What's Lagging in our Understanding of Interruptions?: Effects of Interruption Lags in Sequential Decision-Making

Jennifer Sloane (j.sloane@unsw.edu.au)
Chris Donkin (c.donkin@unsw.edu.au)
Ben Newell (ben.newell@unsw.edu.au)
Garston Liang (garston.liang@unsw.edu.au)
School of Psychology, University of New South Wales
Sydney, 2052, Australia

Abstract

Interruptions are an inevitable part of every day life. Previous research suggests that interruptions can decrease performance and increase errors and response time. Additionally, there is evidence that providing a lag time prior to an interruption can mitigate some of the interruption costs. The goal of this paper is to investigate the effects of interruptions and interruption lags and explore possible strategies to attenuate interruption costs. A novel sequential decision-making paradigm was used, where the difficulty of the task and type of interruption were the two experimental manipulations. The results indicate that there is a potential benefit to including a lag time when presented with interruptions.

Keywords: interruption; interruption lag; decision making

Introduction

Interruptions are a common occurrence in daily life. From a telephone ringing in the middle of a conversation with a friend, to a nurse handing an X-ray to a surgeon in the midst of a procedure, interruptions can happen at any moment and in any situation. The interruption literature dates back to the early 1900's when Zeigarnik (1938) surprisingly found that interrupted tasks were better recalled compared to tasks that were uninterrupted. This is often referred to as the "Zeigarnik effect". However, research within other fields, such as aviation, suggests interruptions can have negative impacts on behavior. For example, Fitts and Jones (1947) explain, "forgetting may occur when something unusual happens to interrupt or momentarily distract the pilot from his normal routine." Although there has been conflicting results when trying to replicate the Zeigarnik effect and countless of studies on interruptions since the 1920s, Gillie and Broadbent (1989) argue it is even more important to research how easily can people resume a task after being interrupted and what makes interruptions disruptive?

To answer these questions, Gillie and Broadbent (1989) had participants complete a complex computer-based adventure game and manipulated the types and duration of interruptions within the task. They found that similarity to the primary task and the complexity of the task lead to disruptive interruptions, but not the length of an interruption or when it occurred (Gillie & Broadbent, 1989). However, it is worth noting that there were only 10 participants in the experiment and this study was completed 30 years ago. In a more recent review of interruptions, Borst, Taatgen, and van Rijn (2015) conclude that there are three main disruptive factors: duration of the interruption, complexity of the interruption, and the moment of the interruption. Research on the effects of interruptions has dramatically increased in recent years, especially in fields where interruptions can lead to serious and sometimes even fatal consequences, such as in medicine (Westbrook, Raban, Walter, & Douglas, 2018; Walter, Li, Dunsmuir, & Westbrook, 2014; Westbrook et al., 2010), aviation (Gontar, Schneider, Schmidt-Moll, Bollin, & Bengler, 2017), and driving (Klauer et al., 2014; Young, Salmon, & Cornelissen, 2013) just to name a few.

Here, we will define interruptions as a break from one task in order to complete another task, and in our experiment, resuming the primary task can only occur once the secondary task is completed. Within the literature of interruption lags, studies have often used paradigms that are inherently complex and only include one interruption (Gillie & Broadbent, 1989; Trafton, Altmann, Brock, & Mintz, 2003; Cane, Cauchard, & Weger, 2012). Therefore, the main aim of the current experiment is to explore strategies to minimize interruption costs in a decision-making task with varying levels of difficulty so that we can easily manipulate the frequency, type, and location of interruptions. This is a novel sequential decision-making task that will be referred to as "The Mazing Race", which will be explained in greater depth later.

Theoretical Framework

Theories for understanding human cognition have been around for decades. Adaptive Control of Thought-Rational (ACT-R) is one cognitive architecture to model human memory that has been gradually developing for years (J. Anderson, Lebiere, Lovett, & Reder, 1998). Derived from ACT-R, the Altmann and Trafton's Goal Activation Model (GAM) theorizes whichever goal is most *active* will govern behavior. This contrasts to the basic "last-in, first-out" structure to model goal behavior, which assumes the *newest* goal directs behavior. Although this specific model will not be implemented in this study, the model is important to understand as it motivates the research question and design.

GAM predicts that people can take time to prepare before goals are suspended or interrupted. Therefore, the model suggests it may be important to give a cue before an interruption. Specifically, the GAM "predicts that interruption lag is critical to the ability to resume an interrupted goal" (Altmann

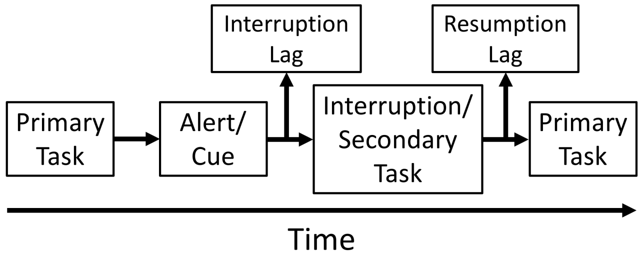


Figure 1: Image modified from Trafton et al. (2003) for a visual representation of the model including interruption and resumption lags.

& Trafton, 2002). Figure 1 illustrates a model to visualize what is happening during an interruption as a function of time. The overall idea of this model is that a person in the middle of completing a primary task is suddenly interrupted with another task, and ultimately has to resume the primary task. Often times there may be an alert, or cue, before the interruption occurs. The interruption lag is the time between the alert and the onset of the interruption. Depending on the context, the duration of the interruption lag may be able to be manipulated. Finally, the resumption lag is the time it takes to resume the primary task once the interruption has ended. This is often the dependent variable in experimental studies investigating interruption lags. One prediction from this model is that the interruption lag gives one time to prepare to resume the primary task after being interrupted.

To further understand this model, we will elaborate on a real-world example alluded to in the introduction. Imagine two people are in the midst of a conversation and suddenly the phone rings. Before the individual goes to answer it, she has the option to quickly end the current conversation, ignore the incoming call, or temporarily pause the conversation. In this scenario, the phone ringing is the alert and choosing to answer the phone would be the interruption to the primary task of the current conversation. If she chooses to pause the conversation, it would be advantageous to take a couple of seconds to remember exactly where the conversation has left off in order to successfully resume the conversation after the call. This is the idea of the interruption lag.

Interruption Lags

Over the past couple of decades, there have been several studies focusing on the effects of interruption lags. However, there are conflicting results with regards to the benefits of interruption lags. On one hand, problem solving tasks (e.g. Tower of London) showed interruption lags lead to faster resumption times compared to no lags (Morgan, Patrick, & Tiley, 2013; Hodgetts & Jones, 2006b, 2006a; Trafton et al., 2003). In fact, Hodgetts and Jones (2006a) found that even a two-second interruption lag can aid resumption on the primary task. Although most research on interruption lags has focused on static contexts, Labonté and colleagues show that a pre-interruption warning can be beneficial in dynamic environments, as well (Labonté, Tremblay, & Vachon, 2019, 2016). On the other hand, there were no benefits to including interruption lags within a reading task (Cane et al., 2012). The authors suggest that the lack of an effect is possibly because interruption lag effects may be dependent on the specific task (e.g. reading task vs. problem solving task).

It is also important to note the complexity of these tasks. For instance, Trafton et al. (2003)'s primary task was a computer game where participants had to keep track of a number of different resources including munitions, fuel, fuel tanks, vehicles, and more. Even the interruption was an involved tactical assessment task lasting 30 seconds. Similarly, the interruption in the reading task was a full minute long. The studies mentioned here investigated the effects of interruption lags in complex primary and secondary tasks. This current study looks to extend the literature by asking what effect, if any, will interruption lags have on a "simpler" task? The "simpler" task will be a novel sequential decision-making task. It is simpler in the sense that participants had to make very quick decisions and the interruptions were relatively short, as well. This paradigm is also novel because the number of interruptions was manipulated, rather than just having one interruption throughout the entire duration of the task. This is arguably a better model of the real world as interruptions are often frequent, unavoidable, and unpredictable.

Method

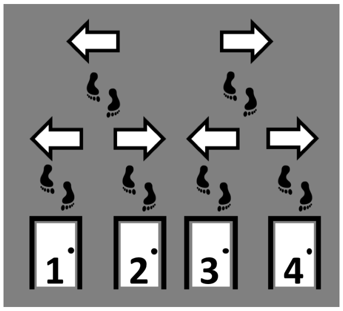
Participants

A total of 64 undergraduate students from the University of New South Wales were recruited to complete the experiment for course credit. Five participants' data were removed from analysis because the programmed crashed, so they were unable to complete the study, leaving 59 participants left for analysis.

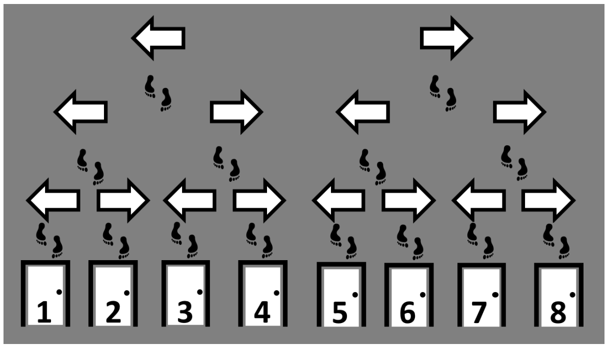
Design

This study was a 3 (difficulty: easy, medium, hard) x 3 (type of interruption: no interruption, interruption, and interruption + lag) fully within-subject design. Participants completed every combination of the conditions once for a total of nine blocks. The Mazing Race was the primary task and the interrupting task was a short-term recognition memory task. The number of interruptions depended on the difficulty level of the block. We were concerned about the difficulty of the task, and so we ensured participants completed the blocks in order of difficulty, from easiest to hardest. Within a set of problems with the same difficulty level, the type of interruption was randomized.

Primary Task: The Mazing Race In The Mazing Race participants had to make a series of decisions to go either "left" or "right" to work their way through a maze to open up doors. Figure 2 shows a visual representation of the underlying structure of the maze. These images were the stimuli used in the experiment and examples of what participants

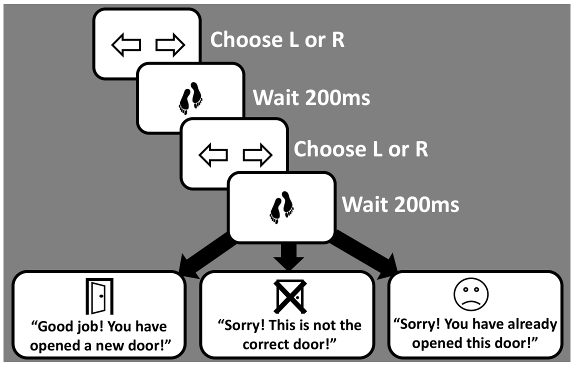


(a) Easy condition

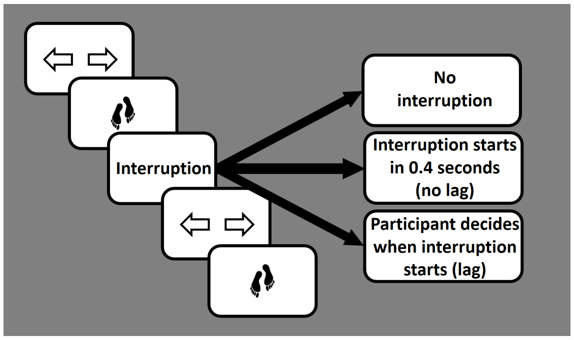


(b) Medium condition

Figure 2: The underlying structure of The Mazing Race for the (a) easy condition and (b) medium condition. The design required participants to open doors in the given order.



(a) Trial outcomes



(b) Interruption types

Figure 3: Schematic representation of the experimental design in the easy condition, depicting the (a) three possible trial outcomes and the (b) interruption types.

saw before each block. This figure illustrates a maze in the easy condition (a) with a total of four doors and a maze in the medium condition (b) with a total of eight doors. Although it is not displayed, the hard condition was of a similar structure, but had one additional decision-making level, resulting in a total of 16 doors. We named it The Mazing Race as it is a race to get to the bottom of every unique path in order to open all of the doors in as few attempts as possible. Once a door was opened, it stayed opened for the remainder of the block. Thus, the main dependent variables were the number of doors successfully opened and the number of trials needed to complete each block. Response times were recorded for further analysis, specifically looking at the response time of every decision (i.e. from when a stimulus is presented until the participant makes a keyboard response).

After participants studied the underlying structure of the maze, they pressed the space bar to start the block. Then, as shown in Figure 3a, two arrows appeared on the screen: one pointing left (L) and one pointing right (R) and participants simply had to choose to go L or R with the respective arrow keys. After every decision, animated footprints appeared for a total of 200ms symbolizing the participant walking down to the next level of the maze, where they made their next decision to go L or R. In the easy condition, for example, after two sequential decisions they reached the bottom of the maze. Every difficulty level had a maximum number of attempts to open all of the doors. In the easy condition it was 8 attempts, medium had 16, and hard had 32. These maximum numbers were included to try to minimize participants' frustration while completing the task.

Participants were required to open the doors in a specific order as shown in Figure 2. The order was always the same: starting by opening the left-most door and systematically working their way to the right-most door. Therefore, on any given attempt to open a door, there was always only one correct response. Feedback was provided every time the participant reached a door (see Figure 3a). If they reached the correct door they received positive feedback saying, "Good job! You have opened the correct door!" If they reached a door that had not already been opened, but was the incorrect door, they received negative feedback saying, "Sorry! This is not the correct door!" Finally, if they opened a door that was already opened, they also received negative feedback saying, "Sorry! You have already opened this door!" To successfully complete an easy block, for example, participants needed to go down the following four paths in this sequence: LL, LR, RL, RR. The block ended when the participant either successfully opened all of the doors or exceeded the maximum number of attempts. The experiment ended when all nine blocks were completed.

Interrupting Task: Recognition Memory Test Past research has shown that similarity and complexity between the primary and interrupting task are factors that determine if the interruption is disruptive (Borst et al., 2015; Gillie & Broadbent, 1989). Because the main interest was the potential benefits of interruption lags, it was necessary that the interruptions were disruptive. Therefore, a recognition memory task was chosen because we assumed that both the primary and secondary tasks relied on a similar subset of memory-related cognitive processes. Figure 3b illustrates the three types of interruptions: no interruption, interruption, and interruption + lag. Participants were explicitly told what type of interruption to expect before the start of each block.

In our memory task there was a study and a test phase. The stimuli included randomly selected words from a list of 1535 words, where all the words were between three and six letters and one-syllable. This is the same word pool as used in Donkin and Nosofsky (2012). Anticipating that some participants may strategically try to keep count of the number of doors they have opened, numbers (randomly generated between 0-999) were also included in the memory test as a way to interfere with any possible counting. In the test phase, participants were presented with one “old” (i.e. previously studied) item and one “new” (i.e. previously unstudied) item and they were instructed to select the word that they believed to be the old item. During every interruption, there was a total of 10 study items and 10 test pairs. Each study item was randomly selected to be either a word or a number and test pairs could be two words, two numbers, or one of each.

The memory test was programmed to occur once in each set of four trials, where a trial is an attempt of opening a door, in The Mazing Race with a set number of interruptions in each condition. There was only one interruption in the easy condition, up to four in the medium condition, and up to eight in the hard condition. The interruptions were purposefully random and spread out to make it harder for the participant to anticipate when they would be interrupted. Before the memory task began, participants completing an interruption block saw a screen that said: “Start memory test **NOW**” (the task began automatically after 400ms) and in the interruption + lag block they saw a screen that said: “**Think about where you are in the Maze.** Press the space bar to start the memory test”. The interruption lag was self-paced, meaning participants decided when to start the memory task. As soon the memory task was completed, participants immediately resumed The Mazing Race at the exact point where they left off and were given no environmental cues about where they were in the maze, which they were told from the start.

Furthermore, after every block, participants were given feedback on their performance for both tasks. For The Mazing Race, they were shown the number of doors they successfully opened and, if there were interruptions, they were shown the percentage of correct answers on the memory test. Lastly, participants were instructed that performance on The Mazing Race and the memory test were equally as important.

Results

We predicted that performance would be best in the no interruption condition and worst in the interruption condition. We expected the interruption + lag condition to fall somewhere between the others, as the lag would provide time to prepare to switch tasks and resume The Mazing Race. As this is a novel paradigm, several different analyses were carried out to try to fully understand the results. We will report the results of both frequentist and Bayesian repeated-measures ANOVAs. The Bayesian anaylses were performed using JASP (JASP Team, 2018), with priors set to their default values within the program. We report Bayes Factors (*BF*), which express the probability of the data given the alternative hypothesis (H_1) relative to the null hypothesis (H_0). A BF = 1-3 indicates weak evidence for the alternative hypothesis and a $BF > 30$ indicates strong evidence for the alternative hypothesis. Also note that for the purposes of this proceedings paper, due to the large number of comparisons, and exploratory nature of this investigation, we will only present the result of omnibus F-tests as a rough indicator of whether there were differences among conditions as a result of the introduction of interruptions. As such, we will focus on describing the qualitative pattern of the means and attempt to provide a more holistic interpretation of the overall pattern of results.

Before looking at specific dependent variables, Table 1 illustrates results from the interruption task. Participants performed equally well in both the interruption and interruption + lag conditions. Although performance decreased slightly as the primary task got harder, performance was still well above chance in all of the conditions. This suggests that participants were engaged in the secondary task and not using all of their cognitive resources on The Mazing Race.

To measure performance on the task, we first observed the average number of doors participants opened (Figure 4). The dotted lines represent the maximum number of doors in each level of difficulty: four doors in easy, eight doors in medium, and 16 doors in hard. Perfect performance would be to open all the doors in four, eight, and 16 trials, respectively. Looking at the Figure, it doesn’t appear that the type of interruption affected the number of doors opened in the easy (BF_{10} = .68; $F(2,116) = 2.85$, $p = 0.06$) or hard (BF_{10} = .12; $F(2,116) = 0.87$, $p = 0.42$) conditions. There may have been an effect of interruptions in the medium condition, but the evidence is

Table 1: Summary Statistics of Interruption Task

	Easy	Medium	Hard
Interruption	0.82 (0.12)	0.79 (0.12)	0.77 (0.13)
Interruption + lag	0.82 (0.14)	0.80 (0.11)	0.77 (0.11)

Average probability of correct responses on the interruption task (memory test) across the different levels of difficulties. Standard deviations are provided in parentheses.

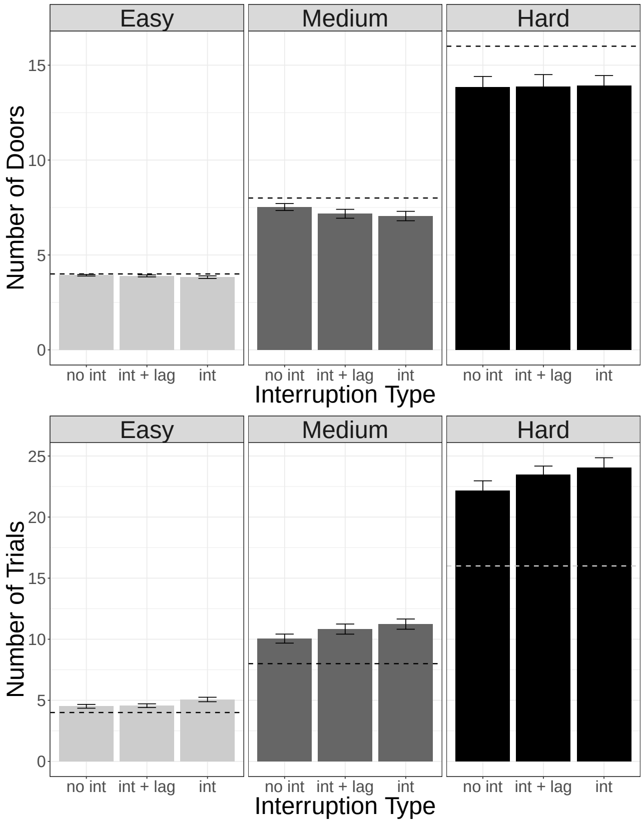


Figure 4: Average number of doors opened (top) and average number of trials needed to successfully complete the maze (bottom) in each condition as a function of interruption type. Dotted lines represent the maximum number of doors. Error bars indicate standard error in this and the subsequent figure.

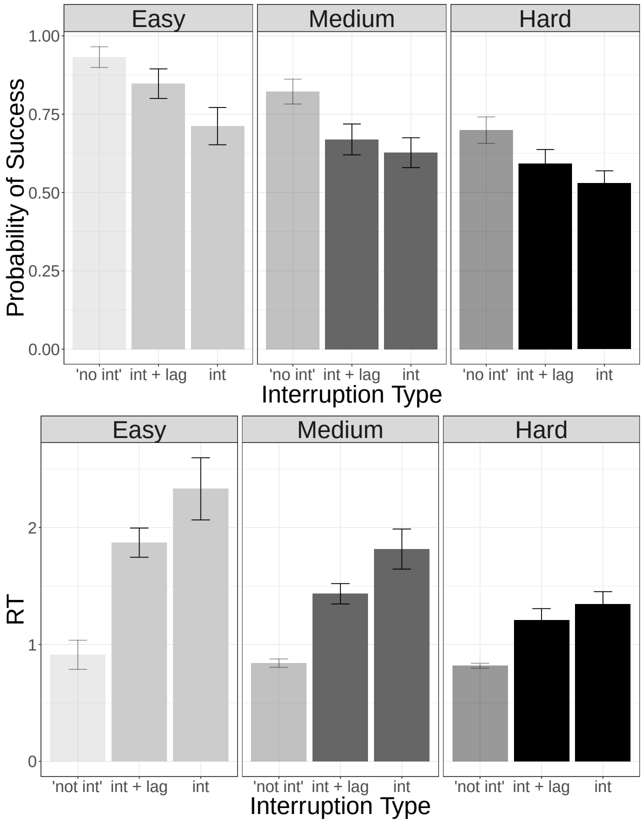


Figure 5: Probability of successfully opening the correct door (top) and median reaction times on trials immediately following interruptions. The "no interruption" condition is a lighter shade to show that it represents a baseline group for comparison even though there was no interruption in these conditions.

inconclusive ($BF_{10} = 1.427$; $F(2,116) = 3.78$, $p = 0.03$). However, it is likely that there were ceiling effects, especially in the easy condition, such that participants were opening all, or close to all, of the doors.

Even if participants successfully opened all of the doors, it is possible that they made more mistakes and needed more trials to open all doors when interrupted. Therefore, we next looked at the average number of trials needed to complete the block (Figure 4). The type of interruption did effect the number of trials in the easy condition (BF_{10} = 38.22; $F(2,116)$ = 7.79, $p < 0.001$), medium condition (BF_{10} = 5.93; $F(2,116)$ = 5.47, p = 0.01), though the statistical evidence was less clear for the hard condition (BF_{10} = 1.82; $F(2,116)$ = 4.10, p = 0.019). Focusing on the mean scores in all difficulty conditions, we see that performance tends to decrease across interruption type with best performance in no interruption, followed by interruption + lag, with poorest performance in the interruption condition without lag.

Our next analyses examined the probability of successfully opening a door and the average median RT on trials immediately following an interruption (Figure 5). In order to have a baseline condition, we created a no interruption ("no int") condition for these two figures. As a reminder, there was one interruption in the easy condition, two in the medium, and four in the hard. That gives one, two, and four data points per participant in the respective conditions. Therefore, for the no interruption condition, we sampled one, two, and four data points for each respective condition from every participant to represent where an interruption may have occurred. It was predicted that the no interruption condition would have the highest probability of success and the fastest RT, the interruption condition would have the lowest and the slowest, and the interruption and lag would fall somewhere in the middle.

Turning first to the probability of success of opening a door immediately after an interruption, there was an effect of interruption type in all of the conditions: easy (BF_{10} = 12.60; $F(2,116) = 5.81$, $p < 0.01$), medium (BF_{10} = 24.91; $F(2,116)$ = 7.13, p = 0.001), and hard (BF_{10} = 15.11; $F(2,116)$ = 6.57, $p < 0.01$). Looking at the means, we can see there was the biggest difference between "no interruption" and interruption + lag, such that the introduction of the interruption had a relatively large effect on the next trial. In all conditions, however, we do still see a benefit of the lag, with worse performance in the interruption without a lag condition.

Next, we looked at median RTs following interruptions as a way to measure resumption lag. The ANOVA analyses reveal very large effects for easy ($BF_{10} > 100$; $F(2,116) = 22.50$, $p < 0.001$), medium ($BF_{10} > 100$; $F(2, 116) = 34.8$, $p < 0.001$), and hard ($BF_{10} > 100$; $F(2,116) = 28.59$, $p < 0.001$) conditions. Again, we saw a similar pattern when looking at the mean scores. The "no interruption" had the shortest RTs and then a big jump up to interruption + lag, and the interruption conditions had the longest RTs. We will turn to the discussion for further possible interpretations of these results.

Discussion

The aim of this study was to analyze the effects of interruptions and explore the possible benefits of interruption lags in a novel sequential decision making task. While performance was, by no surprise, the best in blocks without interruptions, we did find benefits to having lag time when there was an interruption.

Using the number of doors opened as a dependent variable did not show any large effects of interruption type. Additionally, we did not see the quantitative pattern of data like we saw in the other analyses. However, 49 participants successfully opened every door in all three easy blocks, suggesting performance was at ceiling. This has real world implications, such that if one gets interrupted in the middle of an effortless task, the interruption may not disrupt the primary task at all. Performance, however, did begin to decline when the primary task got harder. In the medium and hard conditions, 39 participants successfully opened all of the doors, with a handful of participants opening less than 50% of the doors.

When looking at the maximum number of trials needed to open all the doors, we did begin to see effects of interruption type. This was the first analysis where a consistent pattern of data emerged. Participants were able to complete the task in the lowest number of trials when there were no interruptions. Performance appeared to decrease in the interruption + lag condition and even more so in the interruption condition. This makes sense as participants were explicitly told to use the interruption time to try to remember their place in the maze.

When interrupted, it often takes time to pick up where you left off. For this reason, we were interested in observing the trials that occurred immediately following interruptions, specifically looking at the probability of success and response time when making the subsequent decision. The probability of success was highest and the average RT was the shortest when examining data from the no interruption condition because participants had nothing from which to be distracted. Additionally, we see a similar trend in the data as previously mentioned, where the interruption lag appears to be improving performance (compared to the interruption condition) in both of these analyses.

Limitations

Observing the effects of interruptions is difficult because it is unreasonable and unrealistic to interrupt participants on every single trial. For that reason, we decided to only include interruptions on $\frac{1}{4}$ of the trials. Therefore, we were left with limited data points for each participant. One solution would be to increase the number of interruptions, but that may be too cumbersome and frustrating for participants. Another solution would be to increase either the number of participants or number of trials per participant. Additionally, participants were required to open the doors in the same order (i.e. left to right) in all the blocks and always completed the blocks in order of difficulty (i.e. easy to hard). Therefore, although the overall RTs are longest in the easy condition and shortest in the hard conditions, this is likely due to practice effects. By the time participants get to the more difficult conditions, they can begin to anticipate their next move resulting in quicker decisions. However, models of volitional action control (Heise, Gerjets, & Westermann, 1997) predict that difficult tasks will protect against distractions. For example, Scheiter, Gerjets, and Heise (2014) and Wirzberger, Bijarsari, and Rey (2017) found that irrelevant interruptions only impaired performance in the easy, and not difficult, conditions of their respective experiments. The competing theories of whether practice effects or volitional control are driving the RT effects can be tested in follow up studies by randomizing the difficulty order of the blocks.

Future Directions

Possible avenues for future research would be to make The Mazing Race more challenging, for example, by randomizing the order of doors to open. Another interesting question is would we see the same pattern of results if the interruption task was different? For example, on one hand, the interruption could be as simple as pushing the space bar every time a cue appears. On the other hand, it is possible that a spatial recognition memory task may be even more disruptive. It is necessary to implement different types of interruptions to see if these results generalize. The relative simplicity and flexibility of The Mazing Race makes it possible to address these questions in follow up studies.

Conclusions

Taken together these findings illustrate the potential benefit to including a lag time when presented with an interruption. Performance increased from interruption < interruption + lag < no interruption across levels of difficulty and across multiple analyses, suggesting there is evidence from this study that interruption lags can reduce some interruption costs. Furthermore, this complements previous research on interruption lags in problem solving tasks (Hodgetts & Jones, 2006a; Trafton et al., 2003). Follow up studies should aim to include more interruptions (if possible) to provide more data points. Additionally, modeling these results could prove invaluable in trying to understand and predict participants' performance and the types of mistakes they make. Interruptions will always be part of our daily lives, so it is not only important to study the effects and costs of interruptions, but also to study possible strategies to minimize those costs.

References

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: an activation-based model [Journal Article]. *Cognitive Science*, *26*(1), 39-83.

Anderson, J., & Douglass, S. (2001). Tower of hanoi: Evidence for the cost of goal retrieval [Journal Article]. *J. Exp. Psychol.-Learn. Mem. Cogn.*, *27*(6), 1331-1346.

Anderson, J., Lebiere, C., Lovett, M., & Reder, L. (1998). Act-r: A higher-level account of processing capacity [Journal Article]. *Behav. Brain Sci.*, *21*(6), 831-+. doi: 10.1017/S0140525X98221765

Anderson, J. R., & Lebiere, C. (2014). *The atomic components of thought* [Book]. Hoboken: Hoboken : Taylor and Francis.

Borst, J. P., Taatgen, N. A., & van Rijn, H. (2015). What makes interruptions disruptive?: A process-model account of the effects of the problem state bottleneck on task interruption and resumption. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 2971–2980).

Cane, J. E., Cauchard, F., & Weger, U. W. (2012). The time-course of recovery from interruption during reading: Eye movement evidence for the role of interruption lag and spatial memory [Journal Article]. *The Quarterly Journal of Experimental Psychology*, *65*(7), 1397-1413. doi: 10.1080/17470218.2012.656666

Donkin, C., & Nosofsky, R. M. (2012). A power-law model of psychological memory strength in short-and long-term recognition. *Psychological Science*, *23*(6), 625–634.

Fitts, P., & Jones, R. (1947). Analysis of factors contributing to 460 "pilot error" experiences in operating aircraft controls 1947. *Report TSEAA*, 694–12.

Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive? a study of length, similarity, and complexity [Journal Article]. *An International Journal of Perception, Attention, Memory and Action*, *50*(4), 243-250. doi: 10.1007/BF00309260

Gontar, P., Schneider, S. A. E., Schmidt-Moll, C., Bollin, C., & Bengler, K. (2017). Hate to interrupt you, but… analyzing turn-arounds from a cockpit perspective. *Cognition, Technology & Work*, *19*(4), 837–853.

Heise, E., Gerjets, P., & Westermann, R. (1997). The influence of a waiting intention on action performance: Efficiency impairment and volitional protection in tasks of varying difficulty. *Acta Psychologica*, *97*(2), 167–182.

Hodgetts, H. M., & Jones, D. M. (2006a). Contextual cues aid recovery from interruption: The role of associative activation [Journal Article]. *J. Exp. Psychol.-Learn. Mem. Cogn.*, *32*(5), 1120-1132. doi: 10.1037/0278-7393.32.5.1120

Hodgetts, H. M., & Jones, D. M. (2006b). Interruption of the tower of london task: Support for a goal-activation approach [Journal Article]. *J. Exp. Psychol.-Gen.*, *135*(1), 103-115. doi: 10.1037/0096-3445.135.1.103

JASP Team. (2018). *JASP (Version 0.9)[Computer software].* Retrieved from `https://jasp-stats.org/`

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England journal of medicine*, *370*(1), 54–59.

Labonté, K., Tremblay, S., & Vachon, F. (2016). Effects of a warning on interruption recovery in dynamic settings. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 60, pp. 1304–1308).

Labonté, K., Tremblay, S., & Vachon, F. (2019). Forewarning interruptions in dynamic settings: Can prevention bolster recovery? *Journal of experimental psychology. Applied*.

Morgan, P. L., Patrick, J., & Tiley, L. (2013). Improving the effectiveness of an interruption lag by inducing a memory-based strategy [Journal Article]. *Acta Psychologica*, *142*(1), 87-95. doi: 10.1016/j.actpsy.2012.09.003

Scheiter, K., Gerjets, P., & Heise, E. (2014). Distraction during learning with hypermedia: difficult tasks help to keep task goals on track. *Frontiers in psychology*, *5*, 268.

Trafton, J., Altmann, E., Brock, D., & Mintz, F. (2003). Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal [Journal Article]. *Int. J. Hum.-Comput. Stud.*, *58*(5), 583-603. doi: 10.1016/S1071-5819(03)00023-5

Walter, S. R., Li, L., Dunsmuir, W. T., & Westbrook, J. I. (2014). Managing competing demands through task-switching and multitasking: a multi-setting observational study of 200 clinicians over 1000 hours. *BMJ Qual Saf*, *23*(3), 231–241.

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

Westbrook, J. I., Raban, M. Z., Walter, S. R., & Douglas, H. (2018). Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: a prospective, direct observation study. *BMJ Qual Saf*, *27*(8), 655–663.

Wirzberger, M., Bijarsari, S. E., & Rey, G. D. (2017). Embedded interruptions and task complexity influence schema-related cognitive load progression in an abstract learning task. *Acta psychologica*, *179*, 30–41.

Young, K. L., Salmon, P. M., & Cornelissen, M. (2013). Distraction-induced driving error: An on-road examination of the errors made by distracted and undistracted drivers. *Accident Analysis & Prevention*, *58*, 218–225.

Zeigarnik, B. (1938). On finished and unfinished tasks. *A source book of Gestalt psychology*, *1*, 300–314.