tr>
<td>Estimated total number of LCs performed in career to date</td>
<td>Range = 400-12,000</td>
</tr>
<tr>
<td></td>
<td>Mean = 2,093</td>
</tr>
<tr>
<td></td>
<td>SD = 3,529</td>
</tr>
<tr>
<td>Sex</td>
<td>Male (n=9)</td>
</tr>
<tr>
<td></td>
<td>Female (n=1)</td>
</tr>
</tbody>
</table>

*Table 2*: Demographic statistics for the 10 surgeon participants. (SD = standard deviation)
4.2 Categories of ‘Slowing Down’ Moments Captured in a Think Aloud Transcript

A large number of ‘slowing down’ statements were identified in the data set. In this section, these statements are described in the context of Moulton et al.’s (2010c) categorical framework of ‘slowing down’ moments (Figure 2, page 9). However, it should first be noted that, due to the continuous verbalisation of thoughts, individual statements were often multifaceted. For example, consider the following statement which was made as the operating surgeon was placing clips on the cystic duct:

“I personally don’t like it because he didn’t define the anatomy well enough and I would have taken the artery first and that would have opened it up more completely.” (Case 1, Surgeon 2)

This single sentence contains elements of concern about the operator’s performance (clipping was inappropriate at that time), anatomical uncertainties (the participant could not be sure that what was being clipped was in fact the cystic duct), diagnosis of the problem (inadequate dissection of Calot’s triangle) and decision making (the participant would have clipped the cystic artery first, not the duct). Therefore, it is acknowledged that some of the quotes presented in this section (and in this Results chapter in general) could be interpreted as belonging to multiple categories.

4.2.1 Proactively planned ‘slowing down’ moments

There were few explicit examples of procedural-specific ‘slowing down’ moments in the transcripts:
“This is the Calot’s triangle where, you know, some of the vital structures are and you want to take your time here and try to really sort of identify the artery and the duct itself.”

(Case 1, Surgeon 8)

This is not surprising considering that participants were asked to think aloud in response to particular cases rather than describe their general approach to the procedure. However, it could be implied from participants’ responses that one important procedural-specific moment was the clipping and cutting of the cystic structures. During this step, all participants sought out cues which would increase confidence that the isolated structures were in fact the cystic artery and cystic duct:

“There’s not much else that could be aside from the cystic duct given its size. And I could see the insertion in the common bile duct. And the aah, tissue above the insertion site seems to be going into the liver and the dissection has been free.” (Case 3, Surgeon 1)

Patient-specific moments were not observed because participants were not provided with patient information (e.g., clinical presentation, imaging reports) before watching the videos.

4.2.2 Situationally responsive ‘slowing down’ moments

Using the think aloud method, qualitative differences were captured among the ‘slowing down’ moments that were elicited in response to emergent intraoperative cues (i.e., the situationally responsive ‘slowing down’ moments). Four categories of these moments were identified in the data set and are presented below.
4.2.2.1 “Something bad is going to happen”: Concern in anticipation of a mishap

Most ‘slowing down’ moments in the data set were found to be in response to the operating surgeon’s performance. These moments involved concern in anticipation of a bad outcome (ranging from a minor intraoperative mishap to a serious postoperative complication). Some of these concerns were focused on a specific action:

“…aah, aah, I think he is right around the artery there, he is going to have some bleeding there if he tries to cut through that.” (Case 1, Surgeon 2)

In other instances, the participant seemed to be concerned not about a specific action but rather the operator’s general approach to the procedure. For example, during a case involving a patient with severe inflammation, one participant was uncomfortable with the operator’s decision to proceed with the laparoscopic approach:

“I have no idea what he’s doing but I wouldn’t be doing this… This operation needs to stop or something bad is going to happen.” (Case 4, Surgeon 5)

4.2.2.2 “Now we’re in trouble”: Concern in reaction to a mishap

Other ‘slowing down’ moments were initiated by an intraoperative injury that was caused by the surgeon and now required immediate attention:

“Whoa, there we go, now we’re into the hepatic artery. Nice job. Now you’ve got big trouble… you’re not going to stop that. You, you need a laparotomy kit, you need some help and get some blood. That’s going to bleed a lot aah, until you get that under control, so try and keep some pressure on it.” (Case 4, Surgeon 5)
Considerable bleeding was a common initiator of these ‘slowing down’ moments, as was suspected injury to the biliary tree. For example, during one case, several participants were concerned that the large tubular structure that had just been divided was something other than the cystic duct:

“Aah, really I think what you ought to do is a cholangiogram to that structure to try and figure out what it is… Like that may be the hepatic duct. (Short pause.) Yeah, so that’s much too large umm, a structure to be confident about… I don’t think the surgeon can know what has been done here.” (Case 6, Surgeon 6)

These ‘slowing down’ moments were distinct from those in the first category since the participant’s concern was now focused on a critical situation that resulted from the operator’s performance rather than on the operator’s performance itself.

4.2.2.3  “Where are we?”: Anatomical uncertainties

Participants sometimes expressed a desire to ‘slow down’ and exercise caution in response to uncertainty about the anatomy. Levels of uncertainty varied greatly between examples. Sometimes, participants seemed somewhat confident about the anatomy, but would nonetheless take care to confirm their suspicions before taking definitive action. For example, as the operating surgeon was isolating a band of tissue, one participant said:

“…he is around something. Is that an adhesion or is that the cystic duct? Probably just an adhesion but, boy, you’d want to be sure about that first.” (Case 2, Surgeon 2)
At the other extreme, some ‘slowing down’ moments involved a global lack of understanding of the anatomy. Participants struggled to orient themselves, and this struggle was sometimes associated with a sense of unease:

“This is definitely… I’m already thinking that this is not your classic anatomy here and we definitely have to proceed with caution. Something looks funny with that triangle.”

(Case 2, Surgeon 3)

Participants responded to these moments of uncertainty in different ways. In some cases, participants suggested a search for further information to improve their understanding of the situation (“you may want to do a cholangiogram to figure out what that structure is”). In other examples, the uncertainty prompted the participants to deviate from the original planned approach and initiate a more conservative course of action:

“I don’t know where the bile duct is but it could be right there… I’m not sure. Aah, I mean, at this point I would not be pushing the dissection in that location. I don’t know where the bile duct is but it looks to me like he’s too low down… I would get away from there and it’s time either to open and take the gallbladder down from the top down or just to put a tube in the gallbladder and get out of there.” (Case 4, Surgeon 2)

4.2.2.4 “Keep in mind where the bile duct is”: Proactive hypervigilance

Finally, some ‘slowing down’ moments were not in response to an imminent danger. Instead, they involved the recognition of cues that should be considered in order to avoid running into trouble later on. These ‘slowing down’ moments were sometimes initiated by a very specific cue.
For example, as the operating surgeon was isolating what seemed to be a small cystic artery, a participant commented:

“…a small anterior artery suggests that there could be a bigger posterior one… And here I would… mentally note that there could be a big posterior branch and to look out for that as I clear the rest of the triangle.”

In contrast, other ‘slowing down’ moments were based on a more general impression of the situation. One participant took a brief moment at the start of a case to assess the patient:

“Umm, the gallbladder looks very thick, it could be a very inflamed gallbladder. The alternative, it could be a cancer of the gallbladder umm, I’m not sure what preoperative imaging there was. There is some pus leaking out from the top suggesting this is an obstructed gallbladder or empyema, so severe cholecystitis. So this is going to be a very difficult gallbladder umm, I know that.” (Case 4, Surgeon 6)

Based on the appearance of the gallbladder, this participant expected a more challenging operation that would require increased vigilance to safely complete. The intraoperative finding of severe inflammation in a patient would prompt a surgeon to exercise greater caution when performing the dissection of Calot’s triangle because inflammation increases the risk of anatomic misinterpretation (Strasberg, 2002).

4.3 The ‘Slowing Down’ Transition: Discrete Events and a Continuous Process?

The transcripts were complex, with thoughts running into one another because the think aloud method required near-continuous talking. This made the extraction of ‘slowing down’ moments
from the data set much more difficult than first anticipated. There were certain ‘slowing down’
moments that stood out as critical events:

“No, no, no, no, no, he’s around the bile duct, coming around the bile duct here, that’s
going to get really bloody in a second…” (Case 4, Surgeon 5)

However, other ‘slowing down’ moments were more subtle and were found as part of the
participant’s normal, continuous flow of thoughts:

“Okay, that’s nice. That was a nice move. Yeah, I, I just find the anatomy is not well
defined and I’d be going really slow here partly because, aah, Hartmann’s looks like
it’s folded in there, that’s nice. You want to fold the gallbladder out of there to be able
to really see it. That was a nice thing. And actually now, now I can see that looks like
the cystic duct down below.” (Case 2, Surgeon 2)

This excerpt serves as a good example of the challenge that was faced in sorting out the ‘slowing
down’ moments from the surrounding statements. There was no access to a quantitative measure
of the magnitude of cognitive effort associated with each statement (a limitation of the think
aloud method considering the goal was to capture ‘slowing down’ moments). Therefore, other
criteria (such as the participant’s choice of words [e.g., “careful,” “watch out,” “uh oh,”
expletives], volume and tone of voice) were relied on to select the statements that were thought
to represent ‘slowing down’ moments.

However, in between the ‘slowing down’ moments, it seemed that participants spent a great deal
of their time engaging in an ongoing process of monitoring the situation, which resembled the
“heightened state of surveillance” described in Moulton’s previous work (Moulton et al., 2010b).
This process involved a continuous monitoring of the operative field and the progress of the
operation. Consistent with the SA literature, information was gathered through both top-down and bottom-up processing. For instance, participants sometimes looked for specific cues to minimise the risk of a complication (top-down processing):

“H’mm, I think, you know, before I started digging here I’d like to see the gallbladder wall which is probably a few centimetres higher…” (Case 1, Surgeon 4)

Meanwhile, other information arose in the form of unexpected findings (bottom-up processing):

“Is that a stone? Could there be a st…? He’s cut across a stone, it’s hard to see…” (Case 1, Surgeon 5)

As part of the monitoring process, participants made continuous judgments about the progress of the operation (“the surgeon does seem to be in good control of this operation”) and anticipated future events (“I know it’s going to be a long case… this patient may have a longer stay in the hospital because of this”). They also identified particular aspects of the procedure that were suboptimal and required action. The initiating cues were sometimes anatomic (e.g., omentum obscuring the operative field) and sometimes technical (e.g., the choice of instrument). In response to these cues, participants then recommended a specific technical adjustment for the purpose of keeping the operation safe and on track:

“So in this case when you’re… when the gallbladder is falling apart, go up higher where it’s a little healthier and stop grabbing and just retract it by, aah, without grasping.” (Case 2, Surgeon 5)
From a qualitative perspective, monitoring statements were not that different from the ‘slowing down’ statements. Both monitoring statements and ‘slowing down’ statements captured the perception and interpretation of emergent cues, such as anatomic structures and technical manoeuvres. The difference between monitoring and ‘slowing down’ statements seemed to be more a matter of the magnitude of cognitive effort that was engaged at the time. For example, consider the following two statements:

“He’s holding [the cystic duct] quite tightly now and I’m just worried that he is going to avulse it potentially.” (Case 5, Surgeon 7)

“They should get a better, aah, left hand (short pause) retraction.” (Case 6, Surgeon 10)

Both statements are similar in that the participants are noticing that the retraction is inadequate, but in the first statement there is also an element of concern as Participant 7 is worried that the cystic duct might tear due to excessive tension. For this reason, the first statement was coded as a ‘slowing down’ moment while the second statement was coded as a monitoring statement. However, it is reasonable to consider the second statement as also involving an increase in cognitive effort, albeit a minor one that did not seem salient enough to represent a ‘slowing down’ moment as the transcripts were being coded. It therefore appeared that the continuous monitoring activities that were captured in the transcripts might consist of miniature ‘slowing down’ moments that occur countless times over the course of a surgical procedure.
4.4 Other Categories of Think Aloud Statements

Along with the ‘slowing down’ moments and the ongoing process of situation assessment, various other categories of statements were generated as part of the think aloud task. These categories consisted of decision making statements, statements specific to the participant’s role of observer, and extraneous statements (noise in the data), and are detailed below.

4.4.1 Decision making

Emergent cues sometimes prompted participants to propose a specific course of action. In some instances, participants demonstrated forward planning, anticipating a future step for the procedure:

“Oh, they made a hole [in the gallbladder]... So at this point I would... tell my scrub nurse to get a bag because you’re going to need to put the gallbladder in a bag at the end there.” (Case 2, Surgeon 8)

Other examples of decision making consisted of immediate changes in response to emergent information. For example, the conventional approach to LC involves securing the cystic duct with clips. However, in response what seemed to be a wide cystic duct, one participant proposed an alternative method of cystic duct closure:

“I’m noticing that the cystic duct is somewhat aah, wide... at this point I think I would take this entire gallbladder off again, and aah, assess where, where I thought the cystic duct entered the common bile duct and would probably put an Endoloop around it rather than trying to clip that wide thing.” (Case 5, Surgeon 1)
4.4.2 Statements specific to being an observer

Certain statements that were generated as part of the think aloud task were related to the participant’s role as an observer. In other words, a participant would not be having these thoughts if she was performing the operation herself. These statements could be placed into three subcategories.

Interpreting the operator’s actions or rationale

Because the participants had not performed the operations themselves, they sometimes had some trouble understanding what the operator was doing at a particular moment. However, the confusion was often quickly resolved:

“Okay, I think they’re… what are they trying to do? What are you doing? Okay, taking it off the liver.” (Case 6, Surgeon 9)

There were also moments when the operator’s course of action was clear to the participants but the rationale for doing it was not. This sometimes prompted participants to start guessing what the operator might have been thinking at the time of the procedure.

“They’re trying to remove a clip. I guess they’re concerned that the, aah, clip was improperly placed or didn’t go all the way across or was too close to, aah, to the bile duct.” (Case 1, Surgeon 6)
Alternative course of action

Statements in this category formed a subset of the decision making statements presented above. They consisted of moments when participants stated that they would not be doing what the operator was doing, proposing instead an alternative course of action. Examples of alternative courses of action appeared to fall along a ‘spectrum of disagreement.’ At one end of this spectrum, participants were not objecting to the operator’s actions but just stating their own personal preferences. For example, as the structures of Calot’s triangle were being exposed, one participant said:

“Umm, so, umm, you know poking away with that is a reasonable way to do it. I would probably do it with a different instrument, I’d just put a grasping instrument and spread and try and very quickly identify those two structures.” (Case 1, Surgeon 6)

In contrast, other alternative courses of action were offered in response to critical concern about patient outcome and involved significant changes to the operative approach:

“…at this point [I] wonder why we’re not trying to go top down or convert to open, to have a feel of those structures… it’s not clear what they’re actually dividing.” (Case 4, Surgeon 10)

Performance evaluation

Since participants were observing the performances of other surgeons, judgments were sometimes made about a surgeon’s abilities or level of experience. Some of these statements had negative connotations:
“I guess this is where they got the medical student to take over and, aah, got into the gallbladder.” (Case 3, Surgeon 3)

In other examples, participants seemed quite pleased with the surgeon’s performance:

“That’s a good move with the left hand doing a bit of retraction. I wonder if this is a staff sort of taking over now …” (Case 2, Surgeon 3)

4.4.3 Extraneous statements

In addition to the above categories of statements, extraneous data were sometimes observed in the form of teaching or commentating. When this took place, participants were not speaking to themselves but for an external audience (i.e., the researcher). These statements were therefore inconsistent with the instructions to ‘think aloud.’

Teaching

Teaching statements were when participants provided general information about the LC procedure that was not specific to the case being presented. The content of these statements included definitions of medical terms and general guidelines for the procedure:

“And I, I find that a lot of incidents will happen when, aah, you’re, you are trying to fiddle with your right hand forgetting that your left hand is clamped on to something. And then you end up ripping it and moving the target about. You have to always remember that you’re, that you have got two hands in the patient and that you have your instrument on something.” (Case 3, Surgeon 3)
Commentating

Commentating consisted of editorial remarks, anecdotes, and light humour. Similar to teaching, these statements contained superfluous information that was unrelated to the progress of the case.

“This one should be an absolute chip shot and it should be done before the coffee gets cold.” (Case 3, Surgeon 1)

4.5 Comparison of ‘Slowing Down’ Moments between Individuals

The goal at the onset of this study was to compare ‘slowing down’ moments that emerged from the transcripts between individuals. However, over time it became obvious that this would be more challenging than anticipated, due in part to the complex and complicated nature of the data which made it difficult to categorise the ‘slowing down’ moments. It was therefore decided to focus the quantitative component of this study on a subset of ‘slowing down’ moments during which participants expressed concern about the operator’s performance. These moments (which were termed ‘moments of concern’) were chosen 1) because they would be relatively easy to identify in the transcripts, and 2) because it was believed these were the ‘slowing down’ moments that were most likely to be significant and consistent between participants. In this section, the guidelines for analysis are described in detail, followed by an in-depth presentation of the resulting data.
4.5.1 Guidelines for analysis

4.5.1.1 Operational definition

A moment of concern (hereafter referred to as MC) was defined as concern or displeasure expressed in response to the operator’s actions. To be included, the action had to be specific:

“Oh, my goodness, here we are going to get rid of that artery right away… I wouldn’t have necessarily recommended that… there’s been no attempts to aah, find a cr…, a critical view of safety.” (Case 5, Surgeon 1)

Focusing on specific actions made comparisons between participants easier. Concerns unrelated to a specific action (e.g., general concerns about the progress of the case) were excluded:

“I really don’t like the way this is going. Umm, you’ve got a hole up top with blood coming. (Short pause.) Umm. (Short pause.) This is getting kind of scary.” (Case 2, Surgeon 5)

The action also had to involve a direct risk of causing a complication or getting into trouble. Criticisms about poor retraction or camera angle, for example, were therefore excluded.

4.5.1.2 Timing of MCs

Because participants were asked to think aloud throughout each case, how to deal with the timing of MCs was an important question. For example, if surgeon A expressed a concern 3 minutes and 5 seconds into a case and surgeon B expressed the same concern 3 minutes and 15 seconds into the same case, were these MCs equivalent? Or was the 10-second difference important? It was decided that all expressions of the same MC would be considered to be equivalent if expressed while the action in question was still reversible or avoidable. Concerns expressed after the fact
(e.g., “I think I saw it before he did that, I wish… that wouldn’t have happened” – Case 4, Surgeon 1) were not coded as MCs.

4.5.1.3 Redundant MCs

Sometimes, participants expressed the same concern multiple times in the same case. For example, during one case a participant was concerned about the order in which structures were being divided:

“…I might not do that. My standard technique would be to usually deal with the artery first… I usually teach and recommend that the aah, trainees and residents aah, divide the aah, artery first before the cystic duct to avoid avulsion or problems with the artery later.” (Case 1, Surgeon 1)

Later, as the cystic artery was being divided, the same participant stated:

“So, now I don’t like this aah, tenuous artery… I’d be a little nervous of this technique and all the traction…” (Case 1, Surgeon 1)

The second concern could be traced back to the same action that prompted the first concern: the division of the cystic duct before the cystic artery, which had made the cystic artery more susceptible to avulsion injury. For this reason, the two concerns were coded as the same MC rather than two separate MCs. The same was done for other ‘redundant’ MCs. The exception was when a previous concern was repeated (e.g., dissecting too close to the portal structures) but in response to a new emergent event unrelated to the original one; in this case, the new concern was coded as a new MC.
4.5.1.4 Grouping of MCs

It was sometimes clear that several participants were concerned about the same action in a case, but the reasons provided for their concerns were different. Concerns that shared a common explanation were grouped together as one MC. For instance, during one case participants expressed the following concerns as the cystic duct was being divided:

“…I’m not sure what that structure is back there. I don’t know if they’re way too low or not.” (Case 1, Surgeon 4)

“...it’s not really adequate enough dissection in my mind. The triangle of Calot has not been opened, umm, so there is a possibility that this is the bile duct. It’s fairly big.” (Case 1, Surgeon 5)

“I’m thinking, umm, a critical view should be obtained before placing any clips.” (Case 1, Surgeon 10)

While the details of these statements varied, the common overarching concern was that the anatomy had not been defined well enough to be sure that it was safe to divide this structure. These statements were therefore coded as the same MC.

Sometimes, however, the reasons for concern were too different from a clinical perspective to group them together. To continue with the above example, all participants agreed that the division of the cystic duct was premature and that further dissection would have been beneficial. The two main reasons provided were 1) that the cystic duct had not been cleared of fat and 2) that the anatomy was unclear (as described above). The clinical implications of these concerns were different. Fat left on the cystic duct can prevent the clips from completely occluding the duct, which can in turn result in a postoperative bile leak. In contrast, poor definition of the
anatomy can lead to inadvertent injury to the common bile duct, which is a serious complication. It was therefore decided to maintain the distinction between the two reasons for concern. Participants who expressed concerns about both the excess fat on the cystic duct and anatomical uncertainty were considered to have expressed two MCs in response to the same action (the division of the cystic duct).

4.5.1.5 Coding strategies
As mentioned earlier, the audio and video were used for coding the transcripts. The participant’s tone of voice was useful for determining whether there was actual concern (on their own, the transcripts were sometimes misleading). Any doubts or disagreements about whether a statement qualified as an MC were resolved through discussion with the project supervisor, who had extensive experience with the LC procedure. These doubts were often centred on either 1) the overall meaning of the statement or 2) whether the statement described real concern about the action or whether the participant was just pointing out a minor technical issue.

4.5.1.6 Reporting of MCs
Descriptive statistics were used to tabulate the MCs. These results were then further examined from a qualitative perspective to gain further insight from the data.

4.5.2 Moments of Concern: Descriptive Statistics
4.5.2.1 Overview
A total of 54 distinct moments of concern were identified in the transcripts (Figures 5-10). Each participant expressed between 16 and 30 MCs (mean = 21.3; standard deviation = 5.0). The
number of participants who expressed each MC was variable. As Table 3 demonstrates, most MCs (74%) prompted at least 2 participants to express concern, but few MCs were consistent across the sample. The high proportion of inconsistent MCs was unexpected considering the participants had been selected from a ‘homogeneous’ sample population of expert surgeons. The data were further examined from a qualitative perspective to uncover some of the reasons for these results.

<table>
<thead>
<tr>
<th>Number of participants who expressed MC</th>
<th>Number of MCs (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 or more</td>
<td>13/54 (24%)</td>
</tr>
<tr>
<td>5 or more</td>
<td>22/54 (41%)</td>
</tr>
<tr>
<td>2 or more</td>
<td>40/54 (74%)</td>
</tr>
<tr>
<td>Only 1</td>
<td>14/54 (26%)</td>
</tr>
</tbody>
</table>

Table 3: Level of consistency between participants in the expression of individual moments of concern.

---

1 The term ‘consistent’ is used here to describe MCs that were verbalised by most participants. The term ‘inconsistent’ is used to describe MCs that were verbalised by few participants.
**Figure 5:** Moments of concern (MCs) expressed by participants during Case 1. Each bar along the timeline represents a distinct MC. MCs labeled with an asterisk coincided with one another.

- **Cause of concern**
  
- a = Dissecting too low (could injure portal structures)
- b = Could injure the cystic artery with the cautery tool
- c = Grabbing too much tissue with the cautery tool (could injure an unseen structure)
- d = Order of clipping (would have clipped the cystic artery first)
- e = Too early to clip the cystic duct (too much fat left on the duct)
- f = Too early to clip the cystic duct (anatomy unclear)
- g = Pulling off a clip could injure the cystic duct
- h = Leaving stones in the cystic duct could result in a post-operative complication
Figure 6: Moments of concern (MCs) expressed by participants during Case 2. Each bar along the timeline represents a distinct MC. MCs labeled with an asterisk coincided with one another.

Cause of concern

a = Dissecting too low (could injure portal structures)
b = Burning too close to a potential structure
c = Dissecting too fast/operator being too cavalier
d = Wouldn’t use cautery to control bleeding
e = Blind clipping of a bleeding vessel
f = Burning through what could be a structure
g = Too early to clip the cystic duct (anatomy unclear)
h = Clipping the cystic duct too low (could injure the common bile duct)
i = Applying too much pressure on a bleeding vessel (could tear it out of the liver bed)
j = Blind clipping of a bleeding vessel
k = Blind clipping of a bleeding vessel
l = Could injure the liver with the graspers
Figure 7: Moments of concern (MCs) expressed by participants during Case 3. Each bar along the timeline represents a distinct MC.

**Cause of concern**

- a = Dissecting too far from the gallbladder wall
- b = Grabbing too much tissue with the cautery tool (could injure an unseen structure)
- c = Poking through with the L hook (could injure an unseen structure)
- d = Dissecting too far from the gallbladder wall
- e = Dividing the cystic duct before the cystic artery has been identified
**Figure 8:** Moments of concern (MCs) expressed by participants during Case 4. Each bar along the timeline represents a distinct MC. Note that only 9 participants were included in the analysis for this case.

![Bar chart showing moments of concern](image)

**Cause of concern**

- **a** = Uncomfortable with the area of dissection (could cause a thermal injury)
- **b** = Burning through what could be a structure
- **c** = Uncomfortable with the area of dissection (could cause a thermal injury)
- **d** = Uncomfortable with the area of dissection (could cause a thermal injury)
- **e** = Uncomfortable with the area of dissection (could cause a thermal injury)
- **f** = Uncomfortable with the area of dissection (could cause an injury with the graspers)
- **g** = Uncomfortable with the area of dissection (could cause a thermal injury)
- **h** = Uncomfortable with the area of dissection (could cause a thermal injury)
- **i** = Uncomfortable with the area of dissection (could cause a thermal injury)
**Figure 9:** Moments of concern (MCs) expressed by participants during Case 5. Each bar along the timeline represents a distinct MC. MCs labeled with the same symbol coincided with one another.

**Cause of concern**

a = Dangerous dissection technique (could injure the duodenum)
b = Burning too close to the cystic artery
c = Burning too close to the cystic artery
d = Too early to clip the cystic artery (the cystic duct has not yet been isolated)
e = Poor use of graspers to spread tissue (could injure unseen structure)
f = Dissecting too high (could perforate the gallbladder or injure another structure)
g = Might be burning through a branch of the cystic artery
h = Inadequate dissection of the cystic duct / clips not going across
i = Poor clip placement on the cystic duct (leaving a stone behind)
j = Clipping the cystic duct and cystic artery together
k = Clipping too low (could injure the common bile duct)
l = Too much tension on the cystic duct (could avulse it)
Figure 10: Moments of concern (MCs) expressed by participants during Case 6. Each bar along the timeline represents a distinct MC.

Cause of concern

a = Dissecting too low (could injure portal structures)
b = Could injure the cystic artery with the cautery tool
c = Too early to clip the cystic artery (the cystic duct has not yet been isolated)
d = Cystic artery is divided with only one proximal clip placed across it
e = Hasty/cavalier dissection of the gallbladder
f = Wide cystic duct being clipped/anatomy unclear (could be the common bile duct)
g = Cystic duct is divided with only one proximal clip placed across it
h = Removing the clip that was placed on the cystic duct to replace it with a better one (challenging manoeuvre that could result in injury)
4.5.2.2 Differences in the nature of events

The most consistent MCs were often related to the clipping and cutting of the structures of the triangle of Calot. This is not surprising because this represents the most critical part of the procedure. As an example, during Case 1, the operating surgeon removed a clip that had been placed on the cystic duct. This action prompted 8 out of 10 participants to express concern (Table 4).

<table>
<thead>
<tr>
<th>Participant No.</th>
<th>Expression of concern</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>“…it looks like, they’re about to aah, tear off aah, a clip… And I wonder, aah, how or where this is going, whether this is going to create more harm than good.”</td>
</tr>
<tr>
<td>2</td>
<td>“And he’s, he is going to take off the clip, aah, it’s aah, so let’s see what happens… the problem is you can tear the duct. If he tears it then he is going to be really screwed because it’s, it’s quite low down.”</td>
</tr>
<tr>
<td>3</td>
<td>“I don’t know what the hell the operator is doing… I would proceed with clearing out the triangle more before trying to take off this clip because this clip might be totally fine and now you are, you’re probably causing more problems by taking off the clip… you could get a bile leak from there…”</td>
</tr>
<tr>
<td>5</td>
<td>“Umm, so, umm, I’m not sure that I would take the clip off. You can, you can actually injure the duct, put a hole in it aah, and then have trouble finding it.”</td>
</tr>
<tr>
<td>6</td>
<td>“Okay, so they’re removing the clip. One concern is, umm, was there any crush damage when the clip was applied…”</td>
</tr>
<tr>
<td>7</td>
<td>“To manipulate the clip is very challenging. And I would try not to, I would better just try to see where the cystic duct is coming from umm, because I wouldn’t want to injure the cystic duct either, unnecessarily.”</td>
</tr>
<tr>
<td>8</td>
<td>“Now they’re taking it out of there which is not ideal… you can damage the cystic duct.”</td>
</tr>
<tr>
<td>10</td>
<td>“I’m not loving the picking off the clips.”</td>
</tr>
</tbody>
</table>

Table 4: Participants’ expressions of concern in response to the operating surgeon’s decision to remove a clip that had been placed on the cystic duct in Case 1.
Other consistent MCs were related to major case-specific events that are not encountered during a routine LC. For instance, 8 out of 10 participants expressed concern in response to a careless retrograde dissection of the gallbladder (the operating surgeon made no apparent effort to remain in the proper plane of dissection), which resulted in bleeding and spillage of stones that complicated the case:

“So whoever’s operating now is being a little bit more cavalier now… you’re just making a mess. You’re prolonging the operation by not being careful… Now you’re tugging everything and there is going to be more bleeding.” (Case 6, Surgeon 3)

Among the MCs that were less consistent between participants, there were no obvious patterns. Some were linked to major events, such as clipping and dividing structures or dealing with significant bleeding. Others were related to poor technique during more mundane events. For example, as the operating surgeon was taking down adhesions at the start of Case 2, a single participant was worried about injuring the common bile duct:

“You can see the common bile duct down at the bottom of the screen… you don’t have to work so low down into this area with higher risk of getting into injuries… it’s very dangerous the way he’s aah, using his hook down there.” (Case 2, Surgeon 7)

Therefore, it did not appear that inconsistency in the expression of MCs was specific to a particular type of event.

4.5.2.3 Differences in situation awareness

Apparent individual differences in SA could explain some of the inconsistency that was observed in the expression of MCs. Sometimes, these differences were at the level of perception (Level 1
SA), with some participants noticing a particular cue that others did not. For instance, 8 out of 10 participants expressed reservations when the cystic duct was divided in Case 3 because the cystic artery had not yet been isolated:

“I’d never put a clip yet. I really want to see where my cystic aah, artery is… I’m not seeing any of that right yet umm, and umm, that would aah, that would make me much more comfortable in knowing what is what.” (Case 3, Surgeon 7)

One of the 2 participants who did not express concern was convinced that he had seen the cystic artery earlier (“right now without doing anything I can see the artery”). When he could no longer see it, he assumed that it had been cauterised during the dissection of the triangle of Calot:

“…it looks like the artery was cauterised through which was fine, it was tiny and, aah, and haemostasis was satisfactory… there is no blood anywhere.” (Case 3, Surgeon 3)

Of the 10 participants, he was the only one who claimed to have seen the cystic artery at any point during this case. In this example, the cue in question (the cystic artery) had a positive or reassuring effect for the participant who saw it. It increased his confidence that the structure being divided was in fact the cystic duct, which explains why he did not express this particular MC.

In other instances, participants noticed the same cue (i.e., same Level 1 SA) but differed in their interpretation of that cue (i.e., different Level 2 and 3 SA). As an example, as the operating surgeon divided the cystic duct in Case 1, a stone remnant could be seen sitting in the duct. One participant was concerned about this:
“I don’t like those rem…, those stone remnants sitting in the cystic duct because we’ve
seen cases of aah, of inflamed cystic duct, umm, later on after a, after
cholecystectomy…” (Case 1, Surgeon 1)

In contrast, 6 more participants mentioned the stone but did not express concern (“there are some
small stones that you can see in the cystic duct remnant…I don’t think they’re a big deal”). One
participant even found it reassuring:

“Oh, there is a little stone there in the, right where he cut across which is actually a
pretty, probably a good sign. It’s probably a sign that it actually was a cystic duct and not
the artery.” (Case 1, Surgeon 2)

In this particular example, the idiosyncrasy of the MC seemed to be more a matter of
interpretation than perception. In other words, most participants did notice the stone in the cystic
duct but did not appear to be bothered by it.

Even when participants agreed on an MC (i.e., expressed displeasure at the same moment), there
were apparent differences in SA. First, it seemed that some participants were more concerned
than others based on their tone of voice or the language used to describe their concern. For
instance, consider two participants’ responses to the same event (clipping of the cystic duct):

“I mean, aah, I would have… he hasn’t done any dissection really at all posteriorly. He
could have just freed it up completely to open up that triangle so, I mean it’s probably
fine but there’s no reason, ten seconds more could have just opened up that window a bit
better so that he could see the anatomy a little bit clearer.” (Case 3, Surgeon 2)
“No, no, no, no now he’s going to bring clips in, bad news. I wouldn’t put clips on until I found the artery or made sure I’d opened Calot’s triangle a little better.” (Case 3, Surgeon 5)

Both participants felt that there had been inadequate dissection prior to placing the clips on the duct. However, Participant 3 conceded that it should still be safe to clip the duct while Participant 5 seemed more confident that it was a dangerous move, indicating slight differences in their interpretation of the situation (Level 2 SA). It also appeared that participants could arrive at the same MC while focusing on different perceptual cues (Level 1 SA):

“So the gallbladder, I can’t see the gallbladder wall so I’d go higher… I can’t tell if there’s a problem here but it’s a bit on the low side.” (Case 1, Surgeon 5)

“I’m a little worried, there is the porta… he is too far off the, he should be closer on the gallbladder there.”

Both participants were concerned that the operating surgeon might be working in a dangerous area with a risk of injuring the portal structures, but it seemed that one participant came to this conclusion using the distance from the gallbladder wall as a reference point while the other seemed to focus on the porta itself.

4.5.2.4 Differences between cases

Some cases produced more consistent MCs than others. The least consistent results were obtained in Cases 3 and 4. Case 3 was the simplest of the 6 cases, with the entire operation lasting 20 minutes (in contrast, the average operative time for LC is on the order of 1 hour [und Torney, Dell-Kuster, Mechera, Rosenthal, & Langer, 2012]). Of the 5 MCs that emerged during
this case, only 1 (20%) was consistent across participants (Figure 7). This occurred when the cystic duct was divided without clear visualisation of the cystic artery, prompting 8 out of 10 participants to express concern. The remaining 4 MCs were idiosyncratic, as each was expressed by no more than 1 or 2 participants. These reflected more subtle concerns (e.g., dissecting too far from the gallbladder wall while obtaining the critical view). The prevalence of minor, idiosyncratic MCs in this case suggests that perhaps it was too simple and uneventful, and that more challenging cases are required to elicit consistent MCs from surgeon participants.

In contrast, Case 4 was the most challenging of the 6 cases. The extent of disease (severe inflammation, blood and bile obscuring the operative field) made it challenging, but during this case some participants also expressed concern that the operating surgeon was being too bold or cavalier:

“H’mm, burning a big chunk of tissue (short pause) at the aah, (short pause) h’mm, courageous, very brave… Not, not clear what that was.” (Case 4, Surgeon 10)

“That’s aah (short pause) pretty ballsy. Someone actually videotaped this.” (Case 4, Surgeon 2)

None of the participants were comfortable observing this case. In fact, 7 of 9 participants believed that the operating surgeon was working against the common bile duct. However, only 2 of 9 MCs (22%) in this case were expressed by 5 or more participants (Figure 8). This was a surprising result given that participants’ reactions to this case were unanimously negative. It is possible that this case was too difficult. Participants had trouble interpreting the anatomy (“this looks awful… I see no recognisable anatomy yet we’ve got bile, we’ve got mucus, we’ve got blood”), which may have limited their ability to determine whether the operating surgeon’s
actions were safe. It also seemed that participants were distracted by the operator’s reluctance to convert to a more conservative approach (e.g., retrograde dissection, open cholecystectomy).

This prompted some participants to provide detailed accounts of what they would do differently:

“I mean, in this scenario… I would put a suction in and suck all this pus and other stuff out of here to minimise contamination. I would aah, take it off retrograde, leave the back wall on, and cut down and if you can, think you can dissect into Calot’s triangle go ahead. If not, then I would just open up the infundibulum and Hartmann’s and then maybe aah, take all the stones out and put them in a bag and then close the cystic duct from the inside with a stitch or something.” (Case 4, Surgeon 8)

It is possible that while describing their operative strategies, participants were distracted and failed to notice dangerous actions (or perhaps they did notice them but did not verbalise concern because they were in the middle of discussing something else). This suggests that the ability to think aloud may be impaired when the actions taken by the operating surgeon are radically different from the actions that the participants would take themselves.
Chapter 5
Discussion

In this chapter, the conceptual implications of the findings are discussed, relating them to the existing literature as applicable. The methodology used in this study is then evaluated as a potential tool for studying ‘slowing down’ moments, examining both its positive and negative aspects. Finally, a reflection on the research process is provided.

5.1 Conceptual Implications of Findings

The findings of this study add to our understanding of the ‘slowing down’ phenomenon in three ways. First, the existing framework of categories of ‘slowing down’ moments has been expanded. Second, the findings expand the current conceptualisation of ‘slowing down’ as more discrete and obvious events and provide insight into how miniature ‘slowing down’ moments are part of a continuous process of monitoring or vigilance. These findings inform the concept of fine-tuning described in the original ‘slowing down’ studies. Third, the finding that ‘slowing down’ moments are variable between experts contributes to our understanding of the nature of surgical expertise. Each of these main findings will now be discussed in more detail.

As part of the first research question, the ‘slowing down’ moments captured in the think aloud transcripts were qualitatively described. While it was already known that proactively planned ‘slowing down’ moments could be further categorised as procedural-specific and patient-specific, little was known about the qualitative properties of the situationally responsive ‘slowing down’ moments. In Moulton et al.’s (2010c) previous work, situationally responsive moments had been described as occurring in response to emergent intraoperative ‘surprises.’ Participants described pausing during these moments to “regroup” or “reassess” the situation (Moulton et al.,
Situationally responsive ‘slowing down’ moments were therefore understood to be critical events that should be obvious to an outside observer. These findings were supported by observational work which identified observable manifestations of ‘slowing down’ such as ‘stopping’ and ‘removing distractions’ (e.g., asking that the music be turned off) (Moulton et al., 2010b). However, using the think aloud method to capture the continuous flow of a surgeon’s thoughts in response to emergent intraoperative information, the present study found that situationally responsive ‘slowing down’ moments are more diverse in nature than previously thought. While some are in fact critical and obvious (such as immediate concern in response to a mishap), others are more subtle and emerge as part of the normal, ongoing flow of thoughts (such as uncertainty about the anatomy prompting further investigation). The rich qualitative data collected using the methodology explored in this study led to the identification of four distinct categories of situationally responsive moments, therefore making a contribution to the categorical framework of ‘slowing down’ moments (Figure 11).

Figure 11: Refinement of the framework of categories of ‘slowing down’ moments.
A major finding was the presence of background monitoring activities, which challenges our current conceptualisation of ‘slowing down’ moments as fairly discrete events. The monitoring statements resembled the ‘slowing down’ statements in content (both groups of statements involved the perception and interpretation of emergent cues) but seemed more minor in the amount of cognitive resources required at the time (i.e., the ‘slowing down’ moments seemed more critical or important). There is some literature which suggests that effortful thinking is not an all-or-none event, but rather that cognition occurs along a “continuum” with intuition at one extreme and effortful thinking at the other (i.e., cognition can be more effortful in some situations than in others) (Hammond, 1988). It could therefore be argued that these background monitoring activities can be thought of as miniature ‘slowing down’ moments. To continue with this idea, if a surgeon’s level of cognitive effort is fluctuating throughout the entire procedure as she attends to various cues, it might be more accurate to think of ‘slowing down’ as a continuous process that sometimes involves discrete events (Figure 12).

The background monitoring activities could in fact represent what has been described in previous work as the observable manifestation of ‘slowing down’ called fine-tuning (see Section 1.2.3.2). Due to the observational nature of the work, fine-tuning was described as constant technical adjustments in response to emergent cues to keep out of trouble. Based on the results of the present study, however, it is argued that fine-tuning is a continuous internal process. Moulton et al. (2010b) described the observable product of that process (technical adjustments) but, based on the methodology used in this study, fine-tuning also seems to involve fine adjustments from a cognitive perspective which might not lead to an observable behaviour.
Figure 12: Conceptual drawing showing continuous fluctuations in the surgeon’s cognitive effort over the course of a surgical procedure. Large increases represent discrete ‘slowing down’ moments. Minor fluctuations represent fine-tuning (monitoring) activities. Moderate increases in cognitive effort demonstrate the challenge in discriminating between ‘slowing down’ moments and fine-tuning.

For the second research question, a set of guidelines was developed for comparing ‘slowing down’ moments between individual participants, focusing on a specific subset of ‘slowing down’ moments that were termed ‘moments of concern’ (MCs). Few MCs appeared to be consistent across the sample of expert surgeons, which might be considered surprising coming from a sample population (experts) that was hypothesised to be homogeneous. Three main potential explanations may account for the poor consistency that was observed between experts. First, the limitations of the methodology can likely account for some of the inconsistency. De Groot (1965) suggests that verbal reports are reliable for what they contain but are nonetheless incomplete, which means that participants may have experienced certain concerns that were
never verbalised and captured as data. For example, perhaps a participant failed to mention the sudden recognition of an undesirable action because he did not think the action was critical enough to mention out loud. The failure of some participants to mention a concern during a particular event would in effect reduce the consistency observed between participants. Another important limitation of the methodology was the inability to quantitatively measure the relative importance of each MC. It should be safe to assume that participants’ levels of concern were higher during certain MCs than others. Perhaps participants in this study were inconsistent in identifying the least critical (perhaps even unimportant) MCs but at least consistent with respect to the most important ones. Future work would be required to verify this hypothesis and would require returning to the participants to ask them to rate their level of concern during each MC. Second, some of the variability may be the result of sampling error. Since there is no single accepted measure of surgical expertise (Kirkman, 2013), several measures were relied on to select expert participants (total number of LCs completed, whether LCs were performed on a routine basis in current practice, peer nomination). However, there are limitations to these measures. For example, it could be argued that peer nomination is subjective. Therefore it is possible that some participants were not ‘true’ experts. A third explanation is that the initial hypothesis that ‘slowing down’ moments would be consistent between experts was in fact incorrect. There is some evidence which suggests that clinical judgment is variable among a group of experienced clinicians, even for routine problems (Backlund, Danielsson, Bring, & Strender, 2000; Skaner, Strender, & Bring, 1998; Vancheri, Alletto, & Curcio, 2003). For example, when Vancheri et al. (2003) asked cardiologists and internists with at least 5 years of experience to estimate the probability of congestive heart failure for 30 case histories of real patients, they found variability within each group of specialists (but not between the two groups) both in terms of their probability estimates for a given case and in terms of the clinical features
that most strongly influenced their estimates (edema, history of infarction, cardiac rhythm, etc.). The variability of judgment between experts might make sense considering that each clinician has experienced a unique mixture of clinical scenarios, each one helping to shape the clinician’s interpretation of and response to particular cues (Klein, 1993; Schmidt, Norman, & Boshuizen, 1990). Whether a particular cue will initiate a transition to a more effortful mode depends on the individual’s subjective experience of that cue (i.e., whether it is recognised as being troublesome, unusual, unexpected, etc.), which in turn depends on past personal experiences (Louis & Sutton, 1991). In light of this literature, surgeons are likely no exception and, like other clinicians, might sometimes differ in their interpretation of the same available information, resulting in the heterogeneous ‘slowing down’ moments that were observed.

This study found that the ‘slowing down’ phenomenon was a common experience for expert surgeons; each expert felt concerned or uncertain in response to emergent intraoperative cues. However, when these ‘slowing down’ moments occurred seemed to be variable from one expert to the next. Further, participants sometimes experienced the same MC while focusing on different cues (see Section 4.5.2.3). Given the expert level of the participants in this study, it is likely that each surgeon would perform a safe operation and have a similar successful outcome if performing a LC on a given patient. However, the data suggest that the cognitive processes that underlie expert surgical performance could be less consistent than expected, and this could have important implications for how we teach and assess surgical judgment. It is important that surgical trainees learn to appreciate the critical cues that should initiate intraoperative ‘slowing down’ moments. While much of this learning is the result of hands-on experience in the operating room, it had been hoped that this study might be the first step towards a resource for supplemental training, providing trainees with opportunities to learn from the ‘slowing down’ moments of experts outside of the operating room at their own leisure. However, considering
there seems to be limited consensus about ‘slowing down’ moments between experts, what to teach could be a complicated issue with the methodology explored in this study. In a similar sense, an objective assessment of judgment could also be problematic given that there is no apparent ‘gold standard’ of expert performance to serve as a reference point. These issues suggest that further research is required to understand how ‘slowing down’ moments might best be taught and assessed; studies of clinical competence in social work (Regehr, Bogo, Regehr, & Power, 2007) and professionalism in medicine (Ginsburg, Regehr, & Mylopoulos, 2009) suggest that perhaps some constructs are too complex and nuanced to be amenable to traditional educational strategies (e.g., checklist or numeric scale assessments) and require alternative approaches.

5.2 Evaluation of the Methodology

Although the combination of the think aloud method and videotaped surgical procedures resulted in new conceptual insights about the ‘slowing down’ phenomenon, it was also more complicated than anticipated. In this section, an outline of some of the advantages and limitations of this methodology as a tool for studying ‘slowing down’ moments is provided.

Despite the challenges, this methodology allowed the capture of many instances of ‘slowing down’ moments that would have been difficult (if not impossible) to capture using other methods. For example, the detection and deliberate avoidance of a critical structure (see Section 4.2.2.4) might not be associated with an observable behaviour. Moreover, the standardised format made it possible to compare ‘slowing down’ moments (or at least a subset of ‘slowing down’ moments) between individual participants. The methodology was able to demonstrate some (although limited) evidence of expert consistency in ‘slowing down’ moments, suggesting
there are certain critical events during a procedure that should be recognised. These consistent ‘slowing down’ moments could be used for teaching purposes and could perhaps form the basis of an assessment tool. For example, if a group of experts expressed concern in response to a particular event while watching a videotaped surgical procedure, it might not be unreasonable to expect a senior resident who is about to enter independent practice to be able to recognise the same danger in an assessment context.

There were several limitations to the chosen methodology. First, a ‘slowing down’ moment is defined as a recruitment of cognitive resources, but trying to define the boundaries of a ‘slowing down’ moment in a continuous stream of text is a time-consuming and somewhat subjective task. Although certain cues such as the participants’ choice of words or tone of voice were used to determine when their ‘slowing down’ moments were taking place, it is quite probable that certain ‘slowing down’ moments were missed while coding the transcripts. However, based on the earlier discussion of ‘slowing down’ as a continuous process, it can be seen that even if we had access to continuous quantitative measures of the participant’s cognitive effort throughout the case (Figure 12), deciding which peaks were significant enough to represent discrete ‘slowing down’ moments (as opposed to fine-tuning) would again be a subjective matter. Second, due to the detailed and technical nature of the data, it was essential that content experts (i.e., surgeons) be involved in the data analysis. Considering the busy schedules of clinicians and the time-consuming process of coding the transcripts, conducting a study of ‘slowing down’ moments on a larger scale (e.g., expanding to other procedures within general surgery or to other surgical specialties) could be challenging. Third, verbal reports are, as suggested above, incomplete records of a participant’s thoughts and it is possible that certain ‘slowing down’ moments were not captured; if true, this would place important limitations on the value of this methodology for high-stakes assessment purposes. Fourth, this study took place in a standardised environment
removed from the clinical setting. Although the standardisation of the task made it possible to compare ‘slowing down’ moments between individual surgeons (a main objective of this study), it also resulted in a task that lacked certain important elements of actual practice. Participants were unable to interact with the surgical environment and therefore had no control over the outcome of a case. They also lacked any personal connection to the case, as they did not know the patients. It is therefore unlikely that the thought processes that were engaged in this study were exactly the same as those that would be engaged in actual practice. Also, as a result of the passive nature of the task, it is unlikely that participants approached the task with the same level of cognitive effort as they would if they were in the operating room observing an actual case involving a real patient (even though efforts were made to ensure a high level of engagement in the task by selecting what were thought to be interesting cases). It is therefore important to acknowledge the possibility that a participant’s performance in this artificial setting might not be an accurate representation of her or his performance in situ.

5.3 Reflections on the Research Process

Although the research objectives (page 27) were elaborated at the onset of this project, the research process turned out to be more challenging than anticipated. It had been expected that ‘slowing down’ moments would ‘jump out’ of the transcripts as critical moments involving concern, but what was not anticipated were the more subtle forms of ‘slowing down’ moments that were detected. The process of extracting and categorising ‘slowing down’ moments from the transcripts (research question #1) was therefore time-consuming. It had also been assumed that once the ‘slowing down’ moments were identified, the process of comparing them between participants (research question #2) would be relatively simple. However, considering the large
number of ‘slowing down’ statements that were identified in the data set (more than 400 of them) and the complex nuances of these statements, comparing them all proved to be an unrealistic goal and efforts were therefore focused on a well-defined subset of ‘slowing down’ moments (the ‘moments of concern’).

In general, it seemed as though the data in this study were congruent with those of other studies that also used the think aloud method to capture surgical judgment. For example, both the participants in this study and Ghaderi et al.’s (2012) participants made predictions about the anticipated level of challenge of the case and recognised anatomic abnormalities. Instances of perceptual expertise (“…now we’re getting that kind of bleeding I think. That kind of bleeding you see when you’re around the common bile duct, yeah, that just sort of constant ooze”) and predictions (“what’s going to happen now is…it’s going to be a little unclear as to aah, where the cystic duct is”) were also captured that were similar to those described in Dominguez’s (1997) work. These similarities between this study’s data and those of other studies with similar methodologies suggest that the think aloud method was correctly used in this study.

In hindsight, however, there are certain aspects of the methods that could have been improved or done differently. First, the think aloud training session lasted approximately 5 minutes. A longer training session may have helped to eliminate some of the teaching and commentating statements, infrequent as they were (however, it may also have made this study less feasible by increasing the time required of already busy participants). Second, the practice problems were nonvisual tasks (with the exception of the dot counting problem; see Appendix A); a visual task (e.g., thinking aloud while looking at a picture or video) may have been a more suitable choice considering the perceptual focus of the research task. Third, the presence of teaching and commentating statements in the data set suggests that the primary researcher’s presence
sometimes distracted participants from the task at hand. Having the primary researcher sit or stand behind the participant while collecting the data may have helped to reduce the influence of his presence on the participant’s performance, which may have resulted in verbal reports that more accurately reflected the participant’s inner thoughts. Finally, a limitation of the chosen methodology was the inability to establish the relative significance of each MC. A future version of this study could require that participants rate their level of concern while watching the videos (perhaps using a dial that could be adjusted at any time) or that they only speak when they have significant concerns (instead of thinking aloud at all times).
Chapter 6
Conclusions

This study contributes to our current understanding of ‘slowing down’ in surgical practice. Using the think aloud method, the categorical framework of ‘slowing down’ moments was expanded, which increases awareness of the circumstances that prompt surgeons to transition to a more effortful mode of operating. The conceptualisation of ‘slowing down’ moments as discrete events was challenged, and it was proposed instead that ‘slowing down’ can also be thought of as a continuous process involving countless fluctuations in cognitive effort. As a final finding, comparisons of a specific group of ‘slowing down’ moments across expert surgeon participants revealed that many ‘slowing down’ moments seem to be inconsistent between experts, or at least inconsistently demonstrable using this methodology. Although some of the inconsistency may be linked to limitations of the methodology and sampling error, it is also possible that ‘slowing down’ moments are in fact somewhat variable from one expert to another. The inconsistency between experts suggests that formal teaching and assessment of the ‘slowing down’ phenomenon may be challenging and will require further exploratory research.
Chapter 7
Future Directions

Several potential avenues of research have emerged from the findings of this study. Several categories of situationally responsive ‘slowing down’ moments emerged from the think aloud data. Since these categories were identified using an artificial task, a reasonable next step would be to present the categories to a group of general surgeons (perhaps the original group of participants) to make sure that they resonate with the surgeons’ experiences in actual practice. A subsequent study could also be conducted to determine whether the findings of this study can be replicated in other contexts (e.g., another surgical subspecialty, such as orthopaedics).

The apparent inconsistency in ‘slowing down’ moments between expert surgeons suggests that further exploratory work is required before strategies can be developed for teaching and assessing this phenomenon in a more formal manner. As stated in the Discussion chapter, one of the limitations of the methodology used in this study was the inability to determine which of the MCs were the most critical. To improve understanding of the data collected in this study, a next step would be to ask a group of expert surgeons to rate the importance of the 54 MCs that were identified. These ratings would then make it possible to determine whether the most consistent MCs were in fact the most critical and not just the most obvious.

Future work should also consider exploring the ‘slowing down’ moments of surgical trainees. In pilot work conducted with junior residents, the residents failed to notice some of the critical events that elicited consistent concern from expert participants. For example, a variation of the present study could be conducted using surgeons at various levels of training (e.g., junior residents, senior residents, fellows, and staff) to determine how ‘slowing down’ moments vary between groups. To obviate some of the limitations of the methodology used in this study, more
specific instructions (e.g., explicitly asking participants to speak up when they are concerned) could perhaps be used to ensure that the desired data are captured. Another possible modification of the methodology would involve providing participants with a method to provide a quantitative measure of their level of concern (for example, a numbered dial that the participant could adjust at any time during the case).

As a final suggestion, future research should identify the perceptual and cognitive abilities that mediate the ‘slowing down’ transition. These abilities could be compared between surgical residents at various levels of training to identify potential barriers to appropriate ‘slowing down.’ This knowledge could then inform the development of educational interventions that are appropriate for each level. As a hypothetical example, it is possible that junior residents fail to identify certain critical ‘slowing down’ moments because they have a limited understanding of the anatomy. In this case, improving a junior resident’s ability to ‘slow down’ appropriately could involve providing them with a learning intervention that helps develop pattern recognition of the anatomy, such as the perceptual learning module described by Guerlain et al. (2004). In this way, we could help make sure that surgeons in training develop the abilities they need to be able to recognise when an intraoperative situation is non-routine and requires a more effortful mode of thinking.
References


Moulton, C. A. (2010). *Exploring the role of 'slowing down when you should' in operative surgical judgment.* University of Toronto.


Appendix A

Think Aloud Training Exercises

_Researcher:_ I will start by familiarising you with the procedure for thinking aloud. In order to get more comfortable with speaking your thoughts aloud, I am going to ask you to think aloud as you work on the answers to some practice questions. What I mean by think aloud is that I want you to say your thoughts out loud from the moment you finish hearing a practice question until you say the final answer. I would like you to talk aloud as much as you comfortably can during that time. Don’t try to plan or explain what you say. Just act as if you are alone and speaking to yourself. Keep talking while you are coming up with the answer to each question. If you are silent for a long time, I’ll remind you to think aloud. Do you understand what I would like you to do? We will begin with a practice question. First, listen to the question. Then, answer it as soon as you can. Are you ready?

1) _Please name ten animals that live in the zoo._

   Good. Did you have any other thoughts as you came up with the answer to this question? I want you to think those thoughts out loud as they occur to you. I want to remind you that you don’t have to explain your thoughts to me. Just say what you are thinking, even if it’s not a complete thought or you think it doesn’t make sense. Listen to the next question and try to think of the answer as soon as you can. Are you ready?

2) _What is the sixth letter after B? (Answer: H)_

   Thank you. Chances are that the letter “H” didn't immediately occur to you after hearing the question. You probably had to go through several steps to find the answer. Had you summarised your thinking during the last question rather than reported the sequence of actual thoughts aloud, you might have said that you found the letter H by counting through the alphabet. But, when people solve this problem out loud, they usually say a sequence of individual letters, such as B, then C, D, E, F, and G, before the answer H. And this is what we want. We want to have the most accurate, detailed report of thoughts as possible, instead of a summary of those thoughts.

   Let’s do another practice problem. Think aloud while you generate the answer. I’m going to show you a dot grid and ask you to tell me how many dots are in the grid. Are you ready?

3) _How many dots are there? (Answer: 27)_.

   Great. Can you recall any other thoughts?
4) *How many months begin with the letter J?* (Answer: 3)

Any questions so far?

*(Researcher: At this point decide whether or not to use extra practice problems.)*

Thank you. We will now begin the main experiment or do another practice problem (What is the fifth letter before M? *Answer: H*)