Errors in healthcare: A Proposed Study Investigating a Universal Methodology for Improving Patient Handoffs and Limiting Healthcare Errors

Olivia J. Couture
Bard College, oc3843@bard.edu

Errors in healthcare:
A Proposed Study Investigating a Universal Methodology for Improving Patient Handoffs and Limiting Healthcare Errors

Senior Project Submitted to

The Division of Science, Mathematics, and Computing

of Bard College

by

Olivia Couture

Annandale-on-Hudson, New York

May 2022

Acknowledgments

I would first like to acknowledge my senior project advisor, Justin Hulbert, for his dedication and support throughout this whole project. Without his guidance my project wouldn't be what it is today. Thank you Justin for everything you did for me.

Next I would like to thank Professor Kim and Jude for supporting me by being on my board. Specifically to Professor Jude, thank you for being one of the main reasons I attended Bard, and for your continuous support through my four years. And a special thanks for joining my board.

I would like to thank my friends and family for their support though this all, I wouldn't have been able to do it without them. Thank you to my parents, my sister Jenna, and my brother Evan. A special thank you to Tristan for being my main support over the past four years.

Finally I would like to thank the Tivoli Fire Department for being the most memorable part of my four years here. The two years I spent in the department were some of the best of my college experience. I will never forget the friends and family I gained through the department.

Table of Contents

Abstract

Medical errors are greatly considered to be one of the leading causes of death in the United States, with around 250,000 deaths occurring per year (Anderson & Abrahamson, 2017). Many studies have investigated the causes of these errors, and have found communication to be a major source. The present proposal is a two-study design first planned to investigate the effects of interruptions on the time taken to return to the Tower of London Task. This methodology has first been observed in Hodgetts and Jones (2006) who investigated this using undergraduate students as participants, and found a main effect of both duration and complexity of interruptions. This replication will use registered nurses to complete the London Tower Task while unexpected interruptions occur, attempting to see if these findings replicate. With those expectations, I go on to propose a follow-up study designed to test the effectiveness of an intervention to better handle such interruptions in the healthcare environment. The implementation of the I-PASS mnemonic has proven to be successful in minimizing healthcare errors by 23% in pediatric residency hospitals (Starmer et al., 2014). This proposal will not only replicate this methodology in other residency hospitals but the use of an I-PASS written template will be used to improve this methodology. The implementation of the I-PASS mnemonic and template will help to manage the intrinsic load of nurses during patient handoffs, by increasing their working memory capacity. This implementation will universalize patient handoffs within the healthcare system, and in turn, minimize the occurrence of health care errors and the death that come from it.

Keywords: nursing, cognitive load theory, patient handoff, I-PASS

Errors in Healthcare

Each year around 400,000 hospitalized patients experience some type of preventable harm (James, 2013). Some examples of preventable harm in the healthcare field include hospital-acquired infections, medication errors, or wrong-site surgery. Aside from preventable harm, medical errors cost approximately $20 billion and result in approximately 100,000 deaths per year (Rodziewicz, 2021). Many studies have investigated the root causes of these errors, and potential solutions to this problem. The main causes that have been indicated in some studies are fatigue due to long duty hours, inadequate experience, inadequate supervision, and communication errors (Bari & Rathore, 2016). Although there are potential solutions to these problems, they are difficult to achieve without a full systematic change, costing hospitals money and time they don't have. Hence the need to tailor and test the efficacy of such interventions prior to scaling up their implementation.

Although many of these errors occur in hospitals, they can happen within the whole healthcare system. This system is made up of hospitals, outpatient offices, inpatient facilities, etc. Any place where the care of an individual is given can be quantified as part of this system. The purpose of this proposal is to investigate hospitals, as they contain the most amount of healthcare interactions. The two main types of errors defined in a healthcare setting are errors of omission, and errors of commission. Errors of omission are defined as resulting from actions not taken, for example forgetting to strap a patient into a wheelchair. Conversely, errors of commission are defined as resulting from the wrong action taken, for example administering a medication to a patient who has a known allergy (Rodziewicz, 2021). When looking in a healthcare setting, these errors can take on many forms. Therefore, it is important to identify these types of errors to try and determine which ones are more frequent and if there is one that is more preventable. The review conducted by Rodziewicz, (2021) aimed to provide definitions for the specific types of errors that fall under the umbrella of errors of omission and commission. Some of these errors are more preventable than others and are highlighted as an important part of future research that could maximize patient safety. These errors include latent errors, negligence, near-miss events, root cause error, and–the worst possible outcome–a sentinel event (Rodziewicz, 2021). All of these errors are further explained below, to demonstrate their importance in the healthcare field.

Latent errors occur within the system, equipment, or related to the organization. For example, different hospitals use different equipment for their machinery. Therefore, if someone who was new to the hospital tried to use this machinery they might do it wrong and, in turn, injure the patient. Negligence is another event that increases the chance of error when the quality of a patient's standard of care is failing to be met. This can range from not checking up on a patient enough, or not checking on a pathology report leading to a misdiagnosis. Arguably one of the most important types of events is a near-miss event. This occurs when an action that could have adverse patient consequences doesn't occur either due to chance or someone intervening. It is hard for a healthcare worker to understand the effects an error can have without actually making a mistake. Therefore, a near-miss offers the most opportunity to learn because the individual is able to focus on what could have happened, as opposed to dealing with whatever repercussions occurred due to the error. Another major form of error, root cause, occurs when there is a lack of communication or teamwork. It occurs when a decision made by one individual actually becomes questioned or changed by another. Although many individuals don't like being criticized, when it comes to patient safety it is necessary to inform others about what they are doing. This type of teamwork and communication should be praised and practiced in order to eliminate mistakes made with patients. Finally, the last and most serious type of preventable error, a sentinel event, is defined by the joint commission as “any unexpected occurrence involving death or serious physical or psychological injury” (Rodziewicz, 2021). The Joint Commission currently certifies more than 22,000 health care organizations in the United States (The Joint Commission, 2022). It is recognized as the global leader in healthcare accreditation. This aforementioned commission is the healthcare system's main way of defining and judging the severity of errors in healthcare. All of these errors are highlighted as preventable, and the emphasis of the reviews by Rodziewicz (2021) was to identify the necessity for further studies aiming to minimize these risks.

As described by Laxmisan et al. (2007), there has been a clear paradigm shift over the last century with respect to medical errors. The opinions have changed from a culture of silence to a culture of blame, and finally a culture of safety. The first known opinions of medical errors began with the idea of concealing them. There was a culture of disgrace and many healthcare workers felt it would be better to keep quiet than to report their errors made. As the culture changed, there was a turn to blame. Whenever patients are involved in an error, their loved ones are quick to place the blame somewhere, and hospitals today and in the past are left scrambling to pay for these mistakes. Not only are hospitals paying through patient complaints, but through insurance claims made by patients (Tanne, 2008). One study found that each adverse event cost the hospital an average of $2013, and each negligent injury cost an average of $1246 (Tanne, 2008).

Although this culture can still be seen today, there is a different emphasis on what is important, and that is the patient’s safety. The goal has evolved to learning from our previous mistakes and doing everything possible to prevent them from happening. That is why, today, hours of safety training are required not only before starting a job in healthcare but are mandatory yearly for everyone. However, how much is this helping? Can this training be better adapted to prevent more than just careless mistakes, or is there another underlying culprit that could be leading to these mistakes? For example, researchers have highlighted multitasking, the ability to complete multiple tasks at once, as a potential culprit. This, coupled with the constant interruptions that happen in a healthcare setting, creates potentially dangerous situations where a patient's safety could be at risk.

What is Multitasking?

Multitasking has previously been named an important everyday skill; yet, many studies have recognized it as a major contributor to cognitive overload (Laxmisan et al., 2007). This begs the main question, how much information is too much, and to what threshold can a person multitask without increasing instances of error? How can one investigate the cognitive load on healthcare workers, and establish a threshold of how much someone can multitask without jeopardizing the safety of patients (Laxmisan et al., 2007).

Many studies have investigated not only cognitive load but how it is affected by both multitasking and interruptions. In one study conducted by Coiera et al. (2002), they looked at the communication loads on clinical staff in two separate emergency departments. It was concluded that the combination of interruptions and multitasking can result in more errors due to the disruption of memory processes. Although the authors acknowledge that handoffs, multitasking, and interruptions are very common practices in the emergency department, they propose better communication tools to assist in minimizing errors during these occasions (Coera et al., 2002).

Multitasking can be defined as completing more than one task at the same time. However within this broad definition are two smaller more specific types of multitasking, concurrent and interleaved. Concurrent multitasking is defined as the performance of two or more actions simultaneously, whereas interleaved multi-tasking is the management of multiple tasks in which there is switching between tasks that are progressing in parallel (Douglas et al., 2017). Both of these definitions can be applied to the healthcare field in their own unique ways. An example of concurrent multitasking would be a doctor continuously writing an order while answering a colleague's questions. The key aspect of this type of multitasking is the execution of the tasks simultaneously. Conversely using this same example however with interleaved multi-tasking, the doctor would stop writing the notes to answer the colleague's questions. This in turn forces the individual to perform task switching, which is the crucial differentiation between the two types of multitasking.

Evidence from the literature demonstrates that errors are increased with both types of multitasking. However, when individuals choose to multitask as opposed to being compelled to do so by an outside stimulus, the chance of error is smaller (Douglas, et al., 2017). It also has been found that increasing the time between the presentation of the first and second task reduces the delay in responding to the second task (Douglas, et al., 2017). When it comes to interleaved multitasking, multiple studies have attempted to look at performance and errors made. Interleaved multitasking forces individuals to switch tasks, making them have to reconfigure or activate a new task set, which has been found to increase response time and the likelihood of errors (Douglas, et al., 2017). Finally, in a study conducted by Reiman et al. (2015), they found that voluntary multitasking allows the individual to control the amount of cognitive load they impose on themselves. The participants were instructed to generate their own plan for the order they will complete tasks (including at least one switch between tasks), and then execute said tasks in the order in which they put them. They found that participants tended to generate the least complex sequence of tasks to perform, which suggests that voluntary task switching enables individuals to use the strategic organization to lower their cognitive load (Reiman et al., 2015).

Multitasking in Healthcare

Multitasking can have many different meanings and can be practiced in different ways depending on the situation in which they are used. Many have sought to define what it means to multitask in a healthcare setting with one prominent definition being created. Specifically in the emergency room multitasking has been defined as "the ability to prioritize and implement the evaluation and management of multiple patients in the emergency department, including handling interruptions and task switching in order to provide optimal care" (Table 5, Perina et al., 2012). This definition surpasses the more generic explanations of multitasking by emphasizing the ability to handle interruptions and task switching. Instead of merely defining it as completing two tasks simultaneously, multitasking here incorporates realistic occurrences in an emergency room. Therefore this creates growth in the original definition, and now integrates the ability to cope with an interruption, then to prioritize task switching, and finally return to the original task seamlessly.

In general working as a healthcare professional requires the ability to handle more than one patient at once. Therefore it is difficult to come to the conclusion that a healthcare professional is not multitasking. However many researchers have claimed with scientific reason that it is impossible for humans to multitask. The main argument is that humans are incapable of "true multitasking"; or completing more than one task simultaneously with another. Hence why many researchers have begged the question, can we multitask? A review conducted by Stephens and Fairbanks 2012, brought together sources to support the argument that humans do have the ability to multitask, however just with limited capabilities. They believe that our brains are powerful enough to multitask, however, there are limitations to our short-term memory, and predictable errors made when task switching and task stacking (Stephens & Fairbanks, 2012). The most important aspect of this claim is that the authors are suggesting that the capabilities and limitations of our brains must be addressed when training individuals entering a healthcare profession.

Whenever an interruption occurs, there is a shift in attention from the primary task to the interrupting task (e.g., responding to a question from a coworker). If the primary task is an automatic process, such as suturing, there is a possibility the individual can complete both tasks simultaneously. However, if the primary task is a controlled task, which has the same cognitive domain as the interrupting task (e.g., taking a history from a patient and answering a colleague's question), the individual is forced to decide whether or not they want to task switch (Heng, 2014). If the individual decides to stop the first task of history taking, and start the new task of answering the question then they are relying on short-term memory to return to the original task. In the case of task stacking (the need to switch between multiple tasks), short-term memory is less reliable in making an easy return to the original task (Heng, 2014). Similar findings have been highlighted in a paper by Stephens and Fairbanks (2012), where for some situations multitasking has been found to be possible whereas in others it is not. For example, physicians giving a diagnosis is considered a controlled process, therefore it cannot be done in parallel with another task. On the other hand, completing sutures can be considered an automatic process, therefore it can be done simultaneously while completing a task like answering a question (Stephens & Fairbanks, 2012). One limitation that they highlight is the limited capability of working short-term memory. As tasks are stacked in our brains it becomes more and more difficult to return to the original task without errors. This limitation coupled with the external stimulus of frequent interruptions leaves a breeding ground for potential error that in the end puts the patients at risk.

Alkahtani et al. (2015) reviewed the current state of studies investigating multitasking in healthcare systems. This review emphasizes the need for a study that demonstrates the effect of multitasking on stress and error rates (Alkahtani et al., 2015). Douglas et al. (2017) furthered the research on this topic by connecting cognitive psychology and multitasking in healthcare. They aimed to define two types of multitasking: concurrent multitasking which involves the performance of two or more actions simultaneously (e.g., a nurse writing an order while answering a colleague's question), and interleaved multitasking, or the management of two or more tasks by switching between them (e.g., a nurse writing notes and then stopping to answer a colleague before returning to the notes; Douglas et al., 2017). Although this review looked at both types of multitasking, they focused more on interleaved multitasking as it pertains more to the medical field. One study included in this paper found that when individuals have to switch between tasks they are forced to 'reconfigure,' activating a new task set. When this occurred, it resulted in increased response time to the new task, and an increased likelihood of errors (Douglas et al., 2017). Although much of the literature reviewed included the cognitive aspect of multitasking, the authors noted that most of the studies that connected multitasking to the healthcare system simply looked at the occurrences instead of actual manipulation.

Interruptions

Although the need and danger of multitasking have been established, the question that remains is why is there a need for multitasking? Aside from the need to be able to manage multiple patients, there are constant interruptions that are occurring both internally and externally. One study that observed emergency physicians in three emergency departments found that in 180-minute periods clinicians were interrupted on average 31 times (Chisholm et al., 2000). Although this study was merely observing the number of interruptions occurring in a shift, it paves the way for other studies to be developed. For example, Westbrook et al. (2010) also investigated how interruptions and multitasking can have an effect on the performance of emergency physicians. Similarly, this paper observed the number of interruptions occurring per hour during a shift in the emergency department. It was found that on average clinicians were interrupted 6.6 times per hour, 11% of all tasks were interrupted, and 3.3% of those interruptions occurred more than once (Westbrook et al., 2010). They also found that interrupted tasks were completed more quickly than uninterrupted tasks. When investigating task switching, it was hypothesized that physicians were shortening the primary task in order to compensate for making up time, and were often found making shortcuts (Westbrook et al., 2010). This highlights a major room for medical error that could jeopardize the overall safety of patients.

The current literature highlights the effects that interruptions have on a cognitive base, however, the field is lacking studies that look at the effect in a healthcare setting. Most studies have focused on observing just how many interruptions are occurring, yet they aren't truly investigating the type of interruption and its true effect on the workforce. One study, in particular, conducted by Hodgetts and Jones (2006), investigated the effects of unexpected interruptions on the execution phase of the Tower of London task. The main goal of this paper was to see if there is a relationship between duration and complexity of interruption, and the time taken to return to the original task. The authors acknowledged that although it would be predicted that the longer the interruption the more time it takes to return to the original task, the literature is relatively inconclusive on these effects. Participants were instructed to complete the Tower of London task (Figure. 3) and were interrupted on different trials. Between the two groups, interruptions occurred either for 6 seconds or 18 seconds. In pilot groups, the interruption of 18 seconds, compared to 6 seconds was deemed different enough to produce effects of length, if there are any, and was not too long to make it impossible to return to the original task (Hodgetts & Jones, 2006). After running the trials it was found that there was a significant main effect of the duration of the interruption, and complexity of interruption on the time it took to make the next move in the task. Although this finding is important to cognitive psychology, its ability to generalize to other scenarios is unknown. By investigating this study through a healthcare lens, specifically using nurses as participants, different results could be concluded on the effects of interruptions.

Cognitive Load Theory

Cognitive load theory is a model of cognitive architecture used to better understand how human memory works (Van Merrienboer & Sweller, 2010). The main premise is constructed from the assumption that a limited working memory and an unlimited long-term memory comes from knowledge stored as schemas in long-term memory. Human memory consists of four main aspects: working memory (WM), sensory memory (SM), short-term memory (STM), and long-term memory (LTM). A stimulus begins in the sensory memory where it is first taken in for only a brief second. It then can enter short-term memory where it can be recalled for a short period of time. Working memory is similar to short-term memory in that it holds information for a short period of time, however here is where an individual can manipulate the information. Finally, with proper rehearsal, information can enter long-term memory where it can be recalled in the future due to schemas (Villines, 2020). There are many aspects of everyday life that can interfere with one’s memory. One specific contributor that many researchers have investigated is cognitive load and how that can affect memory, or one’s ability to complete tasks. The three types of cognitive load are as follows: intrinsic load, a direct result of the complexity of performed task and the level of expertise of the learner; extraneous load, a result of a process, not that does not contribute directly to learning, and germane load, which is caused by learning processes that deal with intrinsic cognitive load (Van Merrienboer & Sweller, 2010). An instance within everyday life of how cognitive load can be exemplified would be two people trying to explain the same thing at the same time.

Even though the schemata are stored in long-term memory, the important construction and refinements happen in the working memory. As previously mentioned, one premise of working memory is that each learner has a limit to their capacity. This is the most important aspect of understanding how an individual can learn. Given learning requires immense processing of information, a learner isn’t able to process information when the cognitive load of the task exceeds that working memory capacity. Therefore the ability of an individual to learn is dependent on optimizing the load they are under. Cognitive load theory identifies the three types of load: intrinsic (load associated with the task), extraneous (load not essential to the task), and germane (load imposed by the learner’s use of cognitive strategies to organize information for storage in long term memory). Further development of CLT has better helped to understand complex learning, which makes it able to be applied to the healthcare field.

Extraneous Load

The extraneous cognitive load is defined as the load that is imposed due to poor instructional design or other factors that divert the attention away from the learning environment (Wilby & Paravattil, 2021). In other words, this type of load is often referred to as a "distraction". This could include times when task instructions are not clear or other factors like background noise or interruptions. An example of this within a school setting would be reading while other people are carrying on a conversation in the background. Within the healthcare field, there are a plethora of examples of how the extraneous load can be increased. One example of this would be interruptions, a concept that has been previously exemplified in this paper. Although we understand how often interruptions are occurring in the healthcare field, we cannot fully understand the effects that they have on the cognitive load of an individual. Thus far it is understood that interruptions can increase an extraneous load of an individual. However when looking at the overall cognitive load on a healthcare worker, just limiting extraneous load will not be successful enough in reducing the overall load. This solution cannot be justified in the healthcare field, as interruptions can be a necessary aspect of safe and high-quality care. Regardless of how detrimental these interruptions can have on the cognitive load of a healthcare worker, it is unrealistic to propose limiting extraneous loads of healthcare workers to limit overall errors. Therefore by investigating the other types of loads, like intrinsic load, and how they can be managed a better solution to limit healthcare errors can be proposed.

Intrinsic Load

Intrinsic cognitive load can be defined as the complexity of the task, which is a function of the association between a learner's expertise and the nature of the task. The task complexity is related to the number of elements and the degree of element interactivity necessary to be learned by the WM. The interactivity of elements can be defined as the extent to which information can be learned in isolation from each other (Wilby & Paravattil, 2021). Therefore, elements would be considered to have high interactivity if multiple pieces of information are needed to be processed at the same time to learn a task. Conversely, low element interactivity would be information that can be learned in isolation. To better understand how to distinguish these two things, take the concept of an escape-room-based learning task for example. For some this would be considered a complex task to be learning, however, the mental workload of an individual will vary based on their expertise in completing previous escape rooms and solving puzzles. Therefore as a whole, intrinsic load represents the complexity of a task being learned. For example, in the school setting (for an average high school student) a task with a higher intrinsic load would be completing a calculus problem, whereas a task with a lower intrinsic load would be calculating 2 + 2. Intrinsic load is ultimately governed by the number of elements that interact with each other that are being processed simultaneously.

Generally, the inherent difficulty of information being learned cannot be changed, but it can be managed. When investigating cognitive load, and ways to manage it, typically researchers have highlighted the ability to manage the intrinsic load. For example, when a teacher is presenting a slideshow to a student, the student's intrinsic load can be better managed when the font is plain as opposed to cursive. Although there are ways to limit intrinsic load, complex problems constantly arise especially in the healthcare field. These problems can be addressed by bringing previously processed, organized, and stored information from LTM to WM in the form of schemata or chunks (Fraser et al., 2015). This gives rise to concerns about the capacity of the WM, however since these elements have already been learned they don't increase the capacity of the WM. Fraser et al., (2015) further explained the ability to manage intrinsic load through specific instructional strategies. The first strategy is segmenting effect, which is the process of segmenting information. When tasks are reduced into manageable chunks, these chunks are practiced until they are stored in long-term memory. This means that a large amount of information can eventually be manipulated in WM from the formation of these chunks. These instructional strategies are often seen with students in healthcare training, where they are initially taught the basic sciences and clinical skills before moving on to more complex concepts. Empirical evidence of this strategy was tested by Brydes et al., (2010), in teaching clinical intravenous catheterization. They found that the group of students assigned to "progressive learning", in which the task difficulty was gradually increased, was learned more efficiently than students who trained with easy and complex tasks. This was a perplexing finding because progressive learning was better than just completing a simple task. Brydes et al., (2010) argued that this finding was due to sequencing effects. When knowledge is gradually built in a low to high complex sequence, learners can increase their knowledge stored in long-term memory and ultimately gain exposure to the high complexity. This experiment demonstrates that complex cognitive processes might not be possible if WM is hindered by cognitive overload.

The second instructional strategy proposed by Fraser et al., (2015), is scaffolding. This concept is generally referred to as any guidance given to a learner. As mentioned by Fraser et al., (2015), a study by Clark et al., (2006) referred to scaffolding as an instructional strategy that provides enough support to reach intended learning outcomes. They acknowledge that this can take many forms such as visual aids, and outlines. Within the healthcare field, patient handoffs have already been acknowledged as a vulnerable practice for errors. Currently, there is no generalized methodology for conducting these handoffs, which is ultimately increasing the intrinsic load. Therefore experiment two of these proposals aim to not only generalize this process but to implement a mnemonic and written template to manage the intrinsic load of healthcare workers during a handoff. By managing intrinsic load through this implementation, the overall cognitive load of healthcare workers will be minimized, ultimately helping to reduce errors.

Other studies have investigated ways to manage the intrinsic load. Pollock et al., (2002) conducted an experiment that presented learners with a simple-to-complex sequence of instruction. The intrinsic load of the learner was minimized in the first part, by not presenting all the information at once. This was done by presenting isolated elements that could be processed individually. In the second part, all the information was presented at once, including the interactions among elements. Although the understanding was lower in the first part, this was compensated by the presentation of all interacting elements in the second phase. Therefore, presenting isolated elements before demonstrating how they interact, ultimately increased the ability of the learner to understand (Pollock et al., 2002). This concept of managing intrinsic load can be applied to learning in the medical field. By manipulating the learning environment it can gradually increase the number of interacting elements. For example, after a medical student learns something like diagnosing chronic low back pain, they shouldn't begin with real patients to practice. By starting will low-fidelity environments (e.g. textbook problems), then going to medium-fidelity environments (computer-simulated patients), then to the high-fidelity environment (simulated patients played by actors), and finally ending with the real environment, medical students will have less of an intrinsic load and therefore are less susceptible to errors (Van Merriënboer & Sweller, 2010). This serves as just another example of how intrinsic load can be minimized, and once again highlights this management as necessary to decrease the overall cognitive load of an individual.

Germane Load

Germane cognitive load can be defined by the mental effort an individual must dedicate to create permanent storage of knowledge (Wilby & Paravattil, 2021). This differs from both extraneous and intrinsic load as it deals with the cognitive effort of learning, and not the interfering nature of the other two loads. In basic terms, the germane load is the processing, construction, and automation of schemas and scripts. This load is typically considered a positive thing as it helps to facilitate the formation of long-term memories for easier future recall. Things that can contribute to this load would be assessments, training, receiving feedback, or self-reflection, as they aid in schema formation. Although this load is beneficial and necessary for an individual to learn, it nevertheless adds to the overall load of an individual. However, unlike extraneous and intrinsic load, the germane load isn't managed in order to limit the overall load. It is a necessary aspect of the memory system and loads on an individual that helps them to learn. Therefore when looking at the overall cognitive load of an individual, the best way to minimize it is through the limitation of extraneous load (although this has already been expressed as difficult in the healthcare field), and management of intrinsic load.

Cognitive Load Theory and Transfer of Patients

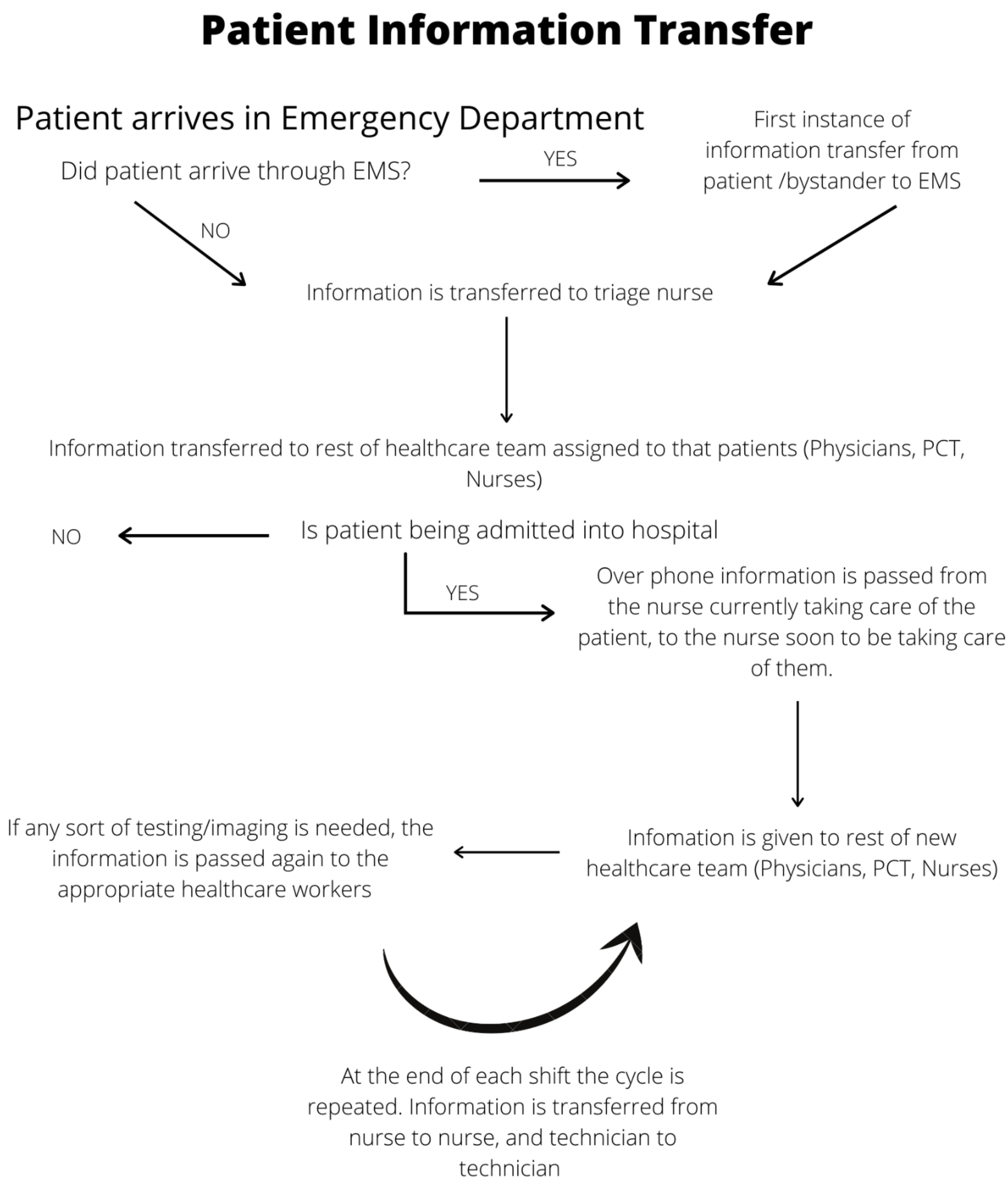
Transfer of information is constantly occurring in any healthcare setting, particularly at the beginning of shifts when patients are handed off to another shift. Although handoffs create vulnerabilities ripe for medical errors, they are necessary for healthcare practices, unless or until healthcare professionals are replaced by automation. Indeed, patient handoffs can be beneficial because they offer a new perspective and help to limit certain fixation errors. As necessary as this practice is, there is enormous room for error which can ultimately put a patient at risk. Previous studies have investigated handoffs in other industries; however, they haven't studied this phenomenon in hospital settings, with manipulation as opposed to mere observation. For example, in an observational study conducted by Patterson et al. (2004), they investigated information handoffs in four settings with high consequences for failure. The four settings included: NASA Johnson Space Center in Texas, nuclear power generation plants in Canada, a railroad dispatch center in the United States, and an ambulance dispatch center in Toronto. Before the intervention, the handoffs left much room for error, and without a standardized methodology, each handoff was conducted in its own fashion. In the study, the authors investigate certain strategies for improving handoff effectiveness and efficiency. Some of these strategies that were used and found to be beneficial include: completing face-to-face handoffs, limiting interruptions during updates, reading back information to ensure it was accurately received, and updating the information in the same order each time. Although this study didn't specifically investigate these strategies in a healthcare setting, they have potential uses for improving handoffs in a hospital.

Patient Handoffs

When looking at major sources of preventable errors in the healthcare system, one practice stands out: patient handoffs. Transfers of patients are constantly occurring and information is always being shared. Within the healthcare system, collaboration is a necessary aspect of patient care. This means that not only is patient information being shared amongst workers but, the movement of patients from the care of one person to another is a very delicate process. This process of patient transfer, along with all their information is a complicated series of events, with the hope that it leads to proper care (Figure 1).

Figure One

How information is transferred in hospital



This flow of communication begins as early as when the patient enters the emergency room. Once there, either the patient or a bystander offers information to a triage nurse, who makes the determination of what room the patient might go in. If this hypothetical patient is taken in by EMS, this information begins even earlier at the scene where the patient was picked up. After that, more information is gathered from the nurse, which is then shared with the patient's patient care technician (PCT) and physician. This would mark just the first occurrence of the transfer of information, from healthcare workers in the hospital. Then, as the patient is being assessed, it could be determined that they need to be admitted into the hospital. The patient would then be moved out of the emergency room and transferred to the appropriate floor. This transfer of information almost always occurs over the phone, where the triage nurse speaks with the nurse that will be caring for the patient. Although the patient might have the same physician, their technician will change. Finally, at the end of each shift, a new technician and nurse will be given the information about the patient in order to take over care. This highlights the patient handoffs that occur at least twice a day every day when shifts change. These shift changes almost always occur at 0700 or 1900. Given this is when the most amount of information is handed off, the chances of error are increased. It is inevitable that these transfers occur; however, when they are not done carefully, there is immense room for error. By forgetting to report how a patient is able to get out of bed themselves, what their allergies are, or even how often their vital signs need to be completed, it opens the door for mistakes to be made.

The Joint Commission currently certifies more than 22,000 health care organizations in the United States (The Joint Commission, 2022). The Commission is recognized as the global leader in healthcare accreditation and is also the healthcare system's main way of identifying errors. When discussing the potential for errors in handoffs the Joint commission wrote, "Potential for patient harm, from minor to severe, is introduced when the receiver gets information that is inaccurate, incomplete, not timely, misinterpreted, or otherwise not what is needed." When looking at the most serious type of error, sentinel errors, which can be defined as errors relating to death or serious harm to the ability to function, it was estimated that 50%-80% of these errors are occurring due to communication failures (Tiwary et al., 2019). By investigating the errors that are made in the healthcare field, it emphasizes the potential negative effect that miscommunication can have. Therefore, it is important to further delve into how these errors can occur, and potential solutions to them.

Transfer of patient care is a valuable and necessary means to ensure continuity of care. These kinds of handoffs include not only the transfer of the patient but the responsibility for all their information from one person to another. The most common occurrences of these transfers occur at the end of every shift, which is also the most vulnerable time for error. The end of the shift is when the caregiver is most tired, making them most susceptible to errors. The previous literature has reported investigations into why errors are occurring during patient handoffs. In a review conducted by Raeisi et al. (2019), it was determined that the leading challenge in this handoff process is a lack of effective communication. Therefore this study aims to investigate ways to minimize these errors, most specifically ones occurring from handoffs at the end of shifts.

When trying to understand the psychological process that can occur during a patient handoff, and why it is a vulnerable place for errors, researchers have often looked into the cognitive load. This process is best understood as to how people process information, and in layman's terms is the cognitive architecture of the brain. Cognitive load theory (CLT) was first introduced by John Sweller (1988). Sweller integrates three components of cognitive architecture: memory systems (sensory, working, and long-term memory), learning processes, and how cognitive load is imposed on working memory. When completing multiple activities simultaneously, each needing its own set of knowledge, skills, etc., there is a potential for an overload of the learner. This can occur in many different professions, or in something as simple as multiple people talking to you at once. Understanding how CLT works can help professionals to understand why it might contribute to errors during patient handoffs.

Handoffs through the lens of CLT

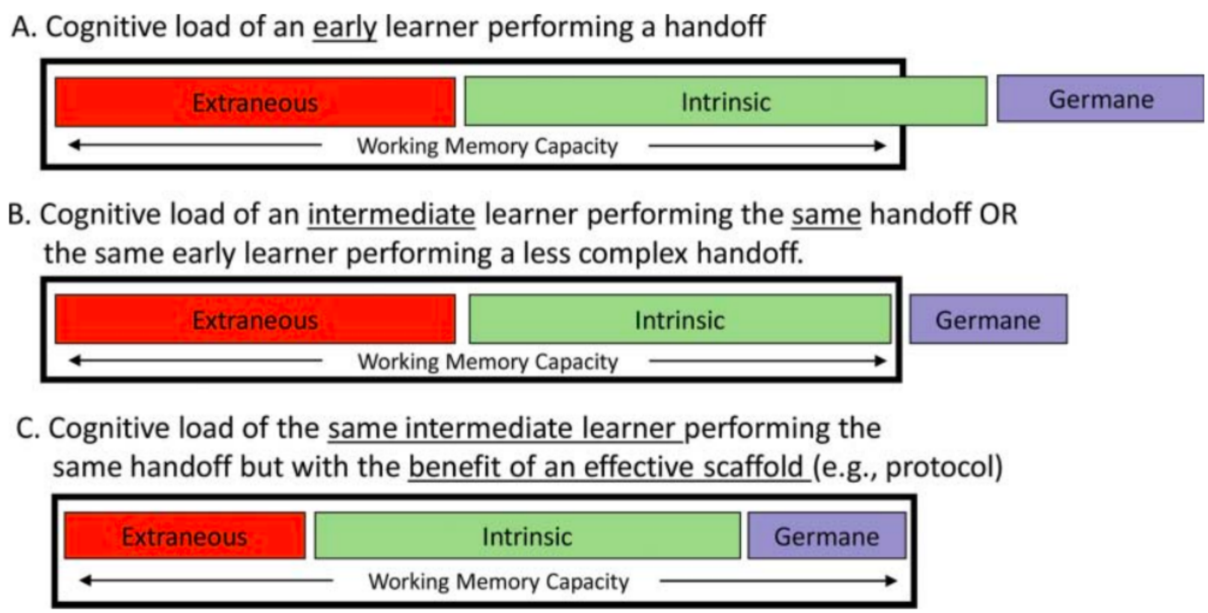
Few studies have fully investigated the effects of cognitive load theory as it pertains to patient handoffs. A recent publication by Young et al. (2016) looked at the complexity of patient handoffs through the lens of CLT. The authors acknowledged the need to explore mental workload during patient handoffs, and just how CLT can be applied. The sensory memory perceives visual and auditory information, with an enormous capacity despite its short retention period. Typically, this information doesn't reach conscious awareness and is retained for less than 0.25 to 2 seconds (Young et al., 2016). Once our attention is brought to the sensory information it then enters the domain of working memory. Here the information is reorganized into schemas in order to be stored in long-term memory (Young et al., 2016). When working in the healthcare setting there is a constant state of information processing. For information that is more pertinent once it moves to the working memory, there is a limitation on the amount of information an individual can hold. The current understanding is that working memory can hold 4 to 7 $\pm$ 2 pieces of independent information, and can only actively process no more than 2 to 4 elements at any given moment (Miller, 1956, as cited by Young et al., 2016).

Whenever there is a transfer of information there is an increase in the intrinsic load associated with processing this information. Through the lens of CLT, there are four aspects that make up an intrinsic load. First is the number of information elements, for example, how many different patients you are receiving information about. Secondly is the time available for a task, like if a rapid handoff is needed, this will require faster processing and more consumption of the working memory resources. The third is the effect to which the information elements interact with one another. When looking at handoffs, CLT brings to focus aspects that are beyond information volume or any time constraints (Young et al., 2016). Any form of uncertainty increases element interactivity, hence making the task more complex (Young et al., 2016). The last aspect of intrinsic load through the lens of CLT is the degree of expertise possessed by the learned. When an individual has previous knowledge of an element they already have the schema that incorporates all the information into one element (Young et al., 2016). Although the intrinsic load refers to the processing of outside information, the extraneous load gives rise to information that consumes working memory but is not essential to making sense of the information. Oftentimes handoffs increase extraneous load because it diverts resources from working memory to searching, rehearsing, and integrating the information (Young et al., 2016). Similarly, interruptions can also increase extraneous load because they divert resources from the working memory, to the sensory information that is occurring from the distraction. By understating how cognitive load theory works, it helps us to piece together how the mind functions, and how it can be distributed by interruptions. More specifically, by looking at how cognitive load can influence patient handoffs can give a better understanding of the potential room for error.

The final aspect of CLT that can be applied to patient handoff is germane load. What distinguishes this type of load from all the others is the fact that it is regulated by the individual. Each person is able to choose the effort they put into learning and self-monitoring to assess when inadequate understanding exists. Even if the individual wants to dedicate effort to learning, there won’t be sufficient working memory resources when the intrinsic and extraneous loads are too high (Young et al., 2016). Young and colleagues use a visual representation (Figure 2) to help demonstrate the additive effects of intrinsic, extraneous, and germane load with regard to the working memory capacity of the learner. In a less-experienced learner, the effects of the extraneous and intrinsic load exceed that of the working memory, leaving no resources for the germane load and limited resources for the intrinsic load (Figure 2A). Conversely, with an intermediate learner and the same handoff, there are working memory resources for both extraneous and intrinsic load, due to better-developed schemata (Figure 2B). Therefore, performance will be better than the less experienced learned; however, learning capabilities are limited due to a lack of resources for the germane load. When the same intermediate learner can reduce extraneous load measures like written templates, then there are working memory resources for all three types of loads, hence allowing the intermediate to perform better and learn (Figure 3A; Young et al., 2016). By looking at the complexity of patient handoffs through CLT, future ways to minimize load and errors can be studied.

Figure 2

Interaction of cognitive load and working memory capacity during a handoff



Note. Figured taken from Young et al. (2014)

Previous studies have investigated techniques used to regulate cognitive load; however, the focus to date has been mostly outside of the healthcare field. Thus far, the techniques highlight three strategies: reducing extraneous load, managing intrinsic load, and optimizing germane load (Young et al., 2016). Although much of these prescriptive strategies were derived outside the medical and clinical education, the fundamentals can still be applied in these situations. The medical field is often chaotic and uncontrollable therefore limiting extraneous load can be difficult. Certain practices like limiting interruptions, standardizing communication, and the use of written templates make all the information more accessible (Young et al, 2016). Even when the extraneous load is minimized, the load will still exceed that of working memory (Figure 2B). Therefore it is necessary to manage the intrinsic load to better limit the resources being used. Changing the way in which the information is presented, can lead to fewer working memory resources used. For example, this could be by simplifying the task (giving reports on fewer patients, or giving more time for reports), or breaking down information into 'partial tasks', or chunks in WM. Finally, enhancing the expertise of the learning by providing preparatory knowledge, or completing training before these tasks are completed (Young et al., 2016). Once the extraneous load is minimized, and intrinsic load is managed, more working memory resources are available for germane load, increasing learning opportunities.

Cognitive Load Theory and Memory

The concept of CLT is built upon many memory models that have been developed by researchers over the years. The first well-known model was proposed by Atkinson and Shiffrin in 1968. This model begins with information entering the mind through the sensory memory system, a subsystem that simultaneously processes visual and auditory information with the ability to retain the information for a matter of milliseconds (Young et al., 2014). In order for the information to be further retained, it is raised to awareness when it enters the working memory. Here, it is reorganized so that what is being taken in can be stored in long-term memory. In theory, long-term memory has limitless capacity; however, a map of sorts is required to find the information. The theory is that once information is retrieved from the LTM, the information goes back into the WM, where it can be utilized in a variety of ways. Therefore, while in the working memory the information is readily accessible in order to be retrieved more easily (Young et al., 2014).

Sensory Memory

Cognitive load theory can be understood both in and out of the healthcare field. Making comparisons between CLT and how an individual can become part of the healthcare field helps researchers to better understand what it takes to be in that field. The first part of the memory system that can be perceived through a healthcare lens is sensory memory. The sensory memory is the start of this pathway, and CLT is based on the principle that learners have separate channels for perceiving and processing auditory and visual information (Issa et al., 2011). In the medical field, the majority of this sensory information comes from sounds (spoken words), and images (printed words). Touch and smell are also important in the healthcare field; however, what is seen and heard is often what will provoke a response in a worker. Visual information is perceived by the eyes and is briefly held in the visual sensory memory system (iconic memory). Similar to what is perceived visually, what is heard is perceived by the ears and briefly held in the auditory sensory memory system (echoic memory). The combination of the visual and auditory systems allows for immense information to be taken in, however only for a very short amount of time (Mayer, 2010). Most of this information doesn't even reach our conscious awareness. However, when the individual brings their attention to the information, then it moves to the working memory.

Working Memory

In order for information to be moved to the working memory, the individual must be able to block out irrelevant stimuli (e.g., a coworker typing on a keyboard) and focus on relevant stimuli (e.g., the sound of an alarm notifying that a patient has gotten out of bed) from the sensory memory (Mayer, 2010). The current understanding of working memory is that it cannot hold more than seven (±2) information elements at a time, and can actively process two-to-four elements at any given moment (Kirschner et al., 2006). Unless this information is actively refreshed through maintenance rehearsal (e.g., repeating a lab value in one's head until it is able to be written down), amongst other methods, it will not be retained. This need to rehearse information limits the ability of an individual to learn. Therefore, complex learning tasks, like clinical responsibilities, require the information to be organized into chunks. These mental representations (or schemas) are able to connect with prior knowledge from long-term memory (Young et al., 2014). For example, if a student were to examine an ECG, they would be able to identify abnormalities based on what has previously been learned and stored in long-term memory. The dual-channel theory of working memory claims that sensory and visual channels are independent of one another (Mousavi et al., 1995). Therefore, this means information for each channel has its own capacity. Although this could be considered a downfall of the working memory capacity, it also means that the capacity can be expanded by utilizing both channels as opposed to one (Young et al., 2014). The ability of working memory to process information in combination with long-term memory is essential for the success of healthcare workers, among countless others. For example, a new nurse who is assessing a patient with chest pain will have a working memory at capacity simply by processing the different symptoms, etc. A more experienced student would be able to limit the information in the working memory capacity, due to the schemas of information they already have stored in the long-term memory.

These schemas are one way for the capacity of the WM to be expanded. However, there are other ways in which individuals can increase that capacity, one being through chunking. The process of chunking information is grouping data points into a larger whole, in order to improve the amount of information an individual can remember. For example, when trying to remember a phone number instead of remembering a sequence like 4-7-1-1-3-2-4 we chunk it into 471-1324. This allows people to remember lists and other information that exceeds the WM capacity. Another example of this is when chess players at the beginner level are briefly presented with middle-game positions in a chess game, they can recall about four pieces in the correct locations (Chase & Simon, 1973). Conversely, players at a more advanced level recall eight or 16 pieces correctly. This begs the question, if WM has limited capacity how can there be such a large difference in memory that is dependent on the level of expertise? It was predicted by Chase and Simon (1973) that the more experienced chess players are able to encode larger chunks, compared to the inexperienced players. In a review conducted by Thalmann et al., (2019), they reported multiple studies investigating the evidence for the fixed-chunk hypothesis. For example, they report an experiment from Chen and Cowan (2005), where participants varied in the chunk size they were instructed to use and found that approximately equal numbers of chunks were recalled across different chunk sizes. Participants were initially trained to remember the pairs in a set of word pairs, and the training continued until the recall was 100% accurate. Next, the individuals trained the words so that each word pair was considered its own chunk. Finally, participants were asked to recall a word list ranging from four to 12 words. It was found that regardless of order participants recalled approximately three chunks across all conditions, which is consistent with the assumption that WM capacity is limited to a fixed number of chunks (Chen & Cowan, 2005).

This experiment in particular gives rise to a couple of possibilities as to how chunks can help immediate recall. The first is that chunks require less capacity, hence freeing up capacity in the WM. Conversely, the second possibility is that information from LTM assists in the reconstruction of the complete chunk from parietal information in WM (Thalmann et al., 2019). This concept would be built on the assumption that chunks are maintained in WM the same way as random pairs. However, at recall, there is more LTM knowledge that is available for a previously learned pair compared to a random pair. In order to distinguish between these two possibilities Thalman et al., (2019) acknowledged the need to test memory for other items in the presence of a chunk. They claim that if chunking reduces the load on WM capacity, then the presence of a chunk in a memory list should improve memory for other items. Conversely, if chunks benefit from being more successfully reconstructed at retrieval, other information in WM wouldn't inherit the benefit. Hence they proposed an experiment to assess whether chunking information in WM frees the capacity to hold information that isn't chunked. They tested this theory by assessing the impact of chunking on the recall of other information that isn't chunked in WM. They best predict this through the use of the following example. Assuming two lists have to be remembered: List 1 = F-B-I-D-Q-B, and list 2 = I-F-B-D-Q-B. In the first list, the first three items form a single chunk (FBI), therefore it reduces the load on the WM from six to four items, and in turn, increases the capacity to hold the second part of the list. Therefore short-term retention of the second half of the list would be predicted to be better in List 1 than in List 2 (Thalman et al., 2019). This prediction demonstrates how chunking can benefit the capacity of the WM, and therefore can help to lay the groundwork for interventions that would help an individual maximize their WM capacity. For the healthcare field in particular by understanding the cognitive load that is placed on individuals, and how to manage it, certain implementations can be put into place to help decrease errors.

Long-Term Memory

In comparison to working memory, long-term memory is considered limitless in the capacity to store information, for all practical purposes. The schemata in long-term memory are stored based on complexity and ability to recall the information (Merrienboer & Sweller, 2010). As mentioned before, schemas are a pre-existing knowledge structure in memory. However, when a more dynamic type of schemata is considered, they are often described as scripts. For example, an illness script allows multiple signs and symptoms to be memorized in one construct. Instead of each individual element being considered separately, they are pieced together in LTM, reducing the load of the WM. There is no known limitation of WM when dealing with information retrieved from LTM. The LTM hold cognitive schemas that vary in their degree of complexity and automation, which ultimately alters the characteristics of WM (Merrienboer & Sweller, 2010).

The ability for long-term memory to enhance the capacity of working memory can be best understood using the analogy of a computer (Young et al., 2014). The hard drive or cloud server of the computer serves as long-term memory (with the ability to serve vast amounts of information), whereas the amount the computer can process differs like working memory. This random access memory has less capacity than the hard drive, however, it is essential to the computers because it allows data to be accessed quickly. When an individual has learned the information so that it is stored in long-term memory, it will ultimately allow their working memory to retain more information. By understanding the sensory, working, and long-term memory, and how it can relate to the healthcare field, researchers can better predict how errors might be made.

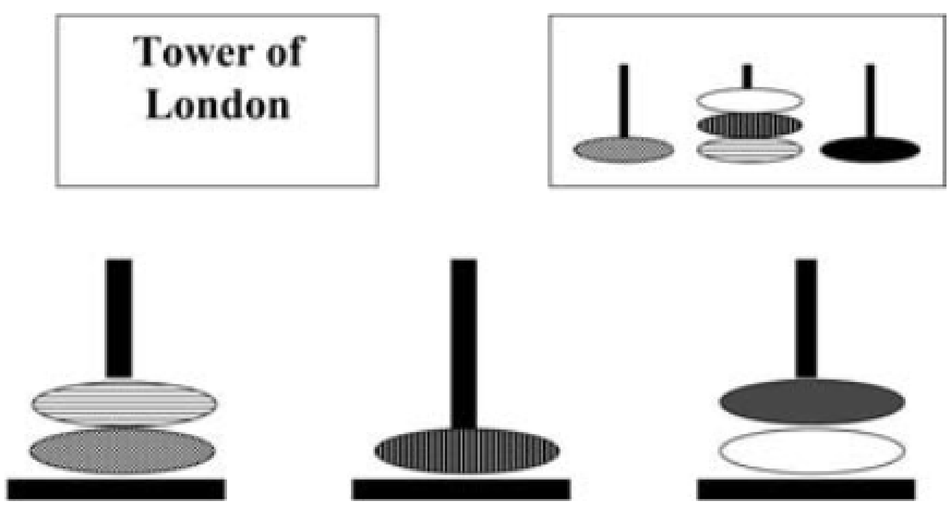
General Methodology

Tower of London Task

The Tower of London task (ToL) has long been used by researchers to test planning capabilities in both normal and abnormal populations (Philips et al., 2002). The task itself consists of three pegs, with different colored or patterned blocks on them (Figure 3). In order to complete the task, participants will be given an image of the desired orientation of blocks, and one by one will move each block to fit this orientation. This task has mostly been used in studies examining planning. For example, in a study conducted by Phillips et al. (2021), this task was used to measure planning abilities across different demographics like culture and age. In this study, they determined that in order to be successful in the task, moves would have to be planned in working memory, evaluated, monitored, and revised prior to the action. They measured success on the ToL task in four different ways: accuracy (how many problems were solved), efficiency (number of moves needed to solve the problem), reaction time (how fast a person completes the task), and rule breaks (whether or not the participants followed directions). In order to achieve a "perfect" score, the participant would solve all the problems without breaking a rule, using the least amount of moves, and doing it in the shortest time possible. The two rules for this task will include: only one block can be moved at a time, and no block can be held or placed outside the pegs while another ball is being moved. The main purpose of using the ToL task for this study was to investigate any differences in success between ages and cultures. The results demonstrate a main effect of age, where younger participants performed better on the task than older ones, but no main effect of culture, and no interaction (Phillips et al., 2021).

Figure Three

Tower of London Task



Note: The 5 disk Tower of London Task (Hodgetts & Jones 2006)

For the purpose of the present experiments, the ToL task will be used to examine the effects of interruption on planning. Being in the medical field, and more specifically a nurse requires planning in order to determine how to tackle the tasks at hand. Such skills are crucial in everyday life, however, the consequences are magnified in the healthcare field, where errors can mean a matter of life and death. Other professions also experience similar consequences when faced with errors from interruptions. Patterson et al. (2004) made direct observations on the NASA Space Center, two Canadian nuclear power plants, a railroad dispatch center in the U.S, and an ambulance dispatch center in Toronto. These observations included end-of-shift handoffs at each facility to evaluate communication strategies. Although errors in handoffs in these fields can have detrimental consequences, the number of interactions in these instances is increased for healthcare workers, specifically nurses. Having multiple patients forces the nurses to arrange each task that needs to be completed. Therefore, the ToL task not only emphasizes the use of planning but provides a simpler task in which response time after an interruption can be measured.

For present purposes, each participant will have to complete 25 trials of the ToL task while interruptions are occurring. A trial in the ToL task will include being presented with a final desired product, and the participant then is asked to move blocks one by one to meet this target state. Each trial can take a different number of moves; however, on certain trials, some form of an interruption will occur. The reason each trial can take a different number of moves is due to the variability of the participants' capabilities. All trials will be considered the same level of difficulty, and participants will complete the task in the fewest moves possible. After each interruption occurs, the time to make the next move will be measured to determine if there is an effect on the type and duration of the interruption.

Using methodology adapted from Hodgetts and Jones (2006), I propose to run a multi-step experiment investigating the effects of interruptions in a healthcare setting. All participants will be completing the ToL (Figure, 3), while unexpected interruptions occur. The original experiment used 20 undergraduate students who received course credit for participation as participants. This replication will use a different demographic of individuals as participants, specifically nurses, in order to determine if these results replicate. A larger participant pool will increase the power of the study and in turn validate the results. Throughout this experiment, the participants will be required to complete 25 trials. In Experiment 1a participants will be interrupted 8 times (Trials 4, 7, 10, 14 & 15, 19, 21, 25) and in Experiment 1b participants will be interrupted 6 times (Trials 4, 7, 12, & 15, 19, 25). In both studies, there will be matched control trials to compare to the time taken to make the fourth move after an interruption. For Experiment 1a these matched control trials will occur on trials 6, 9, 13, 16, 17, 20, 22, 23 and for Experiment 1b they will occur on trials 5, 9, 14, 17, 20, and 22. The fourth move was selected to universalize the move in which the time to return was being measured, in order to ensure all methodology is the same between participants.

This two-part experiment will offer insight into the effects of both duration and type of interruption on time to return to the ToL task. By making the study specific to one group of medical workers, we can better understand the effects of interruptions, which are a major contributor to an extraneous load of an individual. Finally, the results of the first experiment can guide another proposed experiment that investigates solutions to errors made from these interruptions, and how the healthcare field can better protect patients and minimize detrimental errors. More specifically the second proposed experiment will propose a mechanism to universalize the patient handoff system, ultimately managing intrinsic load, and decreasing the overall cognitive load on healthcare workers.

Experiment 1a

Participants

A sample of 500 nurses will be randomly selected to participate in this study. This study will be using nurses specifically with 3-5 years of in-field experience. The minimum number of years is used to prevent any skew in the data due to inexperience. Even though the task isn't specific to nursing, the present study wishes to investigate the effects of interruptions on the ToL task. A nurse right out of school would potentially not be as well adapted to interruptions that occur in the healthcare field. Conversely, the maximum is five years of experience to ensure results aren't skewed on the other side of the scale. An individual with too much experience might be well adapted to interruptions, and could potentially skew results. Selecting this range of experience will hopefully allow the effects of interruptions to be seen in this population. These individuals will be recruited using word of mouth, and advertisements within hospitals, and online. The aim will be to collect participants from areas all around the United States in order to ensure a good range in demographic. The original pilot data will have to be collected within a reasonable area, given the experiment takes place in a lab. Further development of this proposal would be beneficial to make these experiments capable of being completed online. This would allow nurses to participate from all over the country and the world. For this experiment specifically, participants will receive compensation for their participation in the experiment in order to ensure their time is valued.

Materials and Procedure

The first experiment is designed to investigate the effect of duration on interruptions (short vs long), one being 6 s and the other 18 s in duration. For both durations, the interruptions will include a mood checklist that must be completed. A mood checklist will include a box that appears on the screen with six ‘moods’ and a participant must select the one that represents them the most. The 6 s interruption was chosen from pilot work as a reasonable time to complete a mood checklist (Hodgetts & Jones, 2006). The 18 s interruption will consist of three consecutive mood checklists. This length was also selected as it was judged to be enough to produce effects of length but not long enough for goal decay. It is predicted that the time taken to make the next move in the sequence following an interruption would be greater than the time taken to make the equivalent fourth move in a controlled trial without interruptions. Based on the function of the decay model of Anderson and Douglass (2001), it would be expected that increasing the length of an interruption would decrease activation, and therefore cause reinstatement of the suspended goal. In addition to measuring the effect of interruption length, Hodgetts and Jones (2006) also investigated a prediction from Altmann and Trafton’s (1999) earlier goal model. It claims that goals are rehearsed in proportion to the length of time for which they will be suspended. Therefore, for half of the interruption trials, the participants will be told the duration of the interruption (one checklist or three checklists), and for the other half of the trials, the length of interruption will not be specified.

The task will be carried out on a computer, with the same program from Hodgetts and Jones (2006): a ToL program written in Visual Basic 6.0. On the computer screen, two sets of ToL disks and pegs will be displayed: the main display (33 X 25 cm) and, in the top right corner, the goal state (9 X 5 cm). The main display will consist of five different colored movable disks arranged on three pegs. The goal state display will appear with a white background and a diagram of the goal configuration on top. Finally, each initial state and goal state will differ for each trial.

To prevent from the participants rushing into making the first move, a planning phase will be present at the beginning of each trial. When a new trial begins the word *PLAN* will appear in the top left corner of the screen, while the display remains frozen. After eight seconds the word will be replaced with the word *SOLVE,* and participants will be able to make their first move. This time was selected by Hodgetts and Jones (2006) as reasonable for planning that was long enough to prevent participants from executing moves without planning, but short enough to prevent boredom or distractions. If the participants wished to take more time to plan before starting their moves, this could also be done after the words disappeared. The main reason for freezing the display for eight seconds is to emphasize the importance of planning before attempting the task. With the main display frozen, participants could rehearse the suspended goal for as long as they would like before completing the mood checklists. According to Altmann and Trafton's (1999) model, a goal that is suspended for longer will require a greater need for strengthening to mitigate the effects of decay (Altmann & Trafton, 1999, as cited in Hodgetts & Jones, 2006). However, this is contrary to the prediction that Anderson and Douglass (2001) made, that participants would not engage in any rehearsal process and instead would choose to pay the cost of forgetting as retrieval. This experiment will attempt to investigate these models by testing the effects of interruption length, and prior knowledge of interruption length, on both preparation and retrieval times.

Once the planning time is finished, disks on the large display can be moved one at a time by clicking on the peg holding the desired disk (the words *FROM HERE* would then appear) and then on the peg where they wish to move the disk (the words *TO HERE* would be displayed). This will then allow the disk to be moved to the desired position. Before beginning the task, the participants will be informed that no more than three disks can be held on a peg at a time; this is reinforced by a warning sound (beep) playing from the program if a fourth disk is attempted to be placed on one peg. Once each trial is completed, a pop-up will display notifying how many moves the participant took to complete the task, and by clicking *OK,* the next trial will be initiated.

During the interruption trials, after completion of the third move, a small gray button (2 X 4 cm) will appear at the bottom of the screen. Although the main display is still able to be seen, the screen will be frozen to encourage participants to click on the box. For half the trials the box will be labeled "mood (1 list), and on the other half of the trials it will be labeled "mood (3 lists)." This indicates the two types of interruptions that could occur, long (6 seconds) and short (18 seconds). By clicking on the bottom the participants will be brought to the interruption, one mood list (6-second interruption) or 3 mood lists (18-second interruption). The display will consist of a 6 x 4 cm checklist box placed in the center of the screen. Within the box will be six descriptions, one below the other in 12-point Times New Roman font. An example of this interruption would be a mood list as such: "extremely alert," "fairly alert," "slightly alert," "slightly tired," "fairly tired," and "extremely tired". The participant will select the response that most represents how they feel, and the response will be recorded in the program. For the 18-second interruption, the same methodology will be used; however, the three 6-second mood checklists will instead be displayed one after the other. The program will then record the time to resume the primary task, as well as the time between the appearance of the button on the screen, and when it is pressed (preparation time). For this experiment, no mood checklist will be repeated, in order to ensure the participants don't become accustomed to the lists. At the beginning of the experiment, standardized instructions will be given, and two practice trials are allowed so the participants could gain familiarity with the task. The first practice will consist of no interruption, however, the second included a long interruption to demonstrate the amount of time allotted before the list changes. In total the experiment will last on average 30 minutes.

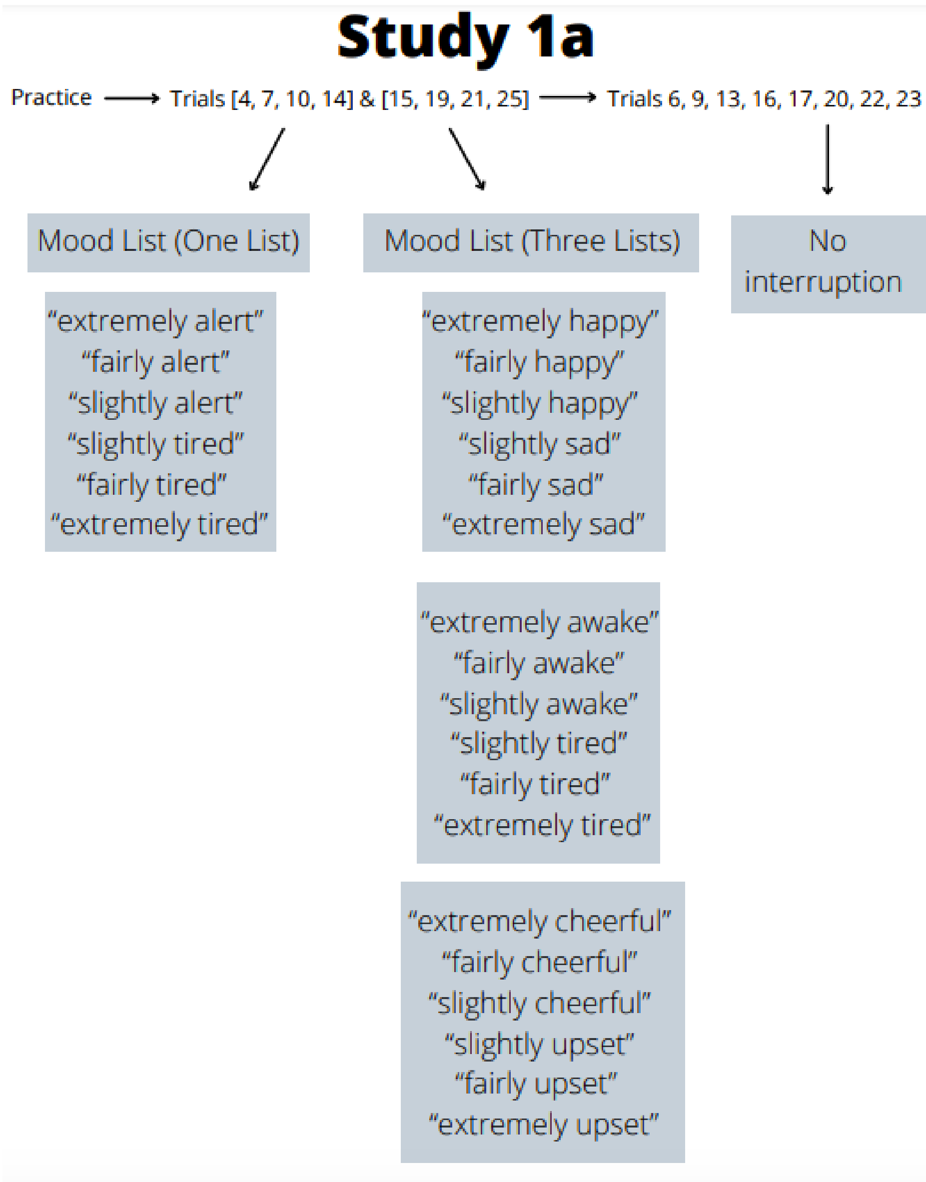
Design

The first part of this experiment will consist of a 2 (interruption duration) X 2 (knowledge of the interruption) repeated measures design measuring the time to return to the ToL task after different length interruptions (6 s or 18 s), and whether or not the participant knew this interruption during or before its onset. Across all the participants (25 trials per participant), the interruptions will always occur on the same trials (4, 7, 10, 14 & 15, 19, 21, 25). The only aspect that is different is the type of interruption that occurred. Half the participants received the short interruption with knowledge of the interruptions, on trials 4, 7, 10, and 14, and the long interruption on trials 15, 19, 21, and 25 (Figure 4). Conversely, the other half received the long interruption with knowledge of the interruption, on trials 4, 7, 10, and 14, and the short interruption on trials 15, 19, 21, and 25. This will total for a possibility of being randomized into four conditions: short interruption first and informed about the upcoming interruptions, short interruption first and not informed about the upcoming interruption, long interruptions first and informed about the upcoming interruptions, and finally long interruptions first and not informed about the upcoming interruption. Therefore this experiment will be conducted using a within-subject design. In order to conduct an analysis, each trial will be matched to a control (no interruption) trial (trials 6, 9, 13, 16, 17, 20, 22, 23). These control trials will consist of essentially the same problem requiring the same solution path but with the disk's color changed.

The remaining trials will consist of “filler problems" only requiring 4-6 moves to complete and will not be included in the final analysis. The dependent measures for this experiment will be preparation time and the time taken to make the fourth move.

Figure 4

Methodology for Study 1a



Predicted Results and Discussion

This procedure is designed to focus on the speed at which the tasks are completed, not the accuracy. As shown in Figure 3, given the simplicity of the ToL task, it would be predicted that the accuracy of the task is high in all conditions (similar findings were reported by Hodgetts & Jones, 2006). Therefore, the accuracy of the task will not be included in any analysis.

The preparation time measured in this experiment is operationalized by the time between the mood button appearing on the screen and it is pressed. This preparation time will be recorded for each condition. A sample distribution of preparation time means that would be expected to be found, based on results from Hodgetts and Jones (2006) can be seen in Table 2. A 2 (interruption duration) X 2 (knowledge of interruption duration) repeated measures analysis of variance (ANOVA) will be conducted to determine if there are any main effects or interactions. An ANOVA will be the best statistical test for this data as it will find differences between groups and find the relationship between the dependent and the independent variables. First, the analysis will determine if there is a main effect of each independent variable: interruption type and duration. If the main effect is found, this would mean that the variable has an effect on the dependent variable (time to make the fourth move) independently of the other. The first part of the experiment (Experiment 1a) not only investigated the different types of interruptions but also whether or not having prior knowledge of the interruption decreased the time taken to make the next move. Based on the memory model of Anderson and Douglass (2001), it would be expected that there is no main effect or interaction of having prior knowledge of an interruption. This would mean that individuals would not rehearse goals more when they are expected to be suspended longer. Hodgetts and Jones (2006) acknowledged the need for further research to determine when rehearsal of goals may occur. They attributed the lack of rehearsal in this experiment to the cost of forgetting. In a simple task like the ToL task, the cost of an error is very minimal therefore diminishing the value of a mistake. Conversely, when the cost of forgetting is higher (for example with nurses), there might be more potential for rehearsing.

The second dependent measure (time to make the fourth move) was predicted based on data from Hodgetts and Jones (2006) (Table 3). In the interruption condition, this is the time from the end of the interruption until the completion of the next move. In the control condition, this is the time from the completion of the third move to the completion of the fourth move. As I previously predicted, the time taken to make the fourth move in the interruption trials would be greater than the time taken to make the fourth move in the control trials. Time to return to the ToL task in the interruption conditions will be compared using a 2 (interruption duration) X 2 (knowledge of interruption duration) repeated measures ANOVA. As previously mentioned, it would be predicted that having prior knowledge of the interruption duration would not have an effect on resumption times. Conversely, a significant main effect of interruption length would be expected. Specifically, I predict that participants would take significantly longer to make the fourth move when the interruption was longer (18 s compared to 6 s). No interaction would be expected to be found between prior knowledge and interruption length.

By comparing differences between interruption and control conditions in terms of the time taken to make the fourth move, it would substantiate the effect an interruption can have on a simple task like the ToL task. Even if this interruption is considered brief, it disrupts the ability to execute the solution to this task. The previous literature has demonstrated that participants become better at dealing with interruptions with practice (Brock & Mintz, 2003). In order to ensure this nuisance factor isn’t driving any observed effects in the study, the resumption rates will be compared to test if participants would recover from the interruption quicker, later in the experiment. All the rates to make the fourth move would be compared across interruption conditions and will be analyzed. A one-way repeated measures ANOVA will analyze the resumption times, and it would be predicted that they wouldn't be distinguishable from one another, regardless of when they occurred in the experiment, based on the results from Hodgetts and Jones (2006). They found there was no effect of where the interruption occurred, compared to resumptions rates. Therefore, it would mean that participants won't differ in the time taken to resume the task after the first interruption, all the way to the eighth interruption. It would be expected that there is no significant difference in the time taken to return to the task following any of the eight interruptions. However, there would be a significant difference between the time to return to the ToL task following a long interruption compared to a short interruption.

Table 2

Preparation Time (in Seconds) for Each Interruption Condition in Experiment 1a

Interruption Condition	*M*	*SE*
Short, duration known	2.83	
Short, duration unknown	2.72	
Long, duration known	2.67	
Long, duration unknown	2.48	

Note. Data predicted based on results from Hoggets & Jones (2006).

Table 3

Time to Make Fourth Move (in Seconds) in Experiment 1a

Interruption Condition	*M*	*SE*
No interruption	2.53	
Short, duration known	5.07	
Short, duration unknown	4.85	
Long, duration known	6.32	
Long, duration unknown	6.11	

Note. Data predicted based on results from Hoggets & Jones (2006).

Research regarding interruptions, and memory has acknowledged that the length of time in which an item is retained is unlikely the only factor in forgetting (Hodgetts & Jones, 2006). Thus far, research has demonstrated that loss of information from memory is defined as a function of decay over time, and factors such as the complexity of intervening tasks are less studied. Interruption complexity has been investigated, and predictions on how it affects memory can be based on the goal activation model. As cited in Hodgetts and Jones (2006), the goal-activation model predicts that a greater number of distractor goals will raise the interference level, and therefore make retrieval of the target goal more difficult. The "complexity of a goal can be considered in terms of "number of actions to perform, difficulty in executing steps, and amount of information managed" (Byrne & Bovair, 1997, p. 46). Given an increased complexity and greater number of goals, it would be predicted that there would be greater retroactive interference when trying to retrieve the goal from memory (Hodgetts & Jones, 2006). In order to test this theory, the second part of experiment one was proposed to interrupt participants with different complexity of tasks.

Experiment 1b

Participants

A sample of 500 nurses will be randomly selected to participate in this study. This study will be using nurses specifically with 3-5 years of in-field experience. The minimum number of years is used to prevent any skew in the data due to inexperience. Even though the task isn't specific to nursing, the present study wishes to investigate the effects of interruptions on the ToL task. A nurse right out of school would potentially not be as well adapted to interruptions that occur in the healthcare field. Conversely, the maximum is five years of experience to ensure results aren't skewed on the other side of the scale. An individual with too much experience might be well adapted to interruptions, and could potentially skew results. Selecting this range of experience will hopefully allow the effects of interruptions to be seen in this population. These individuals will be recruited using word of mouth, and advertisements within hospitals, and online. The aim will be to collect participants from areas all around the United States in order to ensure a good range in demographic. The original pilot data will have to be collected within a reasonable area, given the experiment takes place in a lab. Further development of this proposal would be beneficial to make these experiments capable of being completed online. This would allow nurses to participate from all over the country and the world. For this experiment specifically, participants will receive compensation for their participation in the experiment in order to ensure their time is valued.

Material and Procedure

The first part of this experiment attempted to replicate the findings from Hodgetts and Jones (2006), investigating the effect of duration of interruptions on time to return to the ToL task, however using a different participant pool. This experiment differs as it is looking into different tasks as these interruptions. The participants were either interrupted with a mood checklist (same methodology as Experiment 1a), or a verbal reasoning task like "A follows B - AB", followed by a required true/false response. This experiment disregards the time it takes for each participant to complete the different tasks, as that variable is investigated in Experiment 1a. Although both tasks involve reading and clicking with a mouse on the most appropriate statement, the verbal reasoning task is considered more demanding than the mood checklist. Unlike the verbal reasoning task, the mood checklist can be considered arbitrary given there is no right or wrong answer. Conversely, the verbal reasoning task involves "higher mental processes" because it relates to verbal intelligence (Baddeley, 1968, as cited by Hodgetts & Jones, 2006). A single reasoning task involves multiple elements like ordering A and B in a sentence, whether the sentence is positive or negative, is active or passive, or involves the term *follow* or *precede.* In order to come to a true or false conclusion, it would require processing, retention, and verification of the sentence given. It is predicted that the interruption of the verbal reasoning task would cause an increased time to return to the ToL task, compared to the mood list interruption.

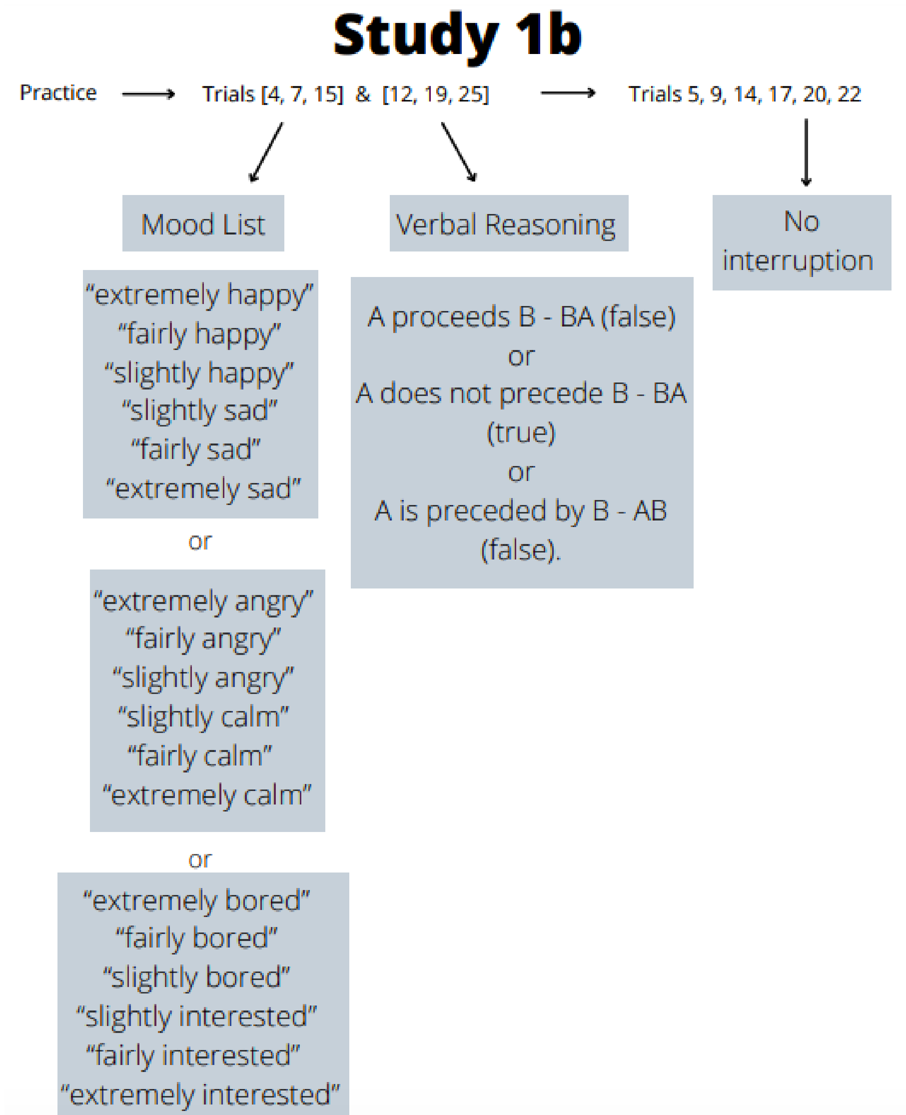
The same ToL program will be used from Experiment 1a, except the onset of interruption will be under the control of the computer. This time around, the interruption will occur automatically after the completion of the third movement. Experiment 1a contained eight different interruptions on 25 trials, whereas Experiment 1b will have six interruptions on 25 trials. The interruptions will either be a mood checklist ("happy-sad," "angry-calm," or "bored-interested"), or a verbal reasoning task which includes verifying the order of two subsequent letters: "A proceeds B - BA" (false), "A does not precede B - BA" (true), or "A is preceded by B - AB" (false). Both the interruptions will have a similar display. At the point of interruption (immediately following completion of the third move), the screen will go black and information will appear in a box in the center of the screen in 12-point Arial font. Participants have to either highlight the appropriate mood statement or *True* or *False* for the reasoning task. There will be no time constraints on the interruptions, and a button labeled *Continue* will appear below the other task which serves to start the timer on the return to the primary task. Before the performance in this study, the participants will be given two ToL practice trials and given examples of the mood and reasoning tasks separate from the ToL problems.

Design

A repeated measures design will be used where each participant completes six interruption trials and six matched control trials. The independent variable, type of interruption, will be counterbalanced so each participant will have both types of interruptions. Half the participants will have the reasoning tasks occur on Trials 4, 7, and 12, and mood checklists on Trials 15, 19, and 25. The other half of the participants will have the mood checklist on Trials 4, 7, and 12, and the reasoning tasks on Trials 15, 19, and 25. Finally, trials 5, 9, 14, 17, 20, and 22, will serve as control trials to compare to the manipulation (Figure 5). In the control trials, the goal state picture will be the same as in the manipulated trials, however with different colored blocks.

Figure 5

Methodology for Study 1b



Predicted Results and Discussion

Time to make the fourth move across conditions would be predicted as seen in Table 4. A one way repeated measures ANOVA will be used to evaluate the main effects and interactions of the independent variable, complexity of interruption. It would be expected that there is a significant effect of condition, meaning participants would take less time to make the fourth move following the mood checking as opposed to following the reasoning task, and participants would be the fastest in the control condition. The paired comparison would reveal a significant difference between all three conditions. In terms of the practice effects mentioned in Experiment 1a, it would be predicted that there would be no significant effect on resumption times in the order in which the six different interruptions occurred.

Unlike in Experiment 1a, Experiment 1b will have no focus on the length of time of the interruption. The short verbal reasoning tasks were selected to be roughly equal in time taken to complete as the mood checklist. Therefore the time taken on each task was in the control of the participant, and not the computer program. Leaving the time taken up to the participants, it will prevent any confounding factors that a time constraint might cause. For example, if the time constraint was too quick a participant could feel rushed, and it might hinder the problem-solving skills needed for the verbal reasoning task. On the other hand, if a time limit was set that was too long, it would leave room for a period of time unfilled by any stimuli. This could be considered a moment of low processing which would confound the results of the experiment, which investigated the effects of high processing tasks. Hence why the interruption length in the different conditions remained relatively similar, but not controlled. In order to ensure the data from this study reflects the complexity of the task, a statistical analysis will be conducted to determine if the effects of complexity are independent of interruption length. Given that in the previous experiment it was predicted that length of interruption did have an effect on resumption time, it is essential to distinguish between length and complexity of interruption in this experiment. It would be predicted that more time would be spent on the verbal reasoning task (M = 5.12 s), compared to the mood checklist (M = 4.08 s) (means are predicted based on results from Hodgetts and Jones (2006)). It would be expected that no significant correlation will be found between time spent on interruption and time to resume the ToL task, following both types of interruptions. If these predictions are supported by the data, then it could be concluded that the difference in resumption times between the two conditions (mood task vs verbal reasoning task) is the result of the complexity and not the time taken to complete the task.

These expected findings can't explicitly be predicted by the model described in Anderson and Douglass (2001), because that model claims that activation decreases are solely a time-based function. The effect of complexity is less studied and suggests that the activity along with the length of the activity can influence primary task performance. A recent study by Radovic and Manzey (2022) found that the complexity of interruptions had an effect on performance on an N-back test. The N-back test is widely used to assess working memory function. During this test, participants watch a series of stimuli and respond whenever a stimulus is presented that is the same as the one presented in previous trials (Owen et al., 2005). Although this study reflects the effect of the complexity of interruption it doesn't align with the Anderson and Douglass (2001) model, which maintains that memory for goals is equivalent to memory for other declarative memory elements. Therefore, further analysis and elaboration of these models are necessary to account for the complexity of the interrupting task.

The Anderson and Douglas (2001) model of memory incorporated memory as a whole as opposed to possible interference. They claim that goals are subject to decay and are typically retrieved in a last-in, first-out manner. Hence, the presence of distracting goals shouldn't have an impact on the process of goal retrieval, and the only factor would be how much time has passed. Conversely, Altmann and Trafton (2002) proposed a model that doesn't specify order, and claims goal selection is dependent on the goal activation. This means older goals are held in "mental clutter," or the mean activation of the most active distractor (Altmann & Trafton, 2002). Therefore, if activation of a target goal exceeds the threshold, then it is more likely to be sampled. During an interruption, if more goals are activated, then the interference level would be higher at the resumption, creating a higher interference for the target goal. Another theory is that goals are stored in declarative memory, and the more declarative information activated during an interruption the more interference (Altmann & Trafton, 2002). This theory helps to support what would be expected as the results from this study, looking at the complexity of interruptions.

Although the results of this study can be explained by the memory model by Altmann and Trafton (2002), there could be other possible explanations for the effects. As mentioned by Hodgetts and Jones (2006), it could be possible that the more complex a task, the more rehearsal is suppressed. The rehearsal of a suspended goal is needed in order to contain it within memory. This means with more complex interruptions a greater degree of strengthening would be needed to overcome any decay that might occur. Another interpretation of this effect relates to the cost of task-switching, instead of goal maintenance. In the current study, switching from a reasoning task is assumed to be more effortful than a mood checklist, regardless of the fact that both tasks don't relate to the goal at hand. Studies looking into task switching costs in more complex interruptions have found evidence to support this (Rubinstein et al., 2001) and to oppose this (Allport et al., 1994). Therefore in order to ensure the complexity of the interruption was affecting goal maintenance, Hodgetts and Jones (2002) completed another experiment in which participants had to complete the ToL task from memory. The current replication of Experiments 1a and 1b, from previous literature, are used to investigate the effects of interruptions on nurses. Therefore, no replication of the third study will be necessary.

Table 4

Time Taken (in Seconds) to Make the Fourth Move in Experiment 2

Condition	*M*	*SE*
Control	2.63	
Mood list	4.08	
Reasoning task	5.12	

Note. Data predicted based on results from Hoggets & Jones (2006).

General Results

Interruptions are known to have an effect on the completion of tasks. The preliminary work done by Hodgetts and Jones (2006), investigated interruptions on the ToL task and the effect they had on time to make the fourth move on a trial. For these studies, they used undergraduate students, who received college credit for participation. Although convenient, this type of sampling can cause problems when it comes to generalization. Given the population is specific to undergraduate students, and a small sample size, it is hard to claim that the findings would be the same in other populations. Therefore, to further investigate whether these findings would be consistent in the healthcare field, a population of nurses will be used and replicated with the Hodgetts and Jones (2006) methodology. It would be expected that there would be both a main effect of duration and type of interruption. Specifically, between the 6 s and 18 s interruptions, participants will take longer to return to the fourth move on the long interruption. In the second part of the experiment, time to return to the task will be measured after a simple vs. complex interruption. It would be expected that time to make the fourth move would be long after the complex interruption compared to the simple interruption. Once the effects of interruptions on the completion of the ToL task are better understood, a second study can be proposed to help limit the negative consequences of these interruptions occurring in the healthcare field. Given interruptions greatly increase an individual's extraneous load, the second proposal will aim to propose a way to manage the intrinsic load. This second experiment will investigate one of the most vulnerable practices for errors in the healthcare field, patient handoffs, and how to improve them to limit errors.

Experiment Two

Participants

The participants for this study will be selected based on the hospitals that agree to take part in the research. Once 10 residency programs from across the U.S. are selected based on if they have been previously identified as data-collection sites through professional academic networks (Sectish et al., 2010), then participants will be recruited. This study in particular will be looking at nurses at these residency hospitals. The nurses themselves are not in the residency program, but individuals who are completing their medical school studies will be. The reason behind a residency hospital being selected is because the healthcare workers at these hospitals are well adapted to learning and interventions. Given they are used to teaching and other programs, it would be easy to implement this program in one of these hospitals. Fifty nurses will be recruited from each hospital, making for 500 total participants. The nurses will be recruited in each hospital by the supervisors in each department. During a staff meeting, the supervisors will explain the study at hand and its potential benefits for changing patient handoffs. Then the nurses will give consent to participate in the study if they wish. Each hospital will be given a $100 gift card that participants have a chance to win when they take part in the study. If researchers are unable to recruit 50 nurses in a hospital, at least 30 will be needed before conducting the implementation at that hospital. This will ensure that the sample size is large enough to validate the results. Given this study will take place over 18 months, it is important to have enough participants at each site to make the data collection significant. There is no importance to the specialty or the number of years of experience a nurse has in this study.

Design

This experiment is a systematic-based intervention study, conducted on 10 inpatient units in 10 residency training programs, across hospitals in the United States. The 10 residency training programs will range in size and will be selected based on if they have been previously identified as data-collection sites through professional academic networks (Sectish et al., 2010). It will be ensured that each hospital doesn't already have a baseline practice for standardized handoffs. After receiving approval from the institutional review board and those participating in the hospitals, the interventions can begin. Data from each hospital will be collected six months pre-intervention and post-intervention. The data collected will be categorized on four different accounts: demographic information, quality of written and oral handoffs, and medical errors. During the 6 month intervention period, half the residents in the hospital will receive training on the I-PASS methodology and will use the I-PASS template. The other half of the residents will serve as a matched control group, in order to rule out confounding factors like changes in hospital protocol. However only residents who provided consent at the beginning of the study will contribute other observations to the data collected.

Intervention

The I-PASS handoff bundle has previously been developed from the best evidence from literature, previous experience, and previously published conceptual models (Starmer et al., 2014). For each hospital, the intervention will include the following: the I-PASS mnemonic (Figure 6); an I-PASS physical outline, which will serve as the anchoring component for both verbal and written handoffs; a 2-hour workshop aimed at improving communications skills along with practicing the I-PASS technique; direct observation tools used by coworkers to provide feedback; supplementary material and literature of where the I-PASS method has been successful; and finally branding tools used to promote the use of this method in the future. Well-established surveillance processes will be used to measure medical errors in each of the hospitals. These techniques include chart reviews, and implementation of research nurses observing both written and oral handoffs (Kaushal, 2002; Starmer et al., 2015). At each research hospital, research nurses will be used to manage most of the observations.

Figure 6

I-PASS mnemonic

I	Illness Severity	• Stable, "watcher," unstable
P	Patient Summary	• Summary statement • Events leading up to admission • Hospital course • Ongoing assessment • Plan
A	Action List	• To do list • Time line and ownership
S	Situation Awareness and Contingency Planning	• Know what's going on • Plan for what might happen
S	Synthesis by Receiver	• Receiver summarizes what was heard • Asks questions • Restates key action/to do items

Data collection on medical errors will be done in multiple different forms: through incident reports, solicited reports from nurses working on the unit, and daily medical-error reports from residents, collected through post-shift surveys. This methodology is adapted from

Starmer et al. (2015), to ensure using previously successful techniques for collecting this type of information. Written and oral handoffs will also be observed in manors similar to those in Starmer et al. (2015). Copies of written handoffs will be collected after shifts for evaluation, and audiotaped verbal handoffs will also be evaluated. Nurses are required to wear some form of technology throughout shifts to help with communication. For example one of the most common systems used is called Vocera (Figure 7). These devices are used to communicate throughout the hospital, making it easier to stay in contact with coworkers. Given nurses are accustomed to wearing these devices, it wouldn't be burdening them if a small recording tool was attached to their Vocera during a handoff. This would allow for both verbal and written handoffs to be observed pre and post-intervention of the I-PASS mnemonic and template.

Figure 7

Vocera technology used to communicate in hospitals

Note. Communication technology can differ between hospitals, however, the main function is to allow calling between healthcare workers.

Assessment of Written and Oral Handoffs

Each time a handoff occurs, both a written document and an in-person verbal exchange will be evaluated. Copies of all written handoffs will be collected each weekday morning and evening. At each hospital oral handoffs will be audiotaped when a research assistant is able to be present conducting time-motion observations. Research nurses who are aware of the intervention periods will evaluate a random sample of written handoff documents (8000 total, 800 per hospital [half from the morning, and half from the evening]), and audio recordings of oral handoffs (2000 total, 200 per hospital [half from the morning, and half from the evening]). This data will be collected both in the pre and post-intervention periods accounting for a total of 10,000 observations. Both the written and oral handoffs will be evaluated for the key handoff data elements. The rates of inclusions of these elements will then be compared to handoffs before and after the intervention. In order to account for confounding factors, length of stay, medical complexity, sex, and age of patients won't differ greatly between preintervention and postintervention periods.

Medical Error Being Measured

As previously mentioned, there are numerous different errors that can occur in the healthcare field. Specifically, this experiment aims to highlight preventable medical errors that can fall into different categories. For the purpose of this experiment medical errors will be broken into subtypes: preventable adverse events, near misses, and nonharmful medical errors (Table 1). The subtypes of medical errors will include: errors related to diagnosis, errors related to therapy other than medication or procedure, errors related to history and physical examination, medication-related errors, procedure-related errors, falls, and nosocomial infections. These subtypes will serve to further specify the errors that are being reported or observed, making it easier to distinguish where the most vulnerable practices for causing errors are. The second type of error, preventable adverse events are defined by the Institute for Healthcare Improvement as "unintended physical injury resulting from or contributes to by medical care (including the absence of indicated medical treatment), that requires additional monitoring, treatment, or hospitalization, or that results in death" (Wolf et al., 2018). An example of a preventable adverse in the healthcare field would include, wrong-site surgery or injury from a fall. The final type of medical error, near miss or nonharmful medical error, are errors that occur when an action that could have adverse patient consequences doesn't happen either due to chance or someone intervening. Although this doesn't leave the patient with physical evidence of the error, they are still important to monitor as they represent what could have happened. These types of errors also offer a large learning experience for the individual involved, as they are not dealing with the consequences of the error, but instead are able to reflect on what went wrong and why it could have happened. By categorizing the different types of errors, it allows for proper investigation into how useful the I-PASS mnemonic and template are in reducing different types of preventable medical errors.

Table 1

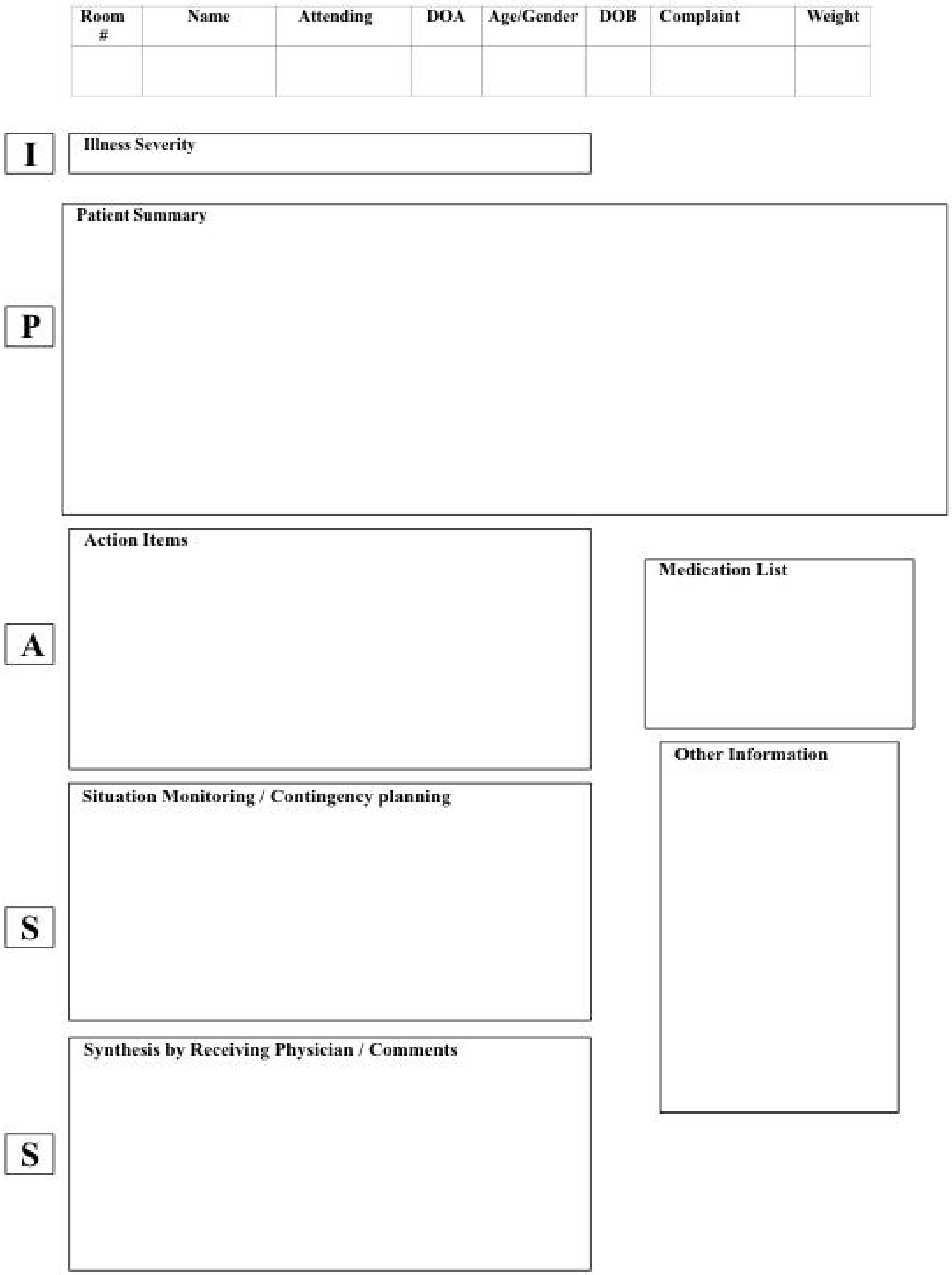
Medical Errors and their potential measures

Variable	*Before Implementation (N = ~ 5,000)*	*After Implementation (N = ~ 5,000)*	*P-value*
Overall medical errors			
Preventable adverse events			
Near Misses/nonharmful medical errors			
Medical-error subtype			
Errors related to diagnosis			
Errors related to therapy (other than medicine)			
Errors related to history and physical examination			
Medication-related errors			
Nosocomial infections			
Falls			

The I-PASS mnemonic will be used along with the template, to help manage the cognitive load of individuals when they are completing a handoff. This template will come in the form of one patient per page (Figure 8), or two patients per page (Figure 9). This allows the healthcare worker to be flexible with how many patients they wish to write per page. For some it might be easier to condense as many patients per page, however, others enjoy the space that comes with one patient per page. As seen in Figure 2, the extraneous, intrinsic, and germane load combine typically exceed that of the working memory capacity during a handoff. Especially when the individual is an early learner, the extraneous and intrinsic load itself exceeds the working memory. Researchers have pondered this problem in previous research, with many of the solutions surrounding minimizing extraneous load. Although in Experiment 1 the effects of interruptions can be seen, the solution to these interruptions isn't just to try and minimize them. It is reasonable to believe that they can be diminished by removing yourself from a situation where interruptions are high (for example completing a handoff in a patient's room where other healthcare workers cannot intrude). However, within the entire healthcare field, it isn't possible to simply just propose limiting extraneous load as the whole solution to these errors. The focus of this experiment is a proposal to minimize intrinsic load via the use of the I-PASS mnemonic and template. This can potentially help to limit errors by making the process universal and to help increase the capability of the working memory capacity by activating two channels (through both verbal and written handoffs). The use of a verbal and written handoff will potentially decrease the intrinsic load of the individual, allowing it to remain contained in the working memory capacity.

Figure 8

I-PASS template - one patient per page



Room #	Name	Attending	DOA	Age/Gender	DOB	Complaint	Weight

I — Illness Severity

P — Patient Summary

A — Action Items

Medication List

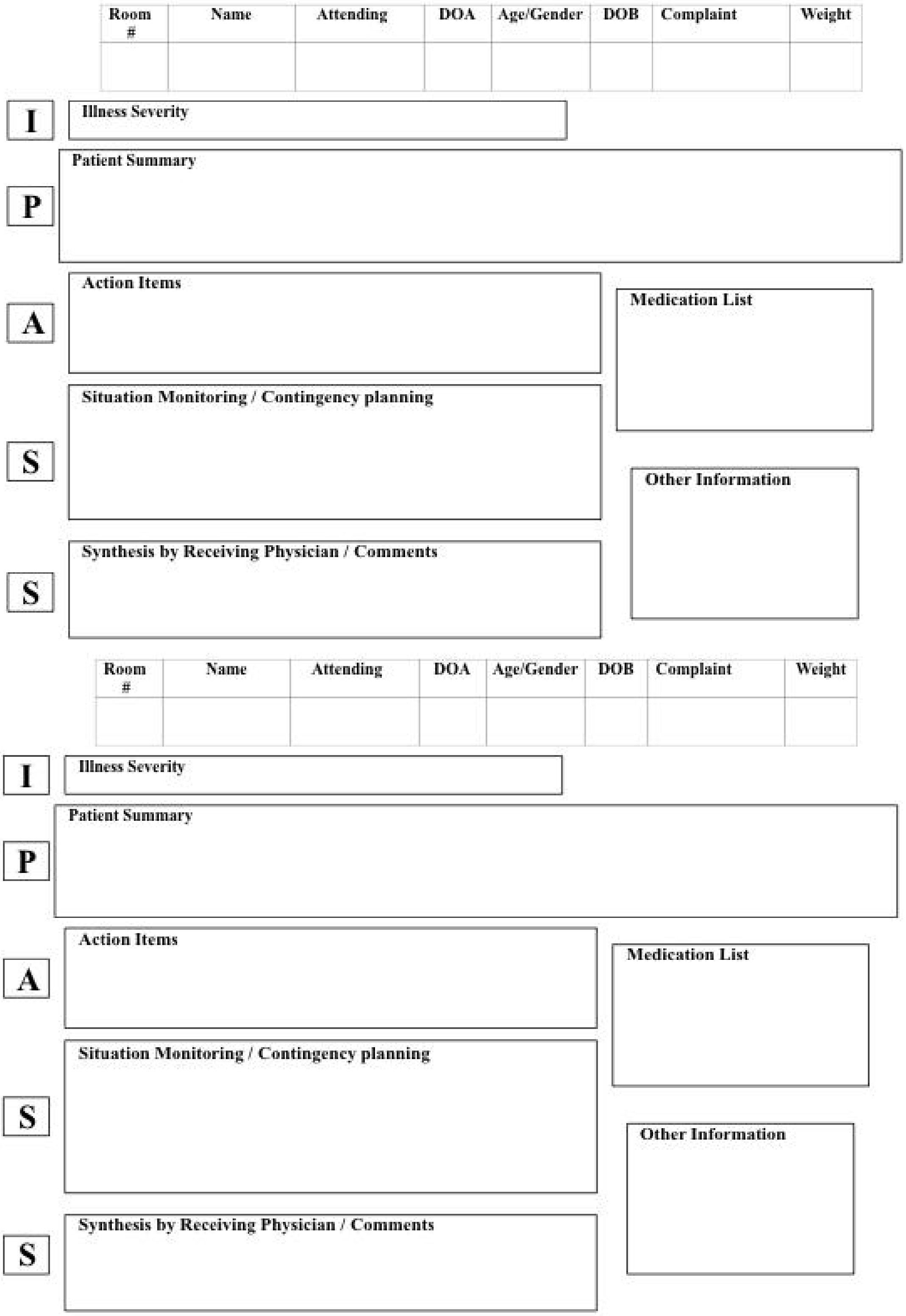
Other Information

S — Situation Monitoring / Contingency planning

S — Synthesis by Receiving Physician / Comments

Figure 9

I-PASS template - two patients per page



Room #	Name	Attending	DOA	Age/Gender	DOB	Complaint	Weight

I — Illness Severity

P — Patient Summary

A — Action Items

Medication List

S — Situation Monitoring / Contingency planning

Other Information

S — Synthesis by Receiving Physician / Comments

Room #	Name	Attending	DOA	Age/Gender	DOB	Complaint	Weight

I — Illness Severity

P — Patient Summary

A — Action Items

Medication List

S — Situation Monitoring / Contingency planning

Other Information

S — Synthesis by Receiving Physician / Comments

Predicted Results/Discussion

Proper implementation of the I-Pass mnemonic and template will be predicted to lead to decreased instances of medical error. Approximately 10,000 handoffs will be observed (around 5,000 in the pre-intervention and 5,000 in the post-intervention). In order to account for potential confounding factors, length of stay, medical complexity, and the sex and age of patients will not differ significantly between the preintervention and postintervention periods. Approximately 500` residents will be approached for this experiment, and written consent to participate will be collected. The response rates will be collected in post-shift surveys, meant as a means to survey accounts of medical error. It would be expected that response rates for the pre-intervention and post-intervention periods would be similar. Throughout the experiment, the research nurses will keep track of these responses to ensure they are being completed.

When comparing the number of medical errors that occurred in the pre-intervention and post-intervention periods, a significant decrease is expected to be seen between the intervention and control groups. Although the data is being compared six months before and six months after, the use of a matched control group will be essential to account for confounding factors that could occur over the six-month time period. For example, if the pre-intervention period occurred before the outbreak of a pandemic, the errors in the post-intervention period could look very different. Regardless if there would be significantly more or fewer errors, the matched control will allow the researchers to attribute differences in errors to the implementation of the I-PASS mnemonic and template.

Poisson Regression

The medical-error rates before and after the intervention will be compared through a Poisson regression with a dichotomous covariate for before versus after the intervention, and a fixed effect for the site. Poisson regression is used to predict a dependent variable that consists of 'count data' given one or more independent variables. Count data consists of data from a dependent variable that occurs in counts (0, 1, 2, 3…), and is discrete in that it consists of non-negative integers. For this experiment, the count data is the occurrences of each type of medical error, during the pre and post-intervention periods. By carrying out the Poisson Regression, it can be determined which of the independent variables, implementation of the I-PASS mnemonic and template, has a statistically significant effect on the dependent variable, medical errors. Given the independent variables are categorical, the percent increase or decrease in counts of one group (preventable adverse events), can be compared to another (near-miss events. Also, the total percent increase or decrease of medical errors can be calculated after I-PASS implementation.

Evaluation of Key Data Elements

Instances of medical error will be the main dependent variable being measured. The second aspect to be investigated is to compare the quality of written and oral handoffs given, pre and post-intervention. Each handoff will be evaluated for key data elements to determine whether they were sufficient in passing patient information. The key data elements will consist of five main aspects: I (illness severity), P (patient summary), A (action list), S (situation awareness and contingency planning), and S (synthesis by the receiver). The illness severity should include whether the patient is stable, or unstable, and what type of code the patient is. The patient summary will include age, events leading up to admission, hospital course, ongoing assessment, and plan. The action list will consist of specific things that need to be done for the patient, for example, if a patient needs their oxygen saturation level checked at a specific time. Situation awareness and contingency planning should include tests that the patient might need to be done.

Finally, synthesis by receiver should include the receiver repeating back the information they just learned, and asking questions if they have any. Each written and verbal handoff will be evaluated to determine if this information is present. It would be expected that there is an increase in the percentage of elements in the post-intervention handoffs compared to the pre-intervention handoffs. More specifically it would be expected that the increase would be more drastic in the written handoff category, as the individual using an I-PASS template can visually see if the information is missing. Overall the implementation of the I-PASS mnemonic and template will lead to a decrease in medical errors, and increase the percentages of key data elements found in each handoff.

The work in this experiment will build upon previous literature, like Starmer et al., 2014, that has attempted to propose handoff-improvement programs. Similarly to that study, this proposal uses a multicenter study approach in order to improve the generalizability. The pre and post-intervention data will be collected in the same year to control for time-of-year confounding factors. Further, this proposal will also include a matched control group that will control for these variables even more. The largest difference in this proposal is the inclusion of the I-PASS template for written handoffs. Previously the I-PASS mnemonic has been simplified to enhance the generalizability, implementation, and sustainability of this implementation. However, based on the literature on cognitive load theory, the use of multiple routes of information processing is beneficial to increase the capacity of the working memory. Promoting the use of auditory and visual methods for a handoff will help to minimize the occurrences of errors. Similarly, by making the handoff process universal, it will also help to limit errors by allowing healthcare workers to be on the same page when it comes to transferring patient information. One major solution that has been previously introduced by hospitals to reduce errors is limiting the hours worked during a shift. Although this might limit errors that could be attributed to being tired, it would increase the instances of handoffs, hence increasing the potential risk for errors. With this proposal that will potentially decrease the number of errors significantly, the implementation of a universal handoff program could be the better solution. Systematic change in hospitals is difficult to achieve, however by changing handoff practices as opposed to hours in a shift, or the number of healthcare workers in the hospital there is more potential for success.

General Discussion

Recent studies have accounted for over 250,000 deaths per year due to medical error, making it one of the leading causes of death in the United States (Anderson & Abrahamson, 2017). More specifically, around 25,000 of these deaths are due to preventable medical errors. Aside from the obvious concern for death, these errors also cost approximately $20 billion a year through insurance payouts (Rodziewicz, 2021). Many solutions have been proposed to limit these errors, like shortening shift hours or increasing the number of healthcare workers per hospital. However these solutions are not feasible in their capability to create the systemic change that is needed. By investigating what these errors are and why they occur, more achievable solutions can be potentially proposed. There are two main types of errors that can be applied to the healthcare field, are those of omission and of commission. Errors of omission are defined as resulting from actions not taken, for example forgetting to strap a patient into a wheelchair. Conversely, errors of commission are defined as resulting from the wrong action taken, for example administering a medication to a patient who has a known allergy (Rodziewicz, 2021). Some of the main causes of these errors that have been investigated are, long work hours, inadequate experience or supervision, and communication errors (Bari & Rathore, 2016). The focus of the present proposal is to investigate how interruptions can negatively affect the completion of Tower of London (ToL) tasks, and how to improve patient handoffs for preventing communication errors.

Previous Research

The first experiment of this proposal was created mainly based on a previous study conducted by Hodgetts and Jones (2006). The first study investigated the effects of interruptions on resumption times while completing the ToL task. Two experiments from that study will be replicated in this proposal while using nurses as the population. These experiments will investigate the time taken to make the fourth move in a trial of the ToL task, after interruptions. The first experiment is looking at the effects of the duration of interruptions. The length of interruption is predicted to have an effect on the time to return to the task, based on the memory model proposed by Anderson and Douglass (2001). They also investigated whether knowledge of upcoming interruptions would have an effect on preparation time for the task. Therefore, they conducted a 2 (interruption duration) x 2 (knowledge of interruption duration) repeated measures ANOVA and found no main effect of knowledge of interruption, a significant main effect of interruption length, and no interaction. The second experiment looks at the effect of interruption complexity through the lens of Anderson and Douglass's (2001) and Altmann and Trafton's (2002) memory models. Through a repeated measure ANOVA they found a significant effect of condition, meaning participants took less time to make the fourth move following a mood checklist compared to a reasoning task. They reported that the time taken to complete both tasks was dependent on the participants and no time limit was set. This was to relieve the pressure that might be created with a limit, which would cause an individual to rush the problem. Conversely setting a duration that is too long would result in a time period of rest, which would be considered low processing. Given this study is investigating the effects of complexity of the task, a period of low processing demand would hinder the results. However, the amount of time taken to complete each task cannot go unassessed, as the first experiment demonstrated a longer interruption has more of an effect than a shorter one. Therefore, Hodgetts and Jones (2006) analyzed the effect of complexity independent of interruption length. They found a slight trend for participants to spend more time on the reasoning task than on the mood checklist, however, the difference wasn't statistically significant. They conducted a Pearson's correlation to determine if time spent on the interruption correlated to the time spent on the two tasks, however, no significant correlation was found. Hence, they presumed that the difference in time to return to task between the two conditions can be attributed to the complexity of the two tasks, and not the time it took to complete them.

The Current Proposal

The current series of proposed experiments is designed to investigate the disruptive effects of interruptions on the ToL task and to propose a universal methodology for patient handoffs. In Experiments 1a and 1b, it is predicted that mid-task interruptions will generate a cost in terms of time taken to make the next move in a solution sequence, and will be exacerbated by longer and more complex interruptions. Generally in both of these parts, it would be expected that interruptions would produce little effect on the incidence of an error on the ToL task. An error on this task could be considered attempting to place four blocks on one peg or trying to move two blocks at once. Similarly, if a participant were to move a block to a position that wouldn't be in the pathway of the final product, this could also be considered an error. However given this task can have multiple pathways, and involves trial and error, it is hard to quantify these errors. This could be a reason why little errors would be expected to be observed. Another reason could have to do with the cost of the task, meaning if the task were completed incorrectly it would have little effect on the participant or people surrounding. When the cost of a task is lower, there is less pressure placed on the individual to perform well.

Despite the relevance of interruptions to both experimental psychology and human factor domains, empirical research has lacked coherence. The proposed studies here will aim to add more coherence to the literature while replicating findings from a previous study in a different group of participants. The use of the ToL task will be able to demonstrate the suitability of its use as a primary undertaking with interruption studies. It allows analysis of performance at each individual move, which therefore enables analysis of basic cognitive costs of interruptions uncontaminated by compensatory strategies. The level of precision in the task, defined by the ability of participants to be interrupted on individual moves, stands in contrast to previous interruptions studies where more general dependent variables have been used, like completion of comprehension tests (Oulasvirta & Saariluoma, 2004).

Practice Effects of Interruptions

In terms of the effects of the practice of interruptions, despite what was found by Hodgetts and Jones (2006), it would be expected that practice would positively affect how disruptive an interruption is. It would be expected that repeated exposure to an interruption would diminish the disruptive effects felt on the individual. Research that has examined the effects of repeated exposure to interruptions, has found support for this claim (Shinar et al., 2005; Compton, 2005). However, these studies do not specifically investigate the source of this improved performance. There are three specific sources that are further investigated in a three-part experiment conducted by Cades et al., (2011). The first form of improvement in performance could be a result of improvement on the primary task itself, which would ultimately be limiting cognitive load. Secondly, the improvement could also originate in the specific primary-interruption task pair, where exposure increases the ability to complete that specific task. And finally, there could be a general learning process where exposure to any type of interruption could lead to improvement in handling them.

Support for Improvement based on Practice with Primary Task

Empirical support supporting that improvement in performance following an interruption is based on practice with the primary task, highlights the use of Long-Term Working Memory (LTWM) (Oulasvirta & Saariluoma, 2004, 2006). They found that when participants were able to encode the primary task, reading an essay on a variety of topics, into LTWM, they scored higher on comprehension tests after interruptions, compared to those who were not able to use LTWM. This work demonstrated that as long as participants were able to encode the information of the primary task, they didn't suffer the negative effects of being interrupted. The first part of the experiment conducted by Cades et al., (2011), found support for this theory in that performance on the primary task improves with training, however only when the individual has had previous training with that specific task and interruptions. Support for these findings can further be mentioned through a previously mentioned study looking at the Memory for Goals model (Altmann & Trafton, 2002). The cognitive architecture suggests that the most active current goal drives behavior, and activation is a measure of the strength of a goal in memory. Once a goal is in memory and a primary task is suspended, the activation level associated with the goal decays. Then when the interruption is complete, the goal associated with the primary task must be retrieved from memory. The time it takes to retrieve this goal, and the probability that it will be correct, are directly related to the activation of the goal at the time of retrieval.

Support for Improvement Based on Practice with Interrupting Task Pairs

When a task is interrupted it bears similarity to the performance of multiple tasks at once, or multitasking. The main difference between completion of a set of tasks, and interrupted task performance, is that completion of one task has priority over the other task. Hence when a primary task is interrupted by a second task, the performer desires to return to the primary task as soon as possible after being interrupted (Wickens, 2008). Therefore, interrupting task performance can be viewed as the performance of a larger, multipart task with smaller individual tasks. By looking at interruptions through this lens, evidence for improvements in task pair training can be investigated. The empirical evidence for this concept is minimal, however in support. For example, one study found that training concentrated on important components of part-tasks found higher performance levels on a whole task during a flight simulator (Goettl & Shute, 1996, as cited by Cades et al., 2011). The same memory model proposed by Althamn and Trafton, (2007), can be applied to this. The model would suggest that people's performance on sequential tasks would improve only when they perform those tasks in the same sequence over time. This demonstrates that interrupted task performance lies in the transition between the primary and interrupting task, and the practice of specific task pairs would improve resumption rates (Cades et al., 2011).

Support for Improvement Based on General Resumption Processes

The final explanation for how people improve at dealing with interruptions over time, is simply by learning how to recover from interruptions in general. This would mean that the specific task doesn't matter, and what does is the actual experience of being interrupted. However, the literature to support or refute this idea is nonexistent. In order for this theory to be successful, it would mean that there are more general goals associated with the interruption and resumption processes. If the goals are more general, then when an individual is interrupted over and over, then these constraints would be strengthened, and individuals would be able to resume more quickly. Hence this explanation would suggest that exposure to the process of interruptions and resumptions would minimize the negative effects felt by interruptions. The third experiment conducted by Caded et al., (2011) ruled out this general-resumption-process view, further confirming that improvements are task-pair-specific.

Despite all the empirical research that supports the concept that the practice of interruptions decreases the negative effect felt by participants, Hodgetts and Jones (2006), found no effect of practice in their experiments. They acknowledge that this issue would require further investigation with more control of manipulation. Specifically by using participants who are well-practiced at the Tol task, in order to disambiguate the effects of practice recovering from interruptions and the effects of ToL expertise. Given Experiment 1a/b of this proposal will be replicating the methodology from Hodgett and Jones (2006), it will be curious to see if the effects of the practice are the same as the original study, or if they will follow what is supported in empirical research.

The Goal Memory Model

The models of Anderson and Douglass (2001) and Altmann and Trafton (2002) can be applied to Experiment 1a/b, to better predict what would be expected. The Anderson and Douglass (2001) model emphasized that memory activation decreases solely as a time-based function. This means that goals are retrieved in a last-in, first-out manner, making the presence of distracting goals ineffective in goal retrieval. This model can specifically be applied to Experiment 1a which investigated the effects of duration of interruption on resumption times on the ToL task. Based on this model it would be expected that longer interruptions would cause longer resumption times, compared to shorter interruptions. The Altmann and Trafton (2002) model claims that goal selection is based on goal activation. Therefore, during an interruption, if goals are activated, then the interference level would be higher, making interference higher for the target goal. This is just another theory that goals are stored in declarative memory, and the more declarative information activated during an interruption the more interference there is. Both these models relate to time, hence why they can be applied to better understand the predictions of the first study. However, interruptions are heterogeneous in that time isn't the only dimension in which they can vary, hence why these models can only be applied up to a certain point. Therefore, further investigation of these models is necessary to be able to better apply them to factors such as interruption complexity. Given these models aren't able to be applied to Experiment 1b to better predict the results, it is important to look at this experiment through the lens of memory. It would be expected that complex interruptions would have a larger effect on resumption times, compared to simple interruptions, due to suppression of rehearsal. In order for a suspended goal to be contained within long-term memory it must be rehearsed, to move from that what makes up the working memory capacity. Therefore in this proposal, a reasoning task is assumed to be more effortful than a mood checklist and would cause more decay in goal attainment.

Although task interruptions are relevant in experimental psychology, empirical research on this topic has been relatively scarce. Especially when relating interruptions to the healthcare field, much of the research has consisted merely of observational studies. This lack of manipulation leads to the inability to establish causation, however, these observations have at least allowed researchers to better understand just how often healthcare workers are being interrupted. Chisholm et al. (2000) found that every 180 minutes a clinician was interrupted on average 31 times. Another study by Monteiro et al., (2015) found that nurses were interrupted between 1 and 13 times an hour. Again, these findings are beneficial to understanding just how often healthcare workers are being interrupted; however, it still begs the question: what effect are these interruptions having on performance? Experiment 1a/b of this proposal will aim to answer this question in the nursing population. Based on the aforementioned reasons, it would be predicted that there is a main effect of duration and complexity of interruptions on resumption times on the ToL task. Once there is a better understanding of how interruptions can affect task completion in nurses, further investigations can be made on the errors made from such interruptions. More specifically, given interruptions are a major way to increase the extraneous load, the overall cognitive load of the individual must be taken into account.

Cognitive Load Theory

Cognitive load theory (CLT), is a model of cognitive architecture used to better understand how human memory works (Van Merrienboer & Sweller, 2010). The main idea is constructed from the assumption that a limited working memory (WM), and unlimited long-term memory (LMT) comes from knowledge stores as schemas in LTM. Within human memory, the capacity for it to function and recall items will depend on an individual's cognitive load. The three types of cognitive load are as follows: intrinsic (load associated with the task), extraneous (load not essential to the task), and germane (load imposed by the learner’s use of cognitive strategies to organize information for storage in long term memory).

An extraneous load of an individual is a factor that diverts the attention away from the learning environment (Wilby & Paravattil, 2021). An example of this would be an interruption or any distractions that prevent the individual from being able to complete the task at hand. One recent publication found that distractions have been shown to play a role in nearly 75% of medical errors, and cognitive overload is a cause of 80% of medical errors (Tariq et al., 2021). Although the extraneous load is a major contributor to cognitive load, minimizing it isn't the only solution to reducing healthcare errors. Regardless of the harmless effects of these interruptions, they are a necessary aspect of the healthcare field to ensure patient care. Therefore, claiming that an extraneous load of healthcare workers is the most reasonable way to limit errors, isn't the main solution. This would require systematic change that is unreasonable to expect each hospital to meet. Hence why it is important to investigate the other types of cognitive load, like intrinsic load, and ways to better manage that.

The intrinsic load of an individual has to do specifically with the complexity of the task, and the learner's expertise in that task. It also depends on the interactivity of the elements while learning (Wilby & Paravattil, 2021). Empirical research has demonstrated that presenting isolated elements before demonstrating how they interact, will increase the ability of a learner to understand (Pollock et al., 2002). Another support for managing intrinsic load has been found in the ability of templates to make information more accessible (Young et al., 2016). This concept gave rise to the second proposed experiment in this proposal. By looking at vulnerable places in the healthcare field for errors, and applying CLT to it, a better understanding can be had of how to potentially minimize these errors.

Patient handoffs can greatly be considered one of the most vulnerable areas in the healthcare field for errors. These handoffs occur at such a great occurrence that it increases the chances of errors. According to the Joint Commission, communication gaps within a handoff are estimated to contribute to 80% of medical errors (The Joint Commission, 2022). Although there are a number of reasons for these errors, the one most prominent to this proposal is the cognitive load placed on the individual. When the cognitive load of an individual is increased, it exceeds the capabilities of the working memory capacity, in turn leading to more potential errors. Experiment 2 of this proposal was created as a way to manage the intrinsic load of healthcare workers. This is proposed by universalizing patient handoffs and promoting handoffs through both auditory and visual pathways in order to maximize the working memory capacity of an individual.

The use of the I-PASS mnemonic has previously been demonstrated to be successful in minimizing errors in the healthcare field by 23% (Starmer et al., 2014). This methodology has investigated the use of this mnemonic only in pediatric hospitals, and as a tool for verbal handoffs. This proposal intends to replicate some of this methodology in 10 teaching hospitals, of all specialties, to better understand its generalizability. The main difference in this proposal is the use of the I-PASS mnemonic coupled with the written template. By universalizing patient handoffs, and promoting the use of visual and auditory pathways, success in minimizing healthcare errors is predicted. Evaluation of written and oral handoffs will also serve as a way to measure healthcare errors. After the implementation of the handoff program, it would be expected that both oral and written handoffs would improve in their quality. By comparing the quality of these handoffs before and after the implementation, and finding that they were improved along with healthcare errors, then the evidence would be had that this decrease in errors would be a result of the implementation.

Limitations

The present proposal builds upon two already published research articles (Hodgetts & Jones, 2006; Starmer et al., 2014). Not only will this replication serve to see if the results are generalizable, but will build on certain limitations within the first studies. Experiments 1a/b are a replication of methodologies from Hodgetts and Jones (2006). The first major limitation of the original study has to do with the participants. They selected 20 undergraduate students to participate in their study. Although this is a convenient method to recruit individuals, there have been many questions about the generalizability of studies that use undergraduate students. This is because students are usually considered more homogeneous than representative samples both within and across countries (Hanel & Vione, 2016). However, some studies have attempted to find data to go against this theory, as most psychological studies rely on student samples. In one study that investigated differences between student samples, and representative samples on 12 personality and attitudinal variables, found that their findings contradicted previous ones. They found that students vary as much as the general population both between and within countries. However, the only time they found sampling students to be problematic is when personal and attitudinal variables are used, because that’s when students vary the most from the general public (Hanel & Vione, 2016). Based on these findings it would be reasonable to think that the results from Hodgetts and Jones (2006), could potentially be generalized to a larger population. Regardless of this, the sample size of the study still includes 20 students which ultimately reduces the power of the study and increases the margin of error. This proposal was created and changed based on both the population and sample size of the original study. The desire is to see if the same results will replicate among nurses, and using a larger sample size will increase the power of the study, hence making the results more significant.

Although the ToL task has proven beneficial as a tool to enable the analysis of basic cognitive costs of interruptions uncontaminated by compensatory strategies, there are some limitations when looking at a population of nurses. Although the main purpose of this proposal isn’t errors on the task, when looking at resumption times, these could be altered based on the basicness of the task. The inherent low cost of the task makes it so there is no pressure to complete the task correctly. When attempting to generalize findings to all nurses it would be beneficial to conduct another study, where the task has a higher cost. An example of this could be conducting a study that investigates interruptions during a med pass, in simulation. Medicine passes are when nurses are giving their patients the medications they need at that moment, and by manipulating interruptions during this time, a number of errors could be collected. Given this is in simulation, it doesn't make this a life or death matter (which would be considered unethical through IRB) however the negative effects can still be felt. If a medication is given in the wrong way, or the wrong dose, the simulation will still demonstrate what would have happened to a real patient. This would be an example of a task with high cost, and could potentially have an effect on the number of errors found. Given within the healthcare field completing a task incorrectly can have negative consequences, using a task with a higher cost would help make results more valid. However, for the purpose of this study, the use of the ToL task will be beneficial in collecting preliminary data on the effects of interruptions within a population of nurses.

The final limitation relating to Experiments 1a/b in this proposal has to do with the statistical analyses. These analyses were selected based on what was done in the Hodgett and Jones (2006) methodology. The first part of the experiment is conducting a 2 (interruption duration) x 2 (interruption knowledge) repeated measures ANOVA. This is the same analysis used in this proposal, however there could be inclusions of one other variable within the analysis. Before the start of the study the participants were randomly assigned to groups based on the order of the interruption. As mentioned in the methodology, half the participants will receive the short interruption first, and the other half will receive the long interruption first. This would add another level to the analysis making it a 2 (interruption duration) x 2 (interruption knowledge) x 2 (interruption order) repeated measures ANOVA. Given the sample size was so small in the

Hodgetts and Jones (2006) study, I would predict that they didn't include order in the analysis because they sorted the data by hand. With a small sample size it would be possible to look at resumptions rates and compare them to the order in which they occured. This also would be found in the second part of the experiment. The analysis included a one way ANOVA looking at the two different complexities of interruptions, and comparing it to resumption rates. However if order were to be included in this analysis, it would make it a 2 (interruption complexity) x 2 (interruption order) repeated measures ANOVA. Again it is possible that since the sample size was so small, that the researchers in the Hodgetts and Jones (2006) study simply sorted the data by hand. For the purpose of this proposal the analyses will be the same that were used in the original study. However for future proposals, that are using a larger sample size, it would be beneficial to include order in the analysis so they don't have to be conducted by hand.

The proposed methodology for Experiment 2 was created to address several limitations of previous studies investigating the I-PASS mnemonic. The first is the use of multicenter studies to improve generalizability. Although the Stramer et al., (2014) study used multiple study centers, they used pediatric residency hospitals. By including residency hospitals of different types (i.e., neurology, cardiology, internal medicine, etc), the proposed work will improve the overall generalizability of the I-PASS handoff system. Second, the data will be collected 6 months pre and post-intervention, within the same year to prevent time-of-year confounding factors. To further limit the effects of these potential confounds, a matched control group will be used at each stage. This will allow the data to be further compared not only between pre and post-intervention but within each condition. Next, the I-PASS mnemonic will be implemented along with the written template. This will enhance the sustainability of this intervention, and, even more importantly, promote the use of multiple sensory channels, expanding the capabilities of the working memory capacity. One final limitation to this proposal is the use of recording devices to assess verbal handoffs. If a participant knows they are being recorded then they are more likely to complete a good handoff. It would be unethical to record the participants without their knowledge, therefore another solution to this problem needed to be proposed. Evaluating the oral and written handoffs based on specific data elements, for both pre and post-intervention, could combat some of the effects. Given the participants don't know what these specific data elements are, having them know they are being recorded might not have a huge effect on the results. By furthering the research on this topic, and improving the methodology to increase the generalizability and sustainability of the I-PASS technique, this proposal is predicted to have great beneficial results in decreasing accounts of healthcare errors.

Mental Health Effects

Although much of the literature has focused on the negative effects of these errors on patients, there is also something to be said about its impact on healthcare workers. Recently a light has been shone on the overall emotional impact of these medical errors. There is much pressure on physicians specifically to continue work without processing emotions made from errors. This is emphasized in a paper by Goldberg et al., (2002), where they quoted Hilfiker, 1984:

> The drastic consequences of our mistakes, the repeated opportunities to make them, the uncertainty about our own culpability when results are poor, and the medical and societal denial that mistakes must happen all result in an intolerable paradox for the physician. We see the horror of our own mistakes, yet we are given no permission to deal with their enormous emotional impacts... . The medical profession simply has no place for its mistakes (p. 289).

Due to human imperfections, medical errors are inevitable, and the impact on healthcare workers is less acknowledged. Hospitals tend to react to errors as an anomaly, in which the solution is to blame the individual, and promise that the mistake will never happen again (Wu, 2000). This approach, unfortunately, diverts attention from systematic improvements that could decrease errors. It goes without saying that patients are the primary victims of medical error, but few people accept that there can be a second victim too, the healthcare professional.

It has been demonstrated in papers that unintentional error can have a lasting impact on healthcare workers causing depression, burnout, lack of concentration, decreased clinical confidence, and impaired work performance (Delbanco & Bell, 2007). One of the main culprits of the lack of preparation for the potential mental effects starts with medical education. Medical school trains individuals in the skills needed to become a physician, however, it doesn't prepare students for dealing with errors (Smith, 2000). This lack of acknowledgment promotes the idea that the healthcare system doesn't tolerate mistakes. Even more, medical education in the past has taught physicians that they are the primary decision-makers, rather than being a part of the healthcare teams. Hence when an error occurs the more natural reaction is to place blame on the individual, instead of improvement in knowledge and skills. Over time, this culture has led to many mental health issues among healthcare workers. A review conducted by Robertson and Long (2018) reported the findings of a 2007 survey of 3,171 physicians that investigated the effects of errors on work and life domains. The results of the survey demonstrated that almost two-thirds of the physicians reported increased anxiety about future errors. Half the clinicians reported a loss of confidence, sleeping difficulties, and reduced job satisfaction. More than 80% of participants expressed interest in counseling after a serious error, but only 10% agreed that their health care organizations adequately supported them in coping with error-related emotional distress (Waterman et al., 2007).

Similar levels of distress were observed in residents from medical errors. West et al., (2009) evaluated burnout and emotional distress in residents in 2006 and 2009. In 2006, the authors found that in 184 internal medicine residents, overall errors were associated with decreased quality of life, increase in burnout symptoms, and increased screening positive for depression (West et al., 2006). They found similar results in 2009 where increased fatigue, burnout, and depressive symptoms were found in residents who reported medical errors during the study period (West et al., 2006). Although many of these studies have looked at the negative effects of errors on clinicians, these negative effects are felt by all healthcare workers. The culture of the healthcare system today preaches the need to be perfect and neglects potential mental health issues resulting from the unintentional error. There is not only a need for systemic change to attempt to limit health care errors to protect the primary victim, patients, however, there is also a need for systemic change in how these errors are handled to help the secondary victim, healthcare workers.

Preventable healthcare errors continue to rise in occurrence and lead to immense death per year. Previous literature has investigated why these errors are occurring and has attempted to come up with solutions. The lack of empirical evidence for improving healthcare practices has given rise to this proposal. With medicine constantly evolving, it is necessary to advance practices specifically relating to patient handoffs. The United States currently has over 3.8 million registered nurses nationwide (*Nurses in the Workforce*, 2022). With all these nurses conducting patient handoffs in their own unique way, it increases the chances of errors drastically. By universalizing the patient handoff system, and promoting the use of auditory and visual pathways to improve working memory capacity, there is a chance to minimize the preventable errors that are killing so many each year.

References

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science, 26*(1), 39–83. https://doi.org/10.1016/S0364-0213(01)00058-1

Anderson, J. G., & Abrahamson, K. (2017). Your Health Care May Kill You: Medical Errors. *Studies in health technology and informatics*, *234*, 13–17.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence, *The psychology of learning and motivation: II.* Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Brydges, R., Carnahan, H., Rose, D., Rose, L., & Dubrowski, A. (2010). Coordinating progressive levels of simulation fidelity to maximize educational benefit. *Academic medicine: journal of the Association of American Medical Colleges*, *85*(5), 806–812. https://doi.org/10.1097/ACM.0b013e3181d7aabd

Cades, D. M., Boehm-Davis, D. A., Trafton, J. G., & Monk, C. A. (2011). Mitigating disruptive effects of interruptions through training: what needs to be practiced?. *Journal of experimental psychology. Applied*, *17*(2), 97–109. https://doi.org/10.1037/a0023497

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

Chisholm, C. D., Collison, E. K., Nelson, D. R., & Cordell, W. H. (2000). Emergency department workplace interruptions are emergency physicians' "interrupt-driven" and "multitasking"?. *Academic emergency medicine: official journal of the Society for Academic Emergency Medicine*, *7*(11), 1239–1243. https://doi.org/10.1111/j.1553-2712.2000.tb00469.x

Coiera, E. W., Jayasuriya, R. A., Hardy, J., Bannan, A., & Thorpe, M. E. (2002). Communication loads on clinical staff in the emergency department. *The Medical Journal of Australia*, *176*(9), 415–418. https://doi.org/10.5694/j.1326-5377.2002.tb04482.x

Delbanco, T., & Bell, S. K. (2007). Guilty, afraid, and alone--struggling with medical error. *The New England journal of medicine*, *357*(17), 1682–1683. https://doi.org/10.1056/NEJMp078104

Douglas, H. E., Raban, M. Z., Walter, S. R., & Westbrook, J. I. (2017). Improving our understanding of multi-tasking in healthcare: Drawing together the cognitive psychology and healthcare literature. *Applied Ergonomics*, *59*, 45–55. https://doi.org/10.1016/j.apergo.2016.08.021

Facts about the Joint Commission. The Joint Commission. (n.d.). Retrieved March 6, 2022, from https://www.jointcommission.org/about-us/facts-about-the-joint-commission/

Fraser, K. L., Ayres, P., & Sweller, J. (2015). Cognitive Load Theory for the Design of Medical Simulations. *Simulation in healthcare: journal of the Society for Simulation in Healthcare*, *10*(5), 295–307. https://doi.org/10.1097/SIH.0000000000000097

Goldberg, R. M., Kuhn, G., Andrew, L. B., & Thomas, H. A., Jr (2002). Coping with medical mistakes and errors in judgment. *Annals of emergency medicine*, *39*(3), 287–292. https://doi.org/10.1067/mem.2002.121995

Hanel, P. H., & Vione, K. C. (2016). Do Student Samples Provide an Accurate Estimate of the General Public?. *PloS one*, *11*(12), e0168354. https://doi.org/10.1371/journal.pone.0168354

Heng K. W. (2014). Teaching and evaluating multitasking ability in emergency medicine residents - what is the best practice?. *International journal of emergency medicine*, *7*, 41. https://doi.org/10.1186/s12245-014-0041-4

Hilfiker D. (1984). Facing our mistakes. *The New England journal of medicine*, *310*(2), 118–122. https://doi.org/10.1056/NEJM198401123100211

Hodgetts, H. M., & Jones, D. M. (2006). Interruption of the Tower of London task: support for a goal-activation approach. *Journal of experimental psychology. General*, *135*(1), 103–115. https://doi.org/10.1037/0096-3445.135.1.103

Issa N, Schuller M, Santacaterina S, Shapiro M, Wang E, Mayer RE, Darosa DA. (2011). Applying multimedia design principles enhances learning in medical education. Med Educ 45:818–826.

James J. T. (2013). A new, evidence-based estimate of patient harms associated with hospital care. *Journal of Patient Safety*, *9*(3), 122–128. https://doi.org/10.1097/PTS.0b013e3182948a69

Kaushal R. (2002). Using chart review to screen for medication errors and adverse drug events. *American journal of health-system pharmacy: AJHP: official journal of the American Society of Health-System Pharmacists*, *59*(23), 2323–2325. https://doi.org/10.1093/ajhp/59.23.2323

Kirschner PA, Sweller J, Clark RE. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educ Psychol 41:75–86.

Laxmisan, A., Hakimzada, F., Sayan, O. R., Green, R. A., Zhang, J., & Patel, V. L. (2007). The multitasking clinician: decision-making and cognitive demand during and after team handoffs in emergency care. *International journal of medical informatics*, *76*(11-12), 801–811. https://doi.org/10.1016/j.ijmedinf.2006.09.019

Malmberg, K. J., Raaijmakers, J., & Shiffrin, R. M. (2019). 50 years of research sparked by Atkinson and Shiffrin (1968). *Memory & Cognition*, *47*(4), 561–574. https://doi.org/10.3758/s13421-019-00896-7

Mayer RE. (2010). Applying the science of learning to medical education. Med Educ 44:543–549

Miller G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

Monteiro, C., Avelar, A. F., & Pedreira, M. (2015). Interruptions of nurses' activities and patient safety: an integrative literature review. *Revista latino-americana de enfermagem*, *23*(1), 169–179. https://doi.org/10.1590/0104-1169.0251.2539

Mousavi SY, Low R, Sweller J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. J Educ Psychol 87:319–334

Nurses in the Workforce (2022). https://www.nursingworld.org/practice-policy/workforce/ American Nurses Association. Retrieved April 22, 2022, from https://www.nursingworld.org/practice-policy/workforce/

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, *25*(1), 46–59. https://doi.org/10.1002/hbm.20131

Patterson, E. S., Roth, E. M., Woods, D. D., Chow, R., & Gomes, J. O. (2004). Handoff strategies in settings with high consequences for failure: lessons for health care operations. *International journal for quality in health care: journal of the International Society for Quality in Health Care*, *16*(2), 125–132. https://doi.org/10.1093/intqhc/mzh026

Perina, D. G., Brunett, C. P., Caro, D. A., Char, D. M., Chisholm, C. D., Counselman, F. L., Heidt, J., Keim, S. M., & Ma, O. J. (2012). The 2011 model of the clinical practice of emergency medicine. *Academic emergency medicine: official journal of the Society for Academic Emergency Medicine*, *19*(7), e19–e40. https://doi.org/10.1111/j.1553-2712.2012.01385.x

Phillips, L. H., Wynn, V. E., McPherson, S., & Gilhooly, K. J. (2001). Mental planning and the Tower of London task. *The Quarterly Journal of experimental psychology. Human experimental psychology*, *54*(2), 579–597. https://doi.org/10.1080/713755977

Phillips, L. H., Lawrie, L., Schaefer, A., Tan, C. Y., & Yong, M. H. (2021). The Effects of Adult Aging and Culture on the Tower of London Task. *Frontiers in psychology*, *12*, 631458. https://doi.org/10.3389/fpsyg.2021.631458

Pollock, E., Chandler, P., & Sweller, J. (2001). *Assimilating complex information*. Learning and Instruction. Retrieved April 28, 2022, from https://www.sciencedirect.com/science/article/pii/S0959475201000160

Radović, T., & Manzey, D. (2022). Effects of complexity and similarity of an interruption task on resilience toward interruptions in a procedural task with sequential constraints. *Journal of experimental psychology. Human perception and performance*, *48*(2), 159–173. https://doi.org/10.1037/xhp0000981

Raeisi, A., Rarani, M. A., & Soltani, F. (2019). Challenges of patient handover process in healthcare services: A systematic review. *Journal of education and health promotion*, *8*, 173. https://doi.org/10.4103/jehp.jehp_460_18

Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly journal of experimental psychology (2006)*, *69*(10), 1969–1995. https://doi.org/10.1080/17470218.2014.914546

Robertson, J. J., & Long, B. (2018). Suffering in Silence: Medical Error and its Impact on Health Care Providers. *The Journal of emergency medicine*, *54*(4), 402–409. https://doi.org/10.1016/j.jemermed.2017.12.001

Rodziewicz, T. L., Houseman, B., & Hipskind, J. E. (2021). Medical Error Reduction and Prevention. In *StatPearls [Internet]*. StatPearls Publishing.

Sectish, T. C., Starmer, A. J., Landrigan, C. P., Spector, N. D., & I-PASS Study Group (2010). Establishing a multisite education and research project requires leadership, expertise, collaboration, and an important aim. *Pediatrics*, *126*(4), 619–622. https://doi.org/10.1542/peds.2010-1793

Shanafelt, T. D. (2006). Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA*, *296*(9), 1071–1078. https://doi.org/10.1001/jama.296.9.1071

Smith, M. L., & Forster, H. P. (2000). Morally managing medical mistakes. *Cambridge quarterly of healthcare ethics: CQ: the international journal of healthcare ethics committees*, *9*(1), 38–53. https://doi.org/10.1017/s0963180100901051

Standard-faqs. The Joint Commission. (n.d.). Retrieved January 5, 2022, from https://www.jointcommission.org/standards/standard-faqs/?utm_medium=optimize&utm_campaign=tjc-home-connect ref=tjchc56

Starmer, A. J., Landrigan, C. P., & I-PASS Study Group (2015). Changes in medical errors with a handoff program. *The New England journal of medicine*, *372*(5), 490–491. https://doi.org/10.1056/NEJMc1414788

Starmer, A. J., O'Toole, J. K., Rosenbluth, G., Calaman, S., Balmer, D., West, D. C., Bale, J. F., Jr, Yu, C. E., Noble, E. L., Tse, L. L., Srivastava, R., Landrigan, C. P., Sectish, T. C., Spector, N. D., & I-PASS Study Education Executive Committee (2014). Development, implementation, and dissemination of the I-PASS handoff curriculum: A multisite educational intervention to improve patient handoffs. *Academic medicine: journal of the Association of American Medical Colleges*, *89*(6), 876–884. https://doi.org/10.1097/ACM.0000000000000264

Stephens, R. J., & Fairbanks, R. J. (2012). Humans and multitask performance: let's give credit where credit is due. *Academic emergency medicine: official journal of the Society for Academic Emergency Medicine*, *19*(2), 232–234. https://doi.org/10.1111/j.1553-2712.2011.01264.x

Shinar, D., Tractinsky, N., & Compton, R. (2005). Effects of practice, age, and task demands, on interference from a phone task while driving. *Accident; analysis and prevention*, *37*(2), 315–326. https://doi.org/10.1016/j.aap.2004.09.007

Tanne J. H. (2008). US hospitals pass on most of the costs of errors. *BMJ (Clinical research ed.)*, *336*(7649), 852. https://doi.org/10.1136/bmj.39551.680417.C2

Tariq, R. A., Vashisht, R., Sinha, A., & Scherbak, Y. (2021). Medication Dispensing Errors And Prevention. In *StatPearls*. StatPearls Publishing.

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(1), 37–55. https://doi.org/10.1037/xlm0000578

Tiwary, A., Rimal, A., Paudyal, B., Sigdel, K. R., & Basnyat, B. (2019). Poor communication by health care professionals may lead to life-threatening complications: examples from two case reports. *Wellcome open research*, *4*, 7. https://doi.org/10.12688/wellcomeopenres.15042.1

Van Merriënboer, J. J., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical education*, *44*(1), 85–93. https://doi.org/10.1111/j.1365-2923.2009.03498.x

Villines, Z. (2020). *Types of memory: Learn more about the different types of memory here*. Medical News Today. Retrieved April 6, 2022, from https://www.medicalnewstoday.com/articles/types-of-memory

Waterman, A. D., Garbutt, J., Hazel, E., Dunagan, W. C., Levinson, W., Fraser, V. J., &

Gallagher, T. H. (2007). The emotional impact of medical errors on practicing physicians in the United States and Canada. *Joint Commission journal on quality and patient safety*, *33*(8), 467–476. https://doi.org/10.1016/s1553-7250(07)33050-x

West, C. P., Huschka, M. M., Novotny, P. J., Sloan, J. A., Kolars, J. C., Habermann, T. M., & Shanafelt, T. D. (2009). Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA*, *296*(9), 1071–1078. https://doi.org/10.1001/jama.296.9.1071

Westbrook, J. I., Coiera, E., Dunsmuir, W. T., Brown, B. M., Kelk, N., Paoloni, R., & Tran, C. (2010). The impact of interruptions on clinical task completion. *Quality & safety in health care*, *19*(4), 284–289. https://doi.org/10.1136/qshc.2009.039255

Wickens C. D. (2008). Multiple resources and mental workload. *Human factors*, *50*(3), 449–455. https://doi.org/10.1518/001872008X288394

Wilby, K. J., & Paravattil, B. (2021). Cognitive load theory: Implications for assessment in pharmacy education. *Research in social & administrative pharmacy: RSAP*, *17*(9), 1645–1649. https://doi.org/10.1016/j.sapharm.2020.12.009

Wu A. W. (2000). Medical error: the second victim. The doctor who makes the mistake needs help too. *BMJ (Clinical research ed.)*, *320*(7237), 726–727. https://doi.org/10.1136/bmj.320.7237.726

Wolfe, D., Yazdi, F., Kanji, S., Burry, L., Beck, A., Butler, C., Esmaeilisaraji, L., Hamel, C., Hersi, M., Skidmore, B., Moher, D., & Hutton, B. (2018). Incidence, causes, and consequences of preventable adverse drug reactions occurring in inpatients: A systematic review of systematic reviews. *PloS one*, *13*(10), e0205426. https://doi.org/10.1371/journal.pone.0205426

Young, J. Q., Ten Cate, O., O'Sullivan, P. S., & Irby, D. M. (2016). Unpacking the Complexity of Patient Handoffs Through the Lens of Cognitive Load Theory. *Teaching and learning in medicine*, *28*(1), 88–96. https://doi.org/10.1080/10401334.2015.1107491

Appendices

Appendix A: Pre Registration

Appendix B: Consent Form Experiment One

Appendix C: Consent Form Experiment Two

Appendix D: Budget Experiment 1a

Appendix E: Budget Experiment 1b

Appendix F: Budget Experiment 2

Appendix A
Preregistration

Experiment 1a/b

1) Has any data been collected for this study already?

This study is a proposal for a Senior Project at Bard College.

2) What's the main question being asked or the hypothesis being tested in this study?

What are the effects of interruptions of time to make the fourth move on the Tower of London task (ToL)? This has been previously studied by Hodgetts and Jones (2006), who found both a significant main effect of duration and complexity of interruptions on the resumption time of the ToL task. It is hypothesized based on the goal memory model of Anderson and Douglass (2001) and Altmann and Trafton (2002), that there will be a main effect of duration and complexity of interruption on the time to make the fourth move on the ToL task.

3) Describe the key dependent variable(s) specifying how they will be measured.

The dependent variable in this experiment is the time taken to make the fourth move on the ToL task. This time will be measured in both the manipulation and control conditions. For the manipulation conditions, the time will be recorded on the computers from the moment the button is pressed to resume the primary task, so when the move is made. In the control condition, the time will be recorded from the end of the third move to the end of the fourth move. Throughout this experiment, the participants will be required to complete 25 trials. In Experiment 1a participants will be interrupted 8 times (Trials 4, 7, 10, 14 & 15, 19, 21, 25) and in Experiment 1b participants will be interrupted 6 times (Trials 4, 7, 12 & 15, 19, 25).

4) How many and which conditions will participants be assigned to?

The first part of experiment 1 will consist of 4 conditions and will be analyzed using a 2 (long vs short interruption) x 2 (informed about upcoming interruptions vs not informed) repeated measures design. Participants will be randomly assigned to any of these 4 conditions along with a matched control group. The short interruptions will consist of a 6-second mood checklist, whereas the long interruption will be three consecutive mood checklists accounting for 18 seconds. All participants are interrupted on trials 4, 7, 10, 14, & 15, 19, 21, 25, with matched control trials. The first four interruptions will be the same and the last four will also be the same. Therefore the four conditions are as follows: short interruption first and informed about the upcoming interruptions, short interruption first and not informed about the upcoming interruption, long interruptions first and informed about the upcoming interruptions, and finally long interruptions first and not informed about the upcoming interruption. Therefore this experiment will be conducted in a within-subject design. It was determined by Hodgetts and Jones (2006) that the difference in interruptions was long enough to produce the effects of length, but not long enough for goal decay. The second part of the study will also be conducted using a within-subject design. It differs from the first part of the study in that participants are interrupted a total of 6 times, and the interruptions have to do with complexity, not duration. Therefore there are two possible conditions for this part of the experiment: complex interruption first, or simple interruption first. The simple interruption will consist of a mood checklist similar to that in the first part of the experiment. The complex interruption will consist of a reasoning task, which had been previously deemed as more complex by Hodgetts and Jones (2006). Similar to the first part of the experiment, there will be matched control trials within the 25 trials that allow for a within-subject design. For both parts of the experiment, the match controlled trials will include the same configuration of the ToL task, in its corresponding manipulation trial, just with different colored blocks.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

A two-way Analysis of Variance (ANOVA) will be used to interpret the data of the first part of the experiment. This will compare the mean differences between the two independent variables: duration of the interruption, and whether or not the participant was informed about the interruption beforehand, and compare them to the dependent variable, time to make the fourth move. The second part of the study will also use repeated measures ANOVA to determine the relationship between the independent variables, complex vs simple interruptions, and dependent variable, time to make the fourth move.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Participants who don't complete all the trials will be excluded from the analysis. This will help to ensure the validity of the data. Hodgetts and Jones (2006) predicted that each of the "filler" trials could be completed within 4-6 moves. If a participant was taking upwards of 20 moves for those trials then their data might also be excluded. This is necessary because if the participant cannot understand how to complete the trials, then it could cause outliers in the data.

7) How many observations will be collected or what will determine the sample size?.

In order to determine the sample size, a power analysis will be conducted. This will ensure that there are enough observations made to conclude the results as significant. Lastly, this will allow for the results to be more generalizable.

8) Anything else you would like to pre-register?

This study will serve as the first of two study proposals. The predicted results that interruptions will have a negative effect on the ability to efficiently complete the Tower of London task, will guide a second proposed study about cognitive load theory during patient handoffs.

Experiment 2

1) Has any data been collected for this study already?

This study is a proposal for a Senior Project at Bard College.

2) What's the main question being asked or the hypothesis being tested in this study?

Miscommunication is one of the leading causes of preventable medical errors. Patient handoffs are one of the most vulnerable areas for these communication errors to occur. Thus far, there is no standard methodology to complete a handoff. Recent literature has highlighted investigations on a new universal way to complete handoffs called the I-PASS method (Starmer et al., 2014). Research to date has found success in minimizing errors in pediatric teaching hospitals. Therefore further investigation using the I-PASS method is needed in non-pediatric hospitals. This study proposes not only using the I-PASS method in non-pediatric hospitals but the adaptation of a physical I-PASS template that can be used during handoffs. The main hypothesis is that there will be a decrease in medical errors observed 6 months post-intervention, compared to the data collected in the 6 months of pre-intervention. This experiment is a systematic-based intervention study, conducted on 10 inpatient units in 10 residency training programs, across hospitals in the United States.

3) Describe the key dependent variable(s) specifying how they will be measured.

The dependent variable in this experiment is medical errors. There will be multiple subtypes of medical errors including errors related to diagnosis, errors related to therapy other than medication or procedure, errors related to history and physical examination, medication-related errors, procedure-related errors, falls, and nosocomial infections. All these different errors will be measured 6 months pre-intervention and 6 months post-intervention. The data collection on medical errors will be done in multiple different forms: through incident reports, solicited reports from nurses working on the unit, and daily medical-error reports from residents, collected through post-shift surveys. This methodology is adapted from Starmer et al. (2015), to ensure using previously successful techniques for collecting this type of information. Written and oral handoffs will also be observed in a manner similar to those in Starmer et al., (2015). Copies of written handoffs will be collected after shifts for evaluation, and audiotaped verbal handoffs will also be evaluated. Nurses are required to wear some form of technology throughout shifts to help with communication. For example one of the most common systems used is called Vocera (Figure 4). These devices are used to communicate throughout the hospital, making it easier to stay in contact with coworkers. Given nurses are accustomed to wearing these devices, it wouldn't be burdening them if a small recording tool was attached to their Vocera during a handoff. This would allow for both verbal and written handoffs to be observed pre and post-intervention of the I-PASS mnemonic and template. The number of errors will be collected in the form of counts, with the main difference being the implementation of the I-PASS mnemonic and I-PASS written template.

4) How many and which conditions will participants be assigned to?

Participants who give consent to take part in this systematic intervention study will all complete the I-PASS training. There will be a 6 month period where data on medical errors is collected pre-intervention. Then the 6-month intervention and training will occur, and the final data on medical errors will be collected 6 months after implementation. For each hospital, the interventions will include the following: the I-PASS mnemonic (Figure 6); an I-PASS physical outline, which will serve as the anchoring component for both verbal and written handoffs; a 2-hour workshop aimed at improving communications skills along with practicing the I-PASS technique; direct observation tools used by coworkers to provide feedback; supplementary material and literature of where the I-PASS method has been successful; and finally branding tools used to promote the use of this method in the future. There will also be a matched control group in each of the hospitals to limit the effects of confounds like time-of-year factors. This will allow for the analysis of the pre-intervention along with the post-intervention to compare, with the matched control group.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Given the data is collected in counts, Poisson regression will be used to analyze the data. For this experiment, the count data is each type of medical error occurring during the pre and post-intervention periods.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Data collected from individuals who didn't complete the full intervention will not be included in the statistical analysis. This will help to remove outliers due to improper training, which could potentially skew the data. errors related to diagnosis, errors related to therapy other than medication or procedure, errors related to history and physical examination, medication-related errors, procedure-related errors, falls, and nosocomial infections.

7) How many observations will be collected or what will determine the sample size?

No need to justify the decision, but be precise about <u>exactly</u> how the number will be determined.

Starmer et al., (2015) calculated based on a single site study that 6 months of data collection at each site is sufficient for more than 90% power to detect a 20% reduction in overall error rates (Starmer et al., 2015). Therefore in this proposal using 10 hospitals (one more than the methodology from Starmer et al.,), 6 months of pre-intervention, intervention, and post-intervention should be enough data to increase the power enough for detecting a real effect.

8) Anything else you would like to pre-register?

(e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

Nothing else to pre-register.

Appendix B

Consent Form Experiment One

Project Title: A proposed study investigating effects of interruption on nurses, and how to limit errors in patient handoffs

Researcher: Olivia Couture

Faculty Advisor: Justin Hulbert

I am a student at Bard College, and I am proposing a multistep study investigating the effects of interruptions on completion of the Tower of London task.

If you agree to participate in this experiment, you will be asked to complete the Tower of London task 25 times. The experiment itself is designed to take around 45 minutes to complete. The moves you take to complete the task will be anonymous and confidential. The experiment will take place on a computer in a lab. Each part of the experiment will be explained with directions on the computer. Two practice runs will be given in order to ensure no confusion. At this point, before the experiment begins, would be the time to ask any questions.

There should be no potential risks to this experiment. However, if you are uncomfortable or frustrated in any way, feel free to leave the lab and end the experiment.

All the information you provide will be anonymous and confidential. If I were to run this experiment the results from the experiment will remain in a file in the lab.

It is suggested that you make a copy of this consent form to keep for yourself. A copy can be made for you by a researcher upon your request.

Participant's Agreement:

I understand the purpose of this research. My participation in this experiment is voluntary. If I wish to exit the study for any reason, I may do so without having to give an explanation.

If I have questions about this study, I can contact the researcher at oc3843@bard.edu or the faculty advisor at jhulbert@bard.edu.

I am at least 18 years of age and I consent to participate in today's study.

________________________________ By signing I agree to participate in the study.

Appendix C
Consent form Experiment 2

Project Title: A proposed study investigating effects of interruption on nurses, and how to limit errors in patient handoffs

Researcher: Olivia Couture

Faculty Advisor: Justin Hulbert

I am a student at Bard College, and I am proposing a study to minimize healthcare errors through the implementation of the I-PASS mnemonic and written template.

If you agree to participate in this experiment, you will participate in data collection for 1.5 years. First, there will be a 6-month pre-intervention period, where verbal and written handoffs will be evaluated. Next, a 6 month intervention period will include the implementation of the I-PASS mnemonic and written template. Finally, another 6-month period of data collection will occur. At the beginning of the experiment, the participants will be randomly assigned to either the implementation or a matched control group. Therefore the pre and post-intervention periods will look the same for all participants. It will consist primarily of the evaluation of both verbal and written patient handoffs.

There should be no potential risks to this experiment. However, if you are uncomfortable or frustrated in any way, feel free to express concerns to the researcher, and potentially leave the study.

All the information you provide will be anonymous and confidential. If this experiment were to run the data collected will remain with the research nurses.

It is suggested that you make a copy of this consent form to keep for yourself. A copy can be made for you by a researcher upon your request.

Participant's Agreement:

I understand the purpose of this research. My participation in this experiment is voluntary. If I wish to exit the study for any reason, I may do so without having to give an explanation.

If I have questions about this study, I can contact the researcher at oc3843@bard.edu or the faculty advisor at jhulbert@bard.edu.

I am at least 18 years of age and I consent to participate in today's study.

________________________________ By signing I agree to participate in the study.

Appendix D
Budget Experiment 1a

The following budgets are used to propose potential costs for both experiments. Although much of the costs attempt to accurately depict these necessities, there is a high possibility of more money needed. The two experiments combined for a total of $46,905. This is obviously a steep price for the completion of two experiments; however, the implementation of the I-PASS system will serve to save significantly more money. Just last year alone, hospital insurance paid $20 billion as a result of preventable medical errors. Previous literature has already found that handoff implementations have led to a 23% decrease in overall medical error (Starmer et al., 2014). Applying this to the money spent last year alone, this would say $4,600,000,000. Although this is estimating numbers, the grant money received for these studies could serve to save millions from insurance companies. Not only will the limitation of healthcare errors lead to a decrease in the money needed to pay insurance companies, but will save many lives. There is no price that can be placed on a loved one's life, and although errors are inevitable, preventable ones should be prevented at all costs. This two-part proposal will serve to identify the effects of interruptions on nurses and implement a universal way to conduct patient handoffs. Ultimately the implementation of the I-PASS mnemonic and template, will manage the intrinsic load of healthcare workers, and lead to a decrease in preventable healthcare errors, decreasing the number of deaths and insurance payouts per year.

Budget Summary

Fees	**$500**
Project staff	$11,250
Materials	$850
Total Budget	**$13,200**

Budget Breakdown

	Role	Number of days	Daily rate ($)	Total cost ($)	Justification
PROJECT STAFF and PARTICIPANTS	Participants (500)	1	$10	$5,000	Participants will be compensated 10 dollars for completion of the experiment.
	Research Assistants (5)	10 (each)	$100	$5,000	Research assistants will work 10 days, running 10 participants a day.
	Project Manager	50	$25	$1,250	Project manager won't be required to run experiments, but available in case of questions.
			Subtotal	$11,250	
MATERIALS	**Description of item**	**Quantity**	**Rate ($)**	**Total cost ($)**	**Justification**
	Computer software	1	$500	$500	Estimated from online ToL software cost
	ToL manual	1	$200	$200	Contains information to help run ToL program
	Dell Desktop w/ monitor	5	$150	$750	How the experiment will take place (with the ability for multiple participants at once).
			Subtotal	$1,450	

Appendix E
Budget Experiment 1b

Budget Summary

Fees	**$500**
Project staff	$11,250
Materials	NA
Total Budget	**$11,750**

Budget Breakdown

	Role	Number of days	Daily rate ($)	Total cost ($)	Justification
PROJECT STAFF and PARTICIPANTS	Participants (500)	1	$10	$5,000	Participants will be compensated 10 dollars for completion of the experiment.
	Research Assistants (5)	10 (each)	$100	$5,000	Research assistants will work 10 days, running 10 participants a day.
	Project Manager	50	$25	$1,250	Project manager won't be required to run experiments, but available in case of questions.
			Subtotal	$11,250	

Appendix F
Budget Experiment 2

Budget Summary

Fees	
Project staff	$16,400
Materials	$5,655
Training	$1,000
Total Budget	**$23,005**

Budget Breakdown

	Role	Number of days	Daily rate ($)	Total cost ($)	Justification
PROJECT STAFF and PARTICIPANTS	Participants (50 per hospital)	na	na	$1,000	Each hospital will be given a $100 gift card that participants have a chance to win
	Research Assistants (1 per hospital)	52	$100	$5,200	Research assistants will go in once a week to review data for 1 year (pre + post intervention periods)
	Project Manager	30	$100	$3,000	Project manager will spend a month reviewing the final data
	Research Assistants (10)	72	$100	$7,200	One research assistant per hospital will go in 3 times a week (for 6 months) to ensure proper implementation
			Subtotal	$16,400	

	Description of item	Quantity	Rate ($)	Total cost ($)	Justification
MATERIALS	I-PASS template printout	100	0.05 c	$5	Actual template used in intervention (only half participants will have implementation)
	Supplemental material (10 pages)	100	0.05 c	$50	Information about where the I-PASS has previously worked and why
	Mini voice recorder	30	$20	$600	3 recorders per hospital (they will be switched off being used)
	I-PASS template paraphernalia	500	$10	$5,000	Sticker, water bottles, etc. will be handed out to promote the longevity of this implementation. They will serve to help brand the project
			Subtotal	$5,655	

	Description of item	Quantity	Rate ($)	Total cost ($)	Justification
TRAINING	Research Assistant (for training)	10	$100	$1,000	A trainer will go to each of the 10 hospitals and complete a 2-hour workshop for participants
	Room to rent for training	10	NA	NA	One room will be rented per hospital to complete the training (prices might vary per hospital)
			Subtotal	$1,000	