

**Navigating the Balance Between Urgency and Importance: Exploring Task Prioritization
Strategies Across Environments**

by
Kristina Wu

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Arts
in
Psychology

Waterloo, Ontario, Canada, 2024

© Kristina Wu 2024

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Kristina Wu was the primary author of the thesis, conceived the experimental design, collected and analyzed the data.

Dr. Samuel Johnson guided the research process, including the experimental design, data analysis, and the writing of the thesis.

Dr. Gordon Brown also guided the research process, including contributions to the experimental design and theoretical framework.

Jaxin Lu coded the experiment, translating the experimental design into executable code.

Abstract

Task prioritization is essential for managing multiple tasks with competing goals, yet psychological research has been limited to basic observations in simplified environments. The current research aims to address this gap by examining how various environmental features influence task management using a novel incentive-compatible game across two experiments. In the game, participants sequenced tasks with different levels of urgency (represented by deadlines) and importance (represented by point values) over multiple rounds, aiming to maximize the number of points earned. In Experiment 1, we manipulated task schedulability (ability to plan task orders in advance) and urgency-importance correlation (degree to which urgent tasks and important tasks conflict). Results showed that planning task orders in advance helped participants better balance urgency and importance, leading to near-optimal performance. In contrast, those who did not plan ahead tended to *overprioritize* importance and performed worse as a result. In Experiment 2, we manipulated task segmentation (whether tasks were to be completed in segmented parts or as a whole), task comparability (uniform versus varied task lengths), and urgency-importance correlation. We found that task segmentation did not affect urgency or importance prioritization, however it decreased overall performance due to suboptimal task-switching. When task lengths varied, performance declined because participants overprioritized importance, while neglecting urgency and task length. Across both experiments, participants balanced urgency and importance more effectively when these factors conflicted the least, but they tended to prioritize importance over urgency more heavily as the conflict between urgency and importance increased. This research highlights the need for strategies that help individuals better balance importance relative to urgency, particularly in dynamic environments with variable task demands. Limitations and future directions are discussed.

Acknowledgements

I would like to express gratitude to my supervisor, Dr. Samuel Johnson, for his guidance and support throughout this project. His knowledge and ceaseless enthusiasm have been instrumental to my development as both a researcher and a professional. I would also like to thank my thesis readers, Dr. Derek Koehler and Dr. Evan Risko. Their efforts and expertise are deeply appreciated.

I am grateful to Dr. Gordon Brown for his contributions during the early stages of this project. I also want to thank my research assistants, Denise Albeanu, Laviessha Puviraj, and Brandon Lam, for their hard work and dedication throughout this project. Thank you to Jaxin Lu for single-handedly coding the experiment—this project could not exist without him.

A special note of appreciate goes out to my lab mates, Dogukan Demircioglu, Jordan Sheen, and Neil Wegenschimmel, as well as every other member of the Decision, Inference, and Cognitive Economics (DICE) lab. Your support has meant a lot, and I am grateful for the unique perspective and encouragement each of you has provided.

Finally, I want to acknowledge the constant support of my friends, family, and partner. I could not have done this without your endless encouragement.

Table of Contents

Author's Declaration.....	ii
Statement of Contributions.....	iii
Abstract.....	iv
Acknowledgements.....	v
List of Figures.....	vii
List of Tables.....	viii
Introduction.....	1
Optimal Task Management.....	2
Human Cognition in Task Management.....	3
Task Environments.....	6
Tradeoff Tension: The Relationship Between Urgency and Importance.....	6
Task Comparability: Uniformity or Variation in Task Length.....	8
Task Schedulability: Planning versus Selecting.....	9
Segmentation: Division of a Task into Parts.....	10
The Current Research.....	12
Experiment 1.....	14
Method.....	14
Participants.....	14
Design and Procedure.....	14
Data Processing.....	18
Exclusions.....	20
Results.....	21
Urgency-Importance Tension Environments.....	23
Psychometric Data.....	28
Discussion.....	31
Experiment 2.....	34
Method.....	34
Participants.....	34
Design and Procedure.....	34
Data Processing.....	36
Exclusions.....	37
Results.....	38
Task Switching.....	42
Task Tension Conditions.....	43
Psychometric Data.....	45
Discussion.....	51
General Discussion.....	53
Relationship to Prior Findings.....	54
Limitations.....	56
Conclusions.....	57
References.....	58

List of Figures

<i>Figure 1.</i> Task Management Game Displays in Experiment 1.....	16
<i>Figure 2.</i> Histograms of Simulated Points Earned Across Tension Conditions.....	20
<i>Figure 3.</i> Proportion of Points Earned by Framing Condition.....	22
<i>Figure 4.</i> Correlation to Each Strategy by Framing Condition.....	23
<i>Figure 5.</i> Proportion of Points Earned Across Tension Conditions by Framing Condition.....	25
<i>Figure 6.</i> Priority Score Across Tension Condition by Framing Condition.....	26
<i>Figure 7.</i> Correlation to Each Strategy Across Tension Condition by Framing Condition.....	27
<i>Figure 8.</i> Relationship Between Proportion of Points Earned and Decision Styles Score in Experiment 1.....	29
<i>Figure 9.</i> Task Management Game Displays in Experiment 2.....	35
<i>Figure 10.</i> Proportion of Points Earned by Task Segmentation and Length Conditions.....	39
<i>Figure 11.</i> Correlation to Each Strategy by Task Segmentation and Length Conditions.....	41
<i>Figure 12.</i> Relationship Between Task Switching and Performance.....	43
<i>Figure 13.</i> Proportion of Points Earned by Segmentation, Length, and Tension Conditions.....	46
<i>Figure 14.</i> Correlation to Strategy by Task Segmentation, Length, and Tension Conditions.....	48
<i>Figure 15.</i> Relationship Between Proportion of Points Earned and Decision Styles Score in Experiment 2.....	49

List of Tables

<i>Table 1.</i> Environmental Features and Anticipated Strategy and Performance Outcome.....	11
<i>Table 2.</i> Results of Linear Mixed Models for Experiment 1.....	24
<i>Table 3.</i> Correlation Matrix for Experiment 1.....	30
<i>Table 4.</i> Number of Participants Excluded from Analysis by Condition.....	38
<i>Table 5.</i> Results of Linear Mixed Models for Experiment 2.....	44
<i>Table 6.</i> Correlation Matrix for Experiment 2.....	50

Introduction

Time management consultant David Allen estimates that the typical person must keep track of about 200 unfinished tasks at any given time (Allen, 2015). It is no wonder people often feel trapped in a perpetual cycle of *time poverty*, with too many things to do and not enough time to do them (Garhammer, 2002; Giurge et al., 2020). Our ever-expanding task lists are proposed to be a symptom of a broader shift toward faster paced life (Gleick, 1999). Studies have found substantial reports of time poverty in countries that are technologically advanced and economically productive, including the United States, Canada, Australia, Russia, Japan, Korea, Norway, and Germany (Levine & Norenzayan, 1999). People living in these countries report feeling rushed all the time, a more hectic pace of life, and perceptions of not having enough time to complete their tasks. Researchers agree that task overload is linked to increased stress and lower wellbeing (Bolino & Turnley, 2005; Ilies et al., 2010; MacDonald, 2003). Thus, understanding how people prioritize tasks is crucial for developing strategies that might mitigate negative outcomes.

People often manage their tasks in suboptimal ways (Ariely & Wertenbroch, 2002; O'Donoghue & Rabin, 2000; Rabin, Fogel & Nutter-Upham, 2011; Steel, 2007). Poor task management skills are associated with reduced feelings of control over one's time and are a recipe for feeling short on time (Aeon et al., 2021; Häfner & Stock, 2010). The need for task management assistance is reflected by the global demand for task management tools, applications, and software. People increasingly rely on external sources to assist them with their to-do lists (Microsoft, 2008), including apps like Motion, Google Calendar, GoblinTools, Todoist, TaskVista, and ClickUp, to name a few. In a study examining task management tendencies among knowledge workers, respondents reported using an average of 8.9 tools for

task-tracking purposes with 79% of these workers using at least one of these tools more than once per day (Hu et al., 2024). It is no surprise that the global task management software market is expected to grow from 2.84 billion USD in 2022 to 9.39 billion USD by 2033 (Regional Research Reports, 2023). The acceleration of the market is underscored by the widely shared desire to manage our ever-increasing task lists, reduce stress, and improve productivity. Understanding the psychology of task management can help guide the development of task-tracking tools so that they align with our cognitive processes, helping us gain control over our time and improve overall task management.

Optimal Task Management

A natural starting point for assessing human task management is to identify optimal benchmarks. To do this, we can look to computer science which has long studied scheduling problems and task management algorithms. Computer scientists have developed countless algorithms to optimize task scheduling, addressing problems from simple task prioritization to complex multi-task optimization (Bugayenko et al., 2023). Much like humans, CPUs with a single-core processor may only execute one instruction at a time. To decide which tasks to prioritize, there are a wealth of algorithms including a few common ones like the Earliest Deadline First (EDF; tasks with the earliest absolute deadline are prioritized), Shortest Job Next (SJN; tasks with the shortest execution time are prioritized), and First Come, First Served (FCFS; tasks are prioritized in the order they are added to the queue; Azmi et al., 2011). More complex algorithms, like priority scheduling, assign a priority score to each task in the queue based on factors such as memory demand, importance, etc. (Ben Alla et al., 2021). Despite the extensive research and practical application of these algorithms in real-time systems, no single algorithm is universally optimal; each has its own limitations and deficiencies and will succeed or fail

depending on the context in which it operates. Scheduling problems are inherently complex and challenging to solve in computing. While algorithmic benchmarks offer valuable insights, applying them to human task management is complicated even further by cognitive biases and motivational factors. To improve task management strategies for people, we must integrate insights from computer scheduling theories with an understanding of human factors.

Human Cognition in Task Management

Like computers, humans have computational limitations and strive to minimize computational expenses. Our cognitive processes can only be pushed so far before becoming overloaded (Bannert, 2002; Sweller, 1988). To manage these limitations, we use strategies that reduce cognitive effort and manage resources efficiently, often relying on heuristics, similar to the way computer systems operate (Shah & Oppenheimer, 2008). However, unlike computers, humans lack precise processing systems, and are instead influenced by cognitive biases and subjective judgments. This can easily lead to suboptimal task prioritization, such that we might prioritize tasks that are more immediately rewarding even if they are less important (Białaszek et al., 2019). While heuristics help manage cognitive load, they can also introduce biases that lead to less-than-rational decisions. The literature has found several cases where these suboptimal processes occur.

First, people often choose to complete easier tasks first. This behaviour can be explained by effort discounting, where the perceived value of a task's reward decreases as the effort required increases (Husain & Roiser, 2018; Westbrook & Braver, 2015). This drives people to prioritize tasks that demand less cognitive effort, often resulting in procrastination or a focus on simpler tasks over more challenging ones. Research supporting this shows that people tend to prefer simpler courses of actions over more difficult ones, even if they result in longer time spent

on the task (Kool et al., 2010). People are willing to sacrifice both time and external rewards to avoid mentally taxing tasks (Westbrook, Kester & Braver, 2013). When we reach the limits of our cognitive bandwidth, we might feel overwhelmed, stressed, or frustrated—a state we are motivated to avoid (Da Silva Cezar & Maçada, 2023; Naveed & Anwar, 2020; Zumbach & Mohraz, 2008). When it comes time to select a task to complete, people might rely on their affective forecasts (predictions about how they will feel) to guide their decision (Mellers et al., 1997). If a task is perceived to be high in effort and therefore might cause stress and fatigue, it may be perceived as a threat to one’s cognitive systems (Dai et al., 2020). This perception can lead to avoidance behavior in an attempt to escape the impending demand (Kool et al., 2010; Li et al., 2020). On the other hand, simple low-effort tasks do not evoke the same threat, allowing individuals to feel productive while delaying more demanding tasks.

A related line of research suggests that people often choose smaller tasks before larger ones (Rusou et al., 2020). There are several theories explaining this smaller tasks trap. For one, people may associate shorter tasks with less effort. Even if a short task requires effort, the total effort exerted is contained to a limited and brief period. In other words, there is reduced threat of prolonged mental effort leading to cognitive fatigue (Dai et al., 2020). Aside from effort avoidance, there may be rational considerations for prioritizing small tasks. First, smaller tasks can function as a “warm up” prior to pursuing a more demanding task. By crossing tasks off a to-do list, people might feel a taste of success, thereby boosting motivation and activating thinking capacities (Job & Brandstätter, 2009; Shellock & Prentice, 1985). For instance, responding to an email at the beginning of a workday can provide a small but immediate reward, leading to an increase in motivation and self-efficacy for future tasks (Liu et al., 2022). Second, clearing smaller tasks from your task list can help reduce the overall number of pending tasks. Addressing

smaller tasks first can free up mental bandwidth otherwise dedicated to maintaining those tasks in mind. As a result, there will be less background noise, and focus can be streamlined toward the large task (Baron, 1986; Zhihong et al., 2024).

People also appear to prioritize urgent tasks over important ones (Zhu et al., 2018). In study by Zhu et al (2018), participants were asked to make a single selection between two identical tasks: Task A, labeled as urgent, and Task B, labeled as not urgent. Task A offered a payout of 12 cents, while Task B offered 16 cents. While the deadlines for both tasks could be completed well within the timeframe of the experiment, participants were asked to select just one task to complete. They found that many participants choose to complete the more “urgent” Task A, thereby sacrificing the larger reward offered by Task B. One proposed mechanism underlying the mere urgency effect is as follows: urgency creates feelings of psychological tension or discomfort, which consequently demands attention. Greater attention is directed toward the discomfort created by the urgency of the task, while at the same time, attention is diverted away from other task attributes, such as importance (Ben Zur & Breznitz, 1981; Svenson & Eland, 1987; Svenson, Edland, & Slovic 1990). This effect is pronounced in situations of time scarcity, where distant deadlines are shelved so urgent tasks can take priority and discomfort can be relieved (Holmes, 2022; Mullainathan & Shafir, 2013).

Taken together, the current literature identifies several processes that influence task selection. The first proposes that task prioritization strategies stem from heuristic processes and our general rule of thumb to preserve cognitive resources, hence why we might procrastinate difficult tasks (Ferrari, 2001). The second suggests there may be rational considerations involved in how we manage task lists, such as strategically completing small tasks to increase intrinsic motivation (Liu et al., 2022). The third emphasizes the role of affective cues, such as discomfort

or anxiety, drawing attention to specific tasks. Although these perspectives are not mutually exclusive and likely work together to guide our task selection strategies, the literature sparingly examines how these processes interact across environments. Addressing the gap requires a comprehensive framework explaining how these factors might collectively influence task prioritization across different contexts.

Task Environments

In the real world, the environment in which we select tasks can differ across many features that plausibly interact with the cognitive mechanisms underlying task management. Some environmental features relate to the tasks themselves (*intrinsic features*), whereas others relate to the conditions under which tasks are chosen (*extrinsic features*). In the current research, we consider two intrinsic features (*trade-off tension* and *task comparability*) and two extrinsic features (*schedulability* and *segmentation*).

Trade-off tension: The relationship between urgency and importance. Both urgency and importance are reasons to prioritize a task. We can think of urgency as reflecting the closeness of a deadline while importance reflects the consequences of completing (or failing to complete) it. Some tasks are both urgent and important (e.g., emergency spinal surgery) whereas others are neither urgent nor important (e.g., catching up on the new episode of *Survivor*). When the tasks on one's to-do list are composed primarily of these two types (i.e., all tasks are *both* urgent and important or *neither* urgent nor important), this constitutes a *low tension* list because one can prioritize both urgency and importance simultaneously. In contrast, other tasks are urgent but not important (e.g., grading undergraduate quizzes) or important but not urgent (e.g., working on a book proposal). When many of the tasks on one's to-do list are composed of these two types

(i.e., many tasks are urgent *or* important but not both), this constitutes a *high tension* list because urgency and importance must be traded off.

How might this aspect of the statistical structure of one's task list influence task prioritization? We consider two dimensions along which performance might vary—trade-off weights (on urgency versus importance) and optimality (relative to the payoff-maximizing task order).

Trade-off weights reflect the degree to which urgency or importance influence task selection decisions. Prioritizing an urgent (but not important task) over an important (but not urgent) task indicates that more weight was given to urgency in that decision. Trade-off weights can be calculated at a broader level by evaluating *repeated* selection choices across a task list. For example, consistently choosing tasks with the closest deadline implies that urgency holds greater weight in the decision-making process. In low tension environments, where urgency and importance are aligned, the weight attributed to both characteristics will naturally be more balanced. However, when urgency and importance conflict, the challenge of making trade-offs becomes more pronounced. In such cases, consistently choosing tasks with the closest deadline is done at the expense of importance, and thus all weight is given to urgency and none to importance. In these high tension environments, we might expect people to favour immediate gains under the principles of delay discounting—the tendency to favour immediate rewards over delayed ones (Critchfield & Kollins, 2001). Individuals might pursue important tasks over tasks that are more time-sensitive but provide a lesser immediate benefit.

Optimality refers to the task selection choices that elicit the greatest *overall* outcome. This requires balancing urgency and importance in a way that accounts for broader goals at this list level rather than individual decisions at the task level. In low tension environments, where

urgency and importance are aligned, the optimal decision at the individual task selection level—whether that be the highest importance or the nearest deadline—tends to naturally align with the optimal decision at the list level. This alignment simplifies decision-making, as selecting tasks based on either deadlines or importance leads to an optimal task sequence. In high tension environments, tasks with pressing deadlines may not align with those of high importance, creating a conflict between immediate demands and long-term outcomes. For instance, solely prioritizing urgency can lead to suboptimality if those tasks are relatively low in importance. On the other hand, focusing solely on importance might result in many missed deadlines. Achieving an optimal outcome requires more complex strategies in these cases. As such, we predict that as tension between urgency and importance increases, behaviour will deviate further from optimality.

Task comparability: Uniformity or variation in task length. Orthogonal to urgency and importance, tasks can also differ in *size*. If a freelance journalist could write a 1000 word article for \$1000 or a 200 word article for \$200, these two tasks are similar in their importance per unit time (assuming it takes 5 times longer to write a 1000 word versus a 200 word article), but the former task is *larger* than the latter. In some task environments, most tasks have uniform sizes (e.g., an Uber driver who primarily makes trips between downtown and the airport) and are therefore readily comparable, whereas in other cases the tasks vary greatly in size (e.g., juggling administrative emails [short], grading student papers [medium], and grant proposal writing [long]), rendering them less comparable.

The shorter a task takes to complete, the more immediate its reward. As proposed by delay discounting, these immediate rewards are often valued disproportionately to rewards that are to be received further in the future (Odum, 2011). For example, our freelance journalist might

be tempted to accept 5 one-day gigs that each offer a payment of \$100, instead of a two-week project that offers \$1500 upon completion. This preference for shorter tasks may be driven by the devaluation of distant rewards. By introducing varying task lengths into the study, we can examine how the appeal of proximal rewards influences task selection. We predict that the presence of varied task lengths will shift trade-off weight away from urgency and importance onto shorter tasks. As a result, behaviour will deviate further from optimality.

Task schedulability: Planning versus selecting. Turning to extrinsic features of the decision-making environment, we first consider whether a series of tasks are sequenced in advance (high schedulability) or selected one at a time in succession (low schedulability). *High schedulability* refers to situations where tasks are planned out ahead of time, allowing for a structured sequence to be established. For instance, a project manager might create a detailed schedule for a long-term project, with each task planned in advance. In contrast, *low schedulability* means there is no pre-existing plan, and tasks are selected in a more reactive way. For instance, a freelancer might choose tasks based on immediate demands or deadlines without a set plan.

Differences in schedulability offer various perspectives on how task lists are viewed. Creating a task schedule provides a zoomed out perspective where task lists can be seen as a whole. In contrast, selecting tasks in real-time provides a more narrow scope where tasks might be viewed individually. The distance associated with scheduling is represented at a higher level of abstraction, primarily concerned with the “why”, i.e., why are we doing this (Baumeister & Tierney, 2012; Liberman & Trope, 1998; Nussbaum et al., 2003; Trope & Liberman, 2014). For example, if you’re planning to start a new career path in five years, you might think broadly about why this change will benefit your life, such as personal fulfillment or financial stability.

The proximity of real-time selection is more likely to be represented at a lower level of abstraction, concerned with the “how”, i.e., how do we do this. When managing your daily tasks, such as deciding which emails need a response, you are more likely to focus on immediate actions and their direct outcomes. Fu and Gray (2004) found that individuals in such settings rely on local cues and less effective “micro-strategies”, prioritizing sub-tasks that offer immediate reward over those aligned with broader goals. Without planning, task selection may be driven by local level heuristics (prioritizing either urgency *or* importance), leading to behaviours which deviate from optimality. Conversely, when planning, task selection is more likely to be viewed from the global level, allowing for a more balanced consideration of urgency and importance, therefore supporting optimal outcomes.

Segmentation: Division of a task into parts. Finally, some task environments facilitate completing an entire task without interruption by other tasks (e.g., recording a video lecture; *unsegmented*), whereas in other cases they can be readily divided into parts that need not be done in a single go (e.g., preparing a manuscript; *segmented*).

Comparing environments with segmented versus unsegmented tasks allows us to examine how task-switching affects the weighting of urgency and importance and its impact on overall performance. Past research has found that task-switching impairs performance due to switch costs, which include the time and cognitive effort needed to reorient attention and refocus (Rogers & Monsell, 1995; Wylie & Allport, 2000). This switch consumes working memory resources that could otherwise be used for maintaining goal-relevant information (Arabacı & Parris, 2020; Liefoghe et al., 2008). We anticipate that while segmentation may not significantly alter the relative weights attributed to urgency or importance, it will affect the efficiency of task execution. When tasks are segmented, the process of switching between tasks might disrupt the

smooth progression of tasks and divert focus away from broader goals. We predict that constantly switching between tasks will lead to inefficiencies and costs that dampen performance, resulting in suboptimal outcomes compared to environments where tasks are unsegmented.

Table 1

Environmental Features and Anticipated Strategy and Performance Outcome

Environmental Features	Outcome	
	Trade-off weights	Optimality
The relationship between urgency and importance (high tension versus low)	Weight attributed to urgency and importance will be fairly equal in low tension conditions. Weight will shift toward importance as tension between urgency and importance increases.	Behaviour will be further from optimal as tension between urgency and importance increases.
Task comparability (uniformity versus variation in task length)	When task lengths are varied, compared to when they are not, weight will shift to tasks with the shortest length.	Behaviour will be further from optimal when task lengths vary, compared to when they are fixed.
Task schedulability (planning versus selecting)	Weight attributed to urgency and importance will be fairly equal when tasks are scheduled. Weight will shift toward urgency <i>or</i> importance when tasks are selected across real-time.	Behaviour will be closer to optimal when tasks are scheduled, compared to when they are not.
Division of a task into parts (segmented versus unsegmented tasks)	There will be no difference in weight given to urgency or importance when tasks are segmented or unsegmented.	Behaviour will be further from optimal when tasks are segmented, compared to when they are unsegmented.

The Current Research

These four environmental features are systematically varied across two experiments. First, we examine schedulability (planning versus selecting) across varying urgency-importance relationship environments (urgency and importance are aligned [low tension], moderately aligned [moderate tension] or misaligned [high tension]). Second, we investigate the effects of task comparability (fixed versus varied task lengths) and task segmentation (tasks are divided into parts [segmented] versus completed in one go [unsegmented]) across the varying urgency-importance relationship environments (low tension, moderate tension, high tension).

Both studies use a novel, incentive-compatible task management game that pits urgency against importance. In this game, participants freely select tasks from a list in order of priority. All tasks are hypothetical in the current iterations of the game. Participants played 9 rounds of the game in total. In each round, participants sequenced a set of tasks with the goal of maximizing the points collected across 30 time-steps (“turns”). Participants sequenced 6 tasks in each round, each assigned a point value (reflecting importance), a deadline (turn by which it must be completed, reflecting urgency), and a length (number of turns required to complete it). The specific distributions in which points, deadlines, and lengths were drawn varied across experiments.

In Experiment 1, we manipulated the task sequencing structure to vary schedulability, categorizing it as either high or low. In the high schedulability condition, participants arranged tasks in a *schedule* and could rearrange the task order as they preferred. In contrast, the low schedulability condition required participants to dynamically select tasks across time-steps, with the turn counter advancing by the length of each task upon each selection. This structure prevented participants from rearranging or revisiting completed (or expired) tasks. In Experiment

2, we manipulated both task segmentation and task comparability. For task segmentation, task turns were either grouped or ungrouped; in the grouped condition, each task was handled as a whole, one task at a time; while in the ungrouped condition, tasks were managed one *turn* at a time. For task comparability, task lengths were either always 5 (fixed) or drawn from a uniform distribution ranging from 2 to 8 turns (varied).

Across both iterations of the game, in each round, point values had a specific correlation with deadlines: (i) positive correlation ($r = 0.5$ to 0.9), negative correlation ($r = -0.5$ to -0.9), or no correlation ($r = -0.2$ to 0.2). When deadlines and points are positively correlated, tasks with closer deadlines (more urgent) are often worth fewer points (less important). As a result, task orders which prioritize deadlines tend to significantly differ from those which prioritize importance. This environment is referred to as the *high tension* condition, where deadline-driven and importance-driven task orders frequently conflict and share little overlap. In contrast, the *moderate tension* condition occurs when deadlines and importance are not correlated with each other, causing a noticeable but more subtle difference between the two task orders. When deadlines and points are negatively correlated, the closer deadline tasks (more urgent) are also often worth the most points (more important), leading to a *low tension* environment such that prioritizing importance tends to automatically prioritize urgency too.

Experiment 1

The current experiment investigated the role of scheduling tasks (relative to selecting tasks) and degree of tension between of urgency and importance on two primary metrics: (i) the trade-off weight attributed to urgency and importance and (ii) overall performance (measured by points earned relative to the point-maximizing [optimal] task order). We anticipate that scheduling will induce greater psychological distance, and in turn, prompt goal-oriented decision making at the global level, leading to an optimal balance of urgency and importance and better overall performance. In contrast, selecting tasks in real-time will decrease psychological distance, and direct decision making toward immediate, proximal concerns. As a result, we anticipate participant behaviour to be driven by specific attributes at the local level, leading to heuristic-based task prioritization and suboptimal performance. As the level of tension between urgency and importance increases, we expect performance decrements and the overprioritization of importance.

Method

Participants

We recruited 320 U.S. participants (201 male, 115 female, 4 other, $M_{\text{age}} = 41$) from Amazon's Mechanical Turk (MTurk) to complete an approximately 12-minute study in exchange for \$2.00 USD plus a bonus payment of up to \$1.00 USD, depending on performance (outlined below in *Design and Procedure*). We excluded 81 participants using the criterion outlined below in *Exclusions*.

Design and Procedure

The study was described to participants as a task management game. Participants were randomly assigned to one of two framing conditions (*planning* or *selecting*). In the planning

condition, participants scheduled the tasks to produce an ordered list, whereas in the selecting condition, participants selected tasks one at a time without the possibility to backtrack (see Figure 1 for game displays).

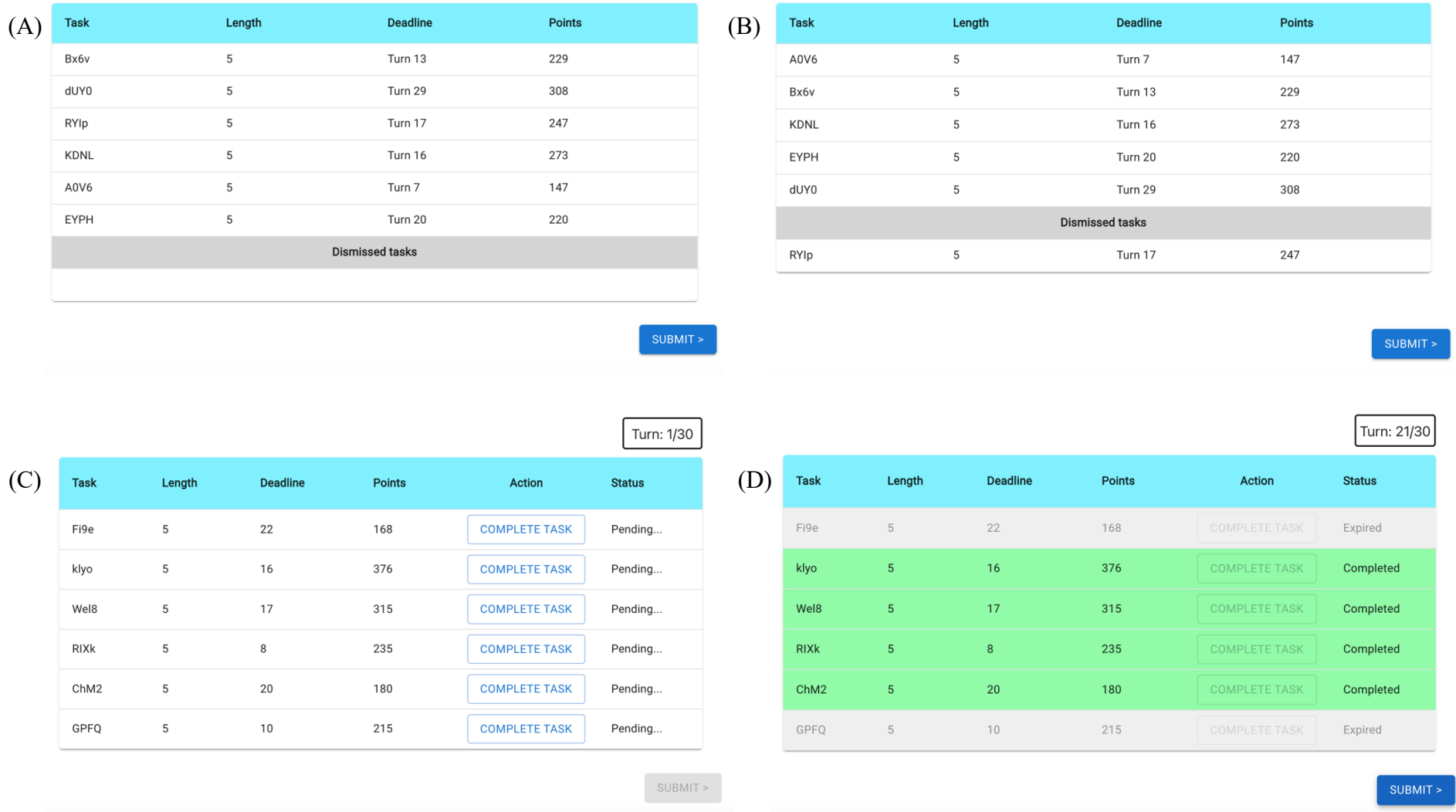
In this iteration of the game, task lengths remained constant at 5. Deadlines were randomly drawn from a uniform distribution (5 to 30). Deadlines were always constructed so that it was not possible to complete all tasks across the 30 turns (number of tasks available: median = 4, min = 3, max = 5). Point values were drawn from a normal distribution ($M = 225$, $SD = 75$). Each urgency-importance tension condition (low, moderate, or high) was presented to participants 3 times. The order of the 9 task sets was random and not blocked by tension condition. Participants were told that the goal of each round was to earn as many points as possible, which would translate into a bonus payment at the end of the study (outlined to them at the end of the instructions block).

To begin the study, participants were given instructions explaining how the game worked, tailored to their condition, and completed one practice trial to familiarize themselves with the structure of the game. In the planning condition, participants were provided with the following instructions:

You are playing a task management game. Each round, you will be presented a unique task list. Each task will have a different point value and deadline. All tasks take 5 turns, and there are 30 turns available each round. Some tasks may expire before you can complete them – any task that cannot be completed within the allotted timeframe should be “dismissed” by dragging and dropping them into the “dismissed” section below the task list. For example, a task with a deadline on Turn 6 must be completed first or

Figure 1

Task Management Game Displays in Experiment 1



Note. Task management game displays: (A) Planning condition before the first turn is taken; (B) Planning condition in a completed state, ready for submission; (C) Selecting condition before the first turn is taken; (D) Selecting condition in a completed state

dismissed, otherwise it will expire and be considered “invalid”. It is your job to rank tasks in the order you wish to complete them.

In the selecting condition, participants were provided with the following instructions:

You are playing a task management game. Each round, you will be presented a unique task list. Each task will have a different point value and deadline. All tasks take 5 turns, and there are 30 turns available each round. Some tasks may expire before you can complete them. The round ends when all tasks are either completed or expired. It is your job to complete tasks in the order you wish to complete them.

Completed tasks were highlighted in green and marked as “completed”. Expired tasks were greyed out and marked as “expired” when the deadline was passed or there were an insufficient number of turns between the current turn and the deadline to complete them. The bonus payment structure was outlined at the end of the instructions block:

At the end of the game, one round will be randomly chosen to convert into a bonus payment. It will be calculated based on the proportion of points you earn in the round relative to the number of points available. For example, if you earn 800 points and there are 1000 possible points available, then you would earn 80% of the maximum bonus payment, which is 1.00 USD, resulting in a bonus of 0.80 USD

After each round, participants received feedback on their performance, which was determined by the number of points they earned compared to the maximum possible number of points that could be scored that round (if the optimal strategy had been followed). For example, “You scored 580 points on the last question. The optimal score is 762 (76.1%).”

After completing the game, participants filled out the Maximization Scale Short (MS-Short) and the Decision Styles Scale (DSS) (Hamilton et al., 2016; Nenkow et al., 2008) in a

random order. These measures were included to gain insight into how different decision-making tendencies might influence task management behaviour.

Data Processing

Benchmarks. We first calculated several benchmarks to which participants' data could be compared, using computer simulations. First, the optimal order for each task set was determined through brute-force. Second, two heuristic strategies were simulated for each task set: (i) the urgency-first strategy which always prioritized the most urgent task and (ii) the importance-first strategy which always prioritized the most valuable task. For each benchmark, we determined both the task order and total points earned. Third, we calculated an estimate of performance given random prioritization by averaging the points earned across 500 simulated rounds, where tasks were randomly selected.

For both simulated benchmarks and actual participant responses, task orders were always coded so that "1" reflected the first-selected task, "2" the second-selected task, and so on. Since not all tasks could be completed for each task list, the lowest-ranked positions (e.g., 5 and 6) were assigned to the non-selected tasks. To ensure non-selected tasks were given equal weight, we averaged their original positions in the selection order. For example, if a strategy completes 4 tasks (occupying ranks 1, 2, 3, and 4), then the ranks of the 2 non-selected tasks (ranks 5 and 6) would be averaged (both 5.5) for data analysis.

Strategy analysis. We used two methods to assess participants' task selection strategies within each round. The first correlated actual performance with the optimal order, the urgency-first order, and the importance-first order. Higher correlations indicated a closer fit with that order. For the second method, we constructed linear models for each of the 9 task sets for each participant. Participants selection orders served as the dependent variable and the task set

features (deadlines and point values) as predictors; all predictors were standardized. Models with a variance inflation factor of 5 or greater were excluded, which accounted for about 1% of the models. Point value coefficients were multiplied by -1 so that larger coefficients represent higher priority given to high value tasks, matching the direction of the deadline coefficients where larger coefficients represent higher priority given to urgent tasks.

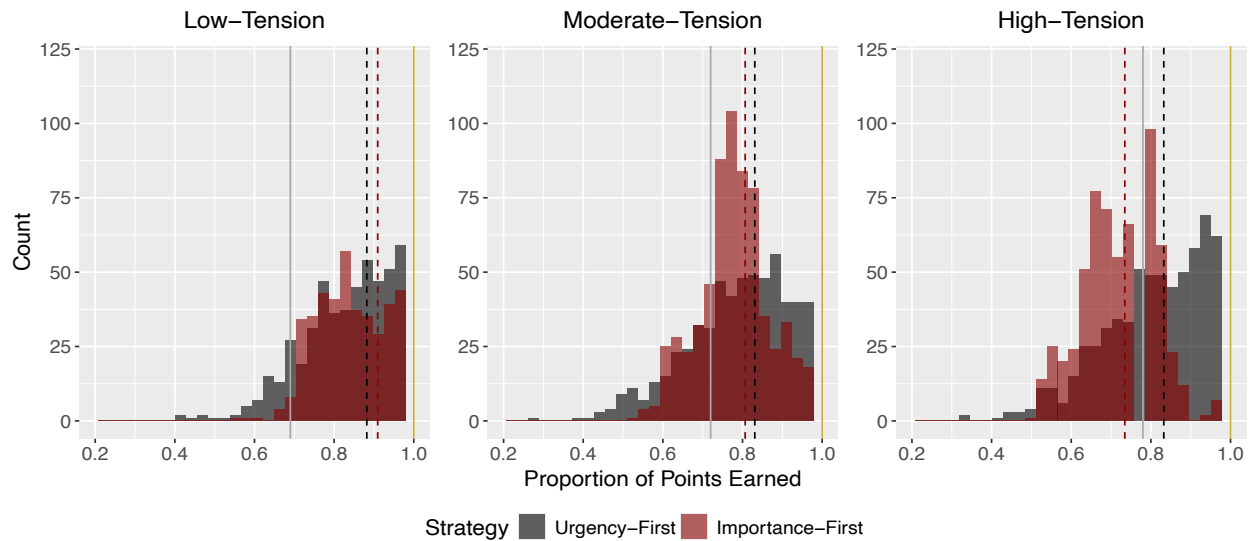
A priority score was calculated for each round by subtracting the deadline coefficient from the point coefficient. The direction of the priority score represented which task characteristic was prioritized higher, with negative values representing priority given to urgency and positive values indicating priority given to importance. The magnitude of the score indicated the degree to which urgency was prioritized over importance or vice versa (a score close to 0 indicated an equal balance of both).

Optimality of heuristics. To help understand the relationship between participants' strategies and performance, we used computer simulations to investigate the performance of each strategy if followed perfectly. We discovered that the effectiveness of each heuristic strategies varied depending on the tension environment (see Figure 2). While urgency-first ($M = 0.85$) outperformed importance-first ($M = 0.82$) overall, it was particularly more advantageous than importance-first in high tension environments. In moderate tension environments, neither strategy performed significantly better than the other, with urgency-first having only a slight advantage. Both strategies fared best in low tension environments, with importance-first taking a slight edge over urgency-first. It is important to note the high degree of overlap between the strategies in this environment. Since urgent tasks are often more important in the aligned conditions, the urgency-first and importance-first benchmarks often produce the same results. In this environment, one can prioritize either highly urgent or highly important tasks without

sacrificing the other characteristic, explaining why both strategies perform closest to each other and optimal. Please see Figure 2 for comparisons to optimal and random chance.

Figure 2

Histograms of Simulated Points Earned Across Tension Conditions



Note. The dotted lines indicate the mean of each strategy. The solid gray line indicates random chance, while the solid yellow line indicates optimal.

Exclusions

To ensure comparability across conditions, we sought to identify participants who responded randomly in order to complete the study quickly. Random earning benchmarks were compared to actual participant earnings to establish whether participants performed at or below chance. If a participant performed worse than chance on 7 or more out of the 9 rounds, they were excluded from analysis. This method excluded two types of participants: (a) those selecting at random and (b) those who selected 2 or fewer tasks each round.

We excluded 81 participants using the performance criterion explained above. Below-chance performance was largely driven by participants who completed very small numbers of

tasks. Excluded participants completed an average of 1.62 tasks (38% of maximum available tasks) across each round, compared to an average of 3.67 (86% of maximum available tasks) completed by included participants.

Exclusions disproportionately afflicted the planning group ($N = 57$ excluded compared to $N = 24$ for the selecting group). This may be explained by differences in structure across conditions. In the planning condition, it was possible to have errors in the selection order (e.g., planning to complete a task at time point 20 that had a deadline at time point 10), which participants were required to resolve before submitting. In the active selection condition, invalid orders were not possible because tasks became unavailable for selection after they expired. Thus, frustrated participants in the planning condition may have dismissed tasks in order to produce a valid (but highly suboptimal) order.

Results

We found that planning improved performance, as participants achieved a more optimal balance between urgency and importance, leading to a consistently higher proportion of points earned across urgency-importance tension environments. In contrast, selecting tasks without prior planning led to decreased performance due to an overemphasis on task importance, which resulted in particularly poor performance in high tension environments where urgency and importance were misaligned.

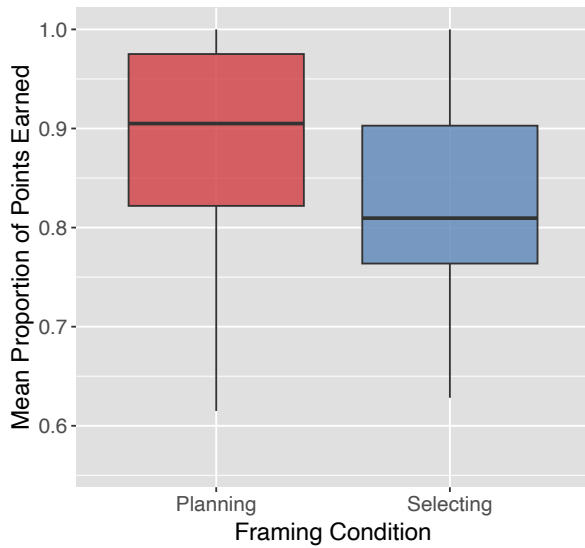
Turning to the results in more detail, we first examined overall performance by conducting Mann-Whitney U test on the proportion of points earned across framing conditions¹. Results revealed that the planning group ($Mdn = 0.90$) earned a significantly higher proportion of points compared to the selecting group ($Mdn = 0.81$), $U = 4368$, $p < .001$, $r = 0.29$ (Figure 3).

¹ We used non-parametric Mann-Whitney U tests because Shapiro-Wilkes tests revealed highly non-normal (negative skews) for many of these distributions.

Selectors ($Mdn = 0.63$) had a significantly higher priority score compared to planners ($Mdn = 0.34$), $U = 8665$, $p < .001$, $r = 0.28$. This indicates that both groups placed more emphasis on importance than on urgency. However, selectors prioritized importance over urgency to an even greater degree than planners.

Figure 3

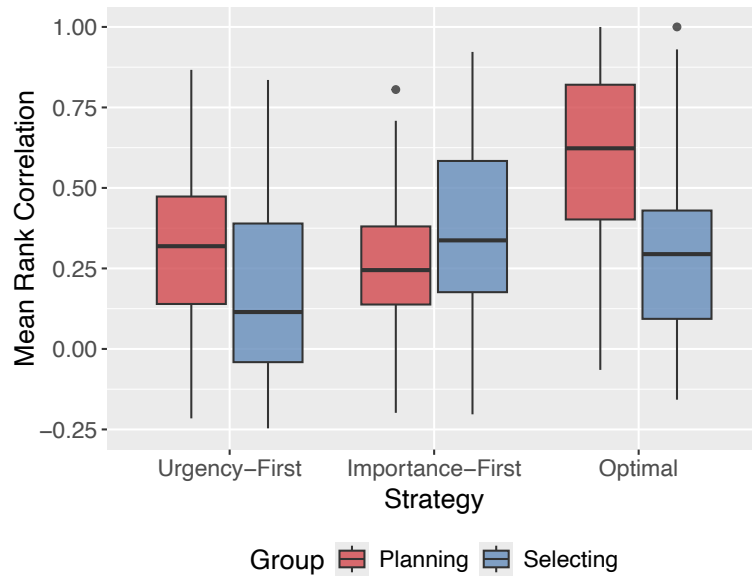
Proportion of Points Earned by Framing Condition



We next calculated the correlation between participants' actual rank orders and the three benchmarks described in Data Processing (urgency-first heuristic, importance-first heuristic, optimal order; see Figure 4). Results showed that planners ($Mdn = 0.32$) followed an urgency-first strategy closer than selectors ($Mdn = 0.10$), $U = 4761.5$, $p < .001$, $r = 0.25$. Conversely, planners ($Mdn = 0.25$) followed an importance-first strategy significantly less close than selectors ($Mdn = 0.34$), $U = 8240$, $p = .004$, $r = 0.19$. Finally, planners ($Mdn = 0.62$) followed the optimal significantly closer than selectors ($Mdn = 0.29$), $U = 3104.5$, $p < .001$, $r = 0.45$. These results highlight that selectors tended to *overprioritize* importance, while planners prioritized urgency and importance in a way that aligned closer to optimality.

Figure 4

Correlation to Each Strategy by Framing Condition



Urgency-Importance Tension Environments

Next, a series of linear mixed effects models, using the lmer (Bates et al., 2015) and emmeans (Lenth et al., 2018) R packages, examined how task framing (planners versus selectors) and urgency-importance correlation (low tension, moderate tension, or high tension) influenced performance. These models used five dependent variables: proportion of points earned, priority score, urgency-first rank correlation, importance-first rank correlation, and optimal rank correlation. For each dependent variable, two models were constructed. In Model 1, fixed effects included framing condition and tension condition, whereas in Model 2, the interaction between these variables was also included. A random intercept for each subject was included in all models. Results are summarized in Table 2.

First, we analyzed the proportion of points earned across conditions (Figure 5). Across conditions, planners performed better than selectors, ($B = -0.05, p < .001$).

Table 2*Results of Linear Mixed Models for Experiment 1*

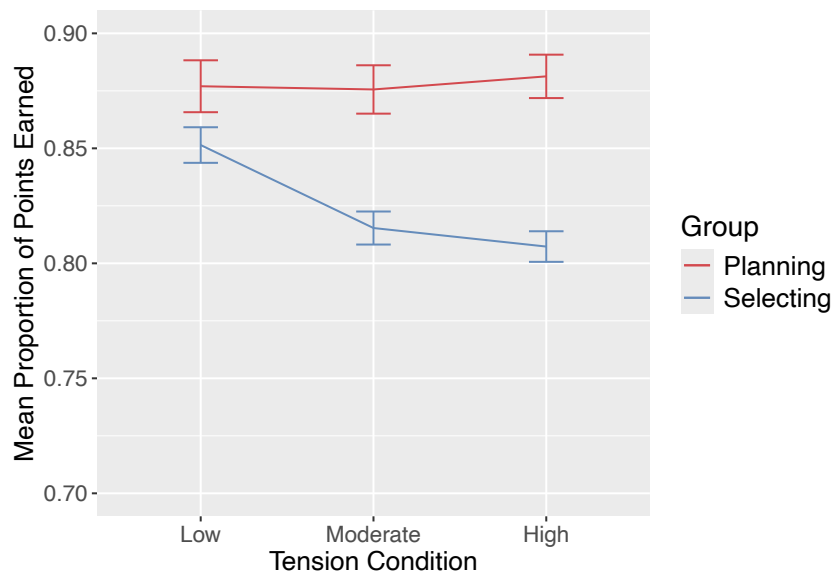
	Points Earned		Priority Score		Cor w/ Urgency		Cor w/ Importance		Cor w/ Optimal											
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>										
Predictor	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>										
Intercept	0.87***	0.88***	0.26***	0.28***	0.29***	0.29***	0.28***	0.30***	0.57***	0.58***										
LowTen	0.02**	0.00	0.39***	0.37***	0.09***	0.08	0.10***	0.15***	0.11***	0.06										
HighTen	-0.00	0.01	-0.13**	-0.18**	-0.11***	-0.10*	-0.17***	-0.26***	-0.12***	-0.10*										
Select	-0.05***	-0.06***	0.25***	0.22**	-0.13***	-0.13**	0.10**	0.07	-0.28***	-0.30***										
LowTen>Select		0.03*		-0.03		0.01		-0.08		0.08										
HighTen>Select		-0.1		0.09		0.01		0.16**		-0.03										
Random Effects	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>
Level 2 (ID)	0.09	0.01	0.19	0.01	0.32	0.10	0.32	0.10	0.23	0.05	0.23	0.05	0.18	0.03	0.18	0.03	0.21	0.04	0.21	0.04
Residual	0.13	0.02	0.13	0.02	0.78	0.61	0.78	0.61	0.50	0.25	0.50	0.25	0.47	0.22	0.47	0.22	0.47	0.22	0.47	0.22
ICC	0.30	0.30	0.14	0.14	0.18	0.18	0.18	0.18	0.13	0.13	0.13	0.13	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$. Moderate tension and the planning condition are the reference point. LowTen = low tension condition; HighTen = high tension condition; Select = selection condition. Model of best fit is italicized.

Participants earned more in the low tension condition than in the moderate tension condition ($B = 0.02, p = .001$), which was in turn similar to the high tension condition ($B = -0.00, p = .101$). The interaction term was found to be significant, $\chi^2(2) = 11.79, p = .003$. Post-hoc Tukey tests revealed that selectors performed best in the low tension conditions compared to the high tension ($B = -0.04, p < .001$) and moderate tension conditions ($B = -0.04, p < .001$), whereas planners' performance did not vary across conditions. This suggests that individuals in the planning condition may be better able to adapt their strategies across environments, leading to more consistent success.

Figure 5

Proportion of Points Earned Across Tension Conditions by Framing Condition



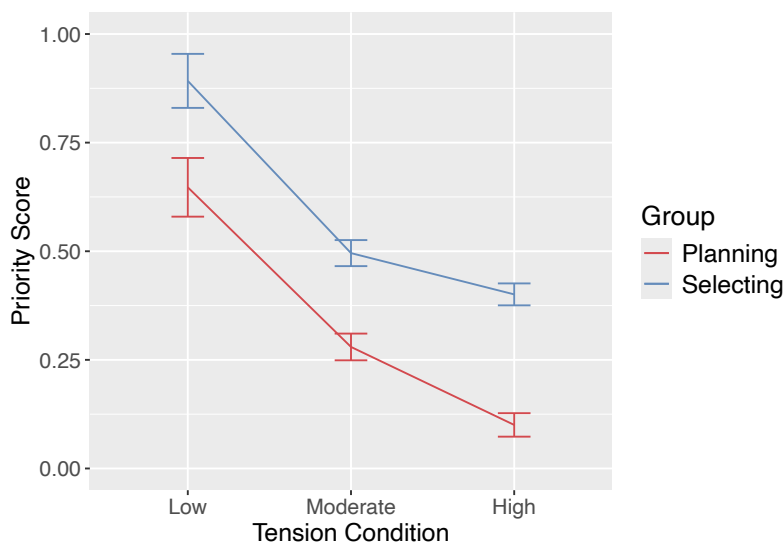
Note. Error bars represent ± 1 standard error of the mean.

Next, we examined whether participants relied more on one strategy relative to the other across tension environments. We found that the interaction term was not significant, $\chi^2(2) = 1.04, p = .595$. Priority scores leaned toward importance significantly more in the selecting condition than in the planning condition ($B = 0.25, p < .001$). It is important to note that despite the

difference in magnitude, priority scores in both framing conditions favoured importance over urgency. Participants prioritized important tasks to a greater degree in the low tension condition compared to the moderate tension conditions ($B = 0.39, p < .001$), which was greater in turn than in the high tension condition ($B = -0.13, p = .002$). Comparing Figure 6 to the theoretical performance of each heuristic in Figure 2, we see that participants prioritized important tasks to a greater degree when it was most adaptive (in the low tension environment) and less when it was most detrimental (in the high tension environment).

Figure 6

Priority Score Across Tension Conditions by Framing Condition



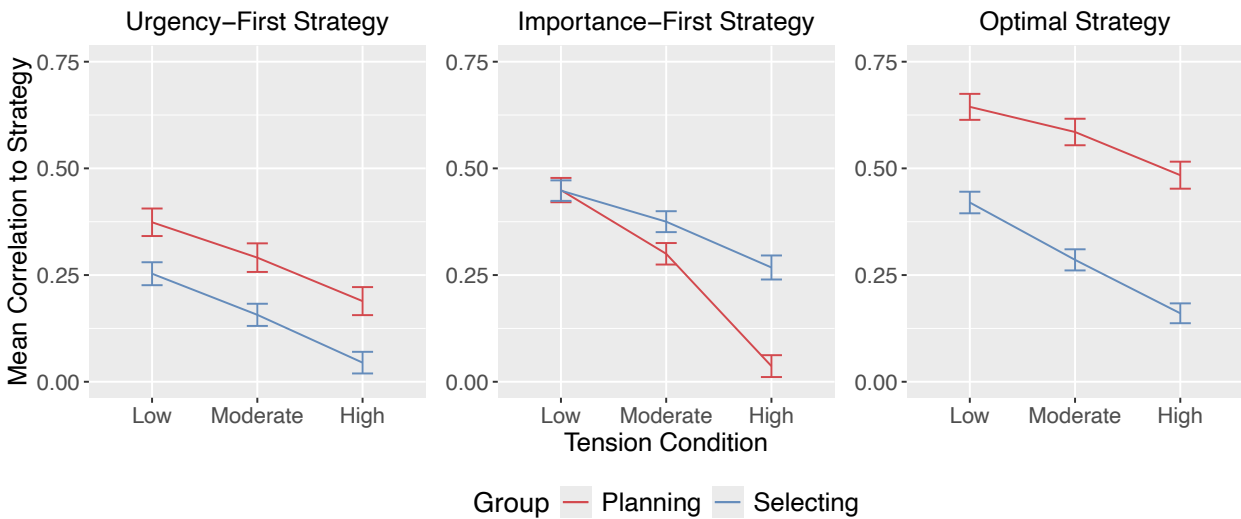
Note. Error bars represent ± 1 standard error of the mean. Priority scores above 0 indicate reliance on importance (> 0), while below one indicates reliance on urgency (< 0).

We next examined the conditions under which behaviour most resembled each strategy (Figure 6). The interaction term was found to be significant, $\chi^2(2) = 21.53, p < .001$. In the case of importance, behaviour was more similar to the importance-first benchmark in the active selection condition compared to the planning condition ($B = 0.10, p = .002$), as found in the priority score analysis. Further analysis showed that the relationship between behaviour and the

importance-first benchmark varied across tension conditions. The similarity between behavior and the importance-first benchmark was higher in the low tension than the moderate tension condition ($B = 0.10, p < .001$), which was in turn higher than the high tension condition ($B = -0.17, p < .001$). When we compare behaviour (Figure 7) to the performance of each heuristic (Figure 2), we can see that planners followed the importance-first heuristic more when the urgency-importance correlation made this strategy most profitable and less when it was least profitable, whereas selectors differentiated much less across correlation conditions. Both priority score and correlation to each strategy results align with the finding that planners earned a higher and more consistent proportion of points across tension correlations.

Figure 7

Correlation to Each Strategy Across Tension Condition by Framing Condition



Note. Error bars represent +/-1 standard error of the mean.

The remaining two strategies were best fit by non-interaction models. For urgency, behaviour was more similar to the urgency-first benchmark in the planning condition compared to the active condition ($B = -0.13, p < .001$), as found in the priority score analysis. The

similarity between behaviour and the urgency-first strategy was higher in the low tension condition than the moderate tension condition ($B = 0.09, p < .001$), which was higher in turn than the high tension condition ($B = -0.11, p < .001$).

Last we examined how closely behaviour resembled the optimal order. Behaviour was more similar to the optimal benchmark in the planning condition than in the selecting condition ($B = -0.28, p < .001$), which aligns with our previous finding that planners earned a higher proportion of points. The similarity between behaviour and the optimal benchmark was higher in the low tension condition compared to the moderate tension condition ($B = 0.11, p < .001$), which was higher in turn than the high tension condition ($B = -0.12, p < .001$).

To summarize, planners earned more points overall compared to selectors, particularly in low tension environments. Across all conditions, participants tended to prioritize importance relative to urgency, with selectors doing so more consistently than planners. Behaviour in the planning condition was better suited to each urgency-importance tension condition and aligned more closely with the optimal benchmark overall. Selectors reliably favoured importance regardless of the condition. Thus, planners demonstrated more adaptive strategies, while selectors more strictly followed the importance-first heuristic.

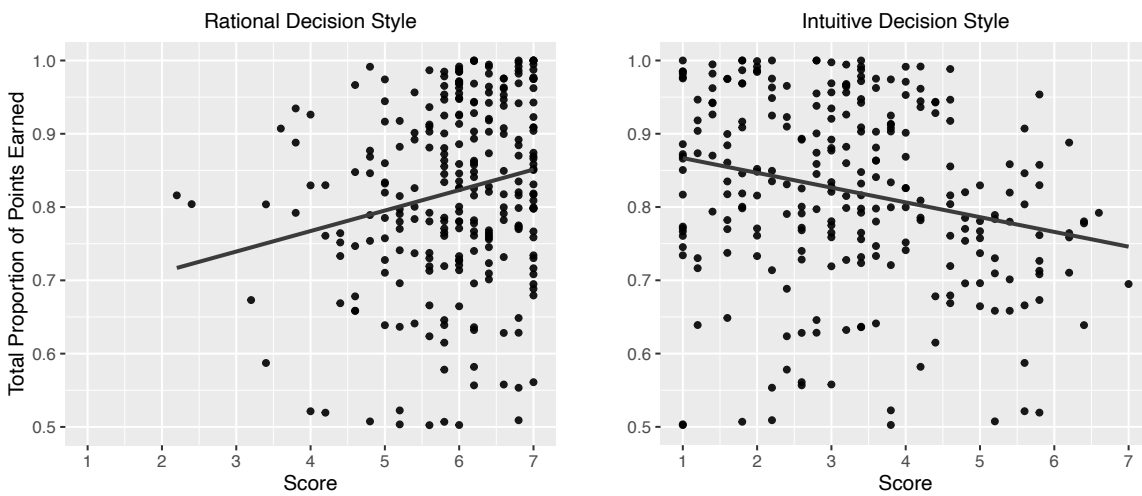
Psychometric Data

Last, we explored whether rational, intuitive, or maximizing decision-making tendencies were related to overall performance and strategy-use (Table 3). Scores were averaged across the 9 trials for each of the dependent variables, including proportion of points earned, priority score, urgency-first rank correlation, importance-first rank correlation, and optimal rank correlation. We found no relationship between maximization score and any of the five dependent variables.

Both subscales of the decisions style scale were found to have significant relationships with several outcome variables (see Figure 8). Rational decision-making was positively correlated with total proportion of points earned, $r(237) = 0.16, p = .012$, while intuitive decision-making was negatively correlated with total proportion of points earned, $r(237) = -0.25, p < .001$.

Figure 8

Relationship Between Proportion of Points Earned and Decision Styles Score in Experiment 1



Rational decision-making had a marginally positive correlation with urgency-first rank, $r(237) = 0.12, p = .056$, and a positive correlation with optimal rank, $r(237) = 0.17, p = .008$. In contrast, intuitive decision-making was negatively correlated with urgency-first rank, $r(237) = -0.23, p < .001$, and optimal rank, $r(237) = -0.25, p < .001$. This indicates that individuals with higher levels of rational decision-making produced task orders more similar to the urgency-first and optimal benchmarks, while those higher in intuitive decision-making tended to produce task orders less similar to those benchmarks. Neither decision style was related to importance-first

order or priority scores (see Table 3 for statistics), meaning none of the decision styles influenced the extent to which the importance-first or priority benchmarks were followed.

Table 3

Correlation Matrix for Experiment 1

Measure	1	2	3	4	5	6	7
1. MS	—						
2. DSS-R	0.24***	—					
3. DSS-I	-0.01	-0.46***	—				
4. P-Score	-0.09	-0.07	0.09	—			
5. Urg-First	0.08	0.12	-0.23***	-0.60***	—		
6. Imp-First	-0.11	-0.03	0.05	0.79***	-0.39***	—	
7. Opt Rank	-0.00	0.17**	-0.25***	-0.13*	0.49***	0.05	—
8. Points	0.09	0.16**	-0.25***	-0.40***	0.64***	-0.23***	0.73***

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$. MS = maximization score; DSS-R = rational decision style score; DSS-I = intuitive decision style score; P-Score = priority score; Urg-First = correlation to urgency-first rank order; Imp-First = correlation to importance-first rank order; Opt Rank = correlation to optimal rank order; Points = proportion of points earned.

It is important to note the strong correlation between the urgency-first rank and the optimal rank, $r(237) = 0.49, p < .001$. It is possible that rational decision makers were following the optimal strategy, which happened to be correlated with the urgency-first rank, $r = 0.25$. Meanwhile, intuitive decision makers may not have adhered to any specific strategy, potentially explaining why their intuitive decision-making score was negatively correlated with both the optimal and urgency-first orders, while uncorrelated with importance-first rank or priority scores.

Further analysis revealed that the effect of decision style on performance was more pronounced in the active selection condition compared to the planning condition. In the active selection condition, rational decision-making was positively correlated with performance ($r = 0.18$), while intuitive decision-making was negatively correlated ($r = -0.26$). In the planning

condition, a similar pattern was observed: rational decision-making was positively correlated with performance ($r = 0.11$), and intuitive decision-making was negatively correlated ($r = -0.17$), however the effects were less pronounced. This suggests that people high in rational decision-making might naturally adapt their strategies to suit their environment. Meanwhile those who rely on intuition may be more swayed by gut feelings and external cues, which leads to suboptimal decision-making in situations where they cannot revise their choices after committing to them.

Discussion

In Experiment 1, we showed that, as expected, participants performed better when scheduling tasks rather than actively selecting them. Planners earned a greater proportion of points overall, which was consistent across tension environments. Individuals given the opportunity to plan found a more optimal balance between urgency and importance, alluding to the value of scheduling. Those actively selecting tasks earned fewer points overall and relied on heuristic-based decisions to select tasks. Contrary to prior findings, individuals overprioritized importance, not urgency. In fact, nearer-deadline tasks were often sacrificed altogether for high-point value tasks. This was demonstrated through fluctuating success across urgency-importance tension environments, where success was highest when urgency and importance did not align and lowest when they most aligned.

We additionally found that rational decision-making was positively associated with both urgency-focused and optimal task strategies. In contrast, intuitive decision-making was negatively associated with both urgency-focused and optimal strategies. The decision styles also affected overall performance, with rational decision-makers generally achieving a higher proportion of points, while more intuitive decision-makers earned fewer points. One interesting

finding was that decision-making styles had a smaller impact on performance in the planning condition compared to the selection condition. This suggests that creating a schedule might help those that tend to rely on their intuition override their impulsive tendencies and improve overall task management abilities.

One possible explanation for why importance was prioritized over urgency could be the psychological distance baked into our paradigm, which we did not originally consider. Due to the hypotheticality of the tasks in our experiment, participants might have felt detached from real task selection processes, as they could “complete” their task lists quickly and with ease. The psychological distance inherent to the game may have prompted desirability considerations (Liberman & Trope, 2014), reorienting focus onto tasks of high-point value. At this distance, feasibility concerns, such as looming deadlines, may have been eclipsed by the desirability of accumulating more points. Thus, in both conditions, there was an overemphasis on task importance at the expense of urgency. Although this effect was less pronounced when participants had the opportunity to plan or revise their task sequences. By allowing for within-round trial and error, participants were better equipped to discover strategies that maximized overall points at the list level rather than at the task level. This flexibility was not available to participants who had to make immediate task selections, which inherently locked in any initial intuitive decisions, such as prioritizing high-point value tasks.

Ultimately, scheduling tasks prior to completing them appears to lead to more optimal outcomes. Individuals striving for improved task management should feel encouraged to create task schedules rather than reacting to task lists in real-time. This reigns especially true for those who rely on their intuition when making decisions. By listing tasks and assigning subjective urgency and importance scores—similar to the approach used in the Eisenhower Matrix—people

can more effectively manage their tasks (Jyothi & Parkavi, 2016). This method promotes thoughtful consideration and intention, helping to systematically prioritize tasks and enhance overall productivity.

Experiment 2

Experiment 2 expands upon the task management game in the previous experiment to examine the role of task segmentation (grouped versus ungrouped turns) and task comparability (fixed versus varied task lengths) on the same primary metrics: (i) the trade-off weight attributed to urgency and importance and (ii) overall performance (measured by points earned relative to the point-maximizing [optimal] task order). We anticipate that ungrouped task turns will lure decision-makers away from the overarching goal to maximize total number of points, leading to diminished overall performance. However, we do not expect task-switching to influence the relative weight attributed to urgency or importance. Upon introducing varied task lengths, we expect people to prioritize shorter tasks over urgent or important ones, with performance suffering as a result. As in the previous experiment, we predict performance to stray further from optimal as the tension between urgency and importance increases.

Method

Participants

We recruited 603 U.S. participants (319 male, 267 female, 17 other, $M_{\text{age}} = 43$) from MTurk to complete an approximately 12-minute study in exchange for \$1.00 USD plus a bonus payment of up to \$1.00 USD, depending on performance. 120 participants were excluded based on the exclusion criterion outlined below in *Exclusions*.

Design and Procedure

In the current iteration of the paradigm, each condition was adapted to align with what was previously referred to as the selecting condition (i.e., tasks were selected across turns, representing time points).

Figure 9

Task Management Game Displays in Experiment 2



Note. Task management game displays in Study 2: (A) fixed task lengths and grouped task turns, (B) varied task lengths and grouped task turns, (C) fixed task lengths and ungrouped task turns, (D) varied task lengths and ungrouped task turns.

Participants were randomly assigned to one of four groups, comprised of a segmentation variable (grouped or ungrouped) and a task comparability variable (fixed or varied; see Figure 9 for game displays). In the grouped condition, tasks were completed as a whole upon selection, precluding the possibility to task-switch. In the ungrouped condition, task could be partially completed, making it possible to task-switch.

Task lengths were either fixed (task lengths were always 5) or varied (task lengths were drawn from a uniform distribution ranging from 2 to 8). In the varied length condition, the total number of turns required to complete the task list within each round summed to 30, consistent with the previous experiment. Deadlines were drawn from a uniform distribution ranging from 2 to 30, with the constraint that task lengths could not exceed their respective deadlines. Point values were drawn from the same normal distribution in the previous experiment, adjusted in proportion to their respective task length as follows: $M = 225 * (\text{length} / 5)$, $SD = 75 * (\text{length} / 5)$. In this condition, two extra columns were added—“Turns Left” and “Turns Taken”—to help participants track the number of turns already taken on each task.

Data Processing

Benchmarks. Several new strategies were possible with the introduction of varied lengths. We calculated additional benchmarks to compare to participant behaviour to add to the lineup of benchmarks calculated in the previous study (urgency-first, importance-first, and optimal). The slack-first strategy prioritized the task with the lowest slack (slack = deadline - turns left), contrary to the urgency-first strategy which prioritized the task with the nearest absolute deadline, irrespective of slack². The unit-importance strategy prioritized the task with the highest point value per turn (unit-importance = points/length), while the absolute-importance strategy

² Slack-first and urgency-first orders were highly correlated ($r = 0.88$) and therefore the slack-first order was dismissed from further analysis.

(previously referred to as just the importance strategy) prioritized the task with the highest point value, regardless of unit importance. The shortest-length strategy prioritized the task with the shortest length. These three additional benchmarks apply *only* to the varied length conditions. When task lengths are fixed, the unit-importance heuristic is identical to the absolute-importance heuristic and the slack-first heuristic is identical to the urgency-first heuristic. The shortest-length heuristic simply is inapplicable to task lists where tasks are of equal length.

Task switching. We calculated the number of times a participant switched from “working on” an incomplete task to working on a new task and averaged the number of task switches across all rounds to produce an “average task switches” score. Behaviour was only considered as task switching when the switch occurred *before* the task was complete. Working on a new task upon the completion of another was not classified as a task switch.

Exclusions

The random prioritization algorithm was updated to reflect behaviour in the ungrouped turn selection conditions. Instead of selecting tasks at random, task *turns* were selected at random. However, randomly selecting by turn produced strikingly low and unrealistic benchmarks that did not capture the “random” or aimless human behaviour we intended to target ($M = 48\%$). As such, simulations were pseudo-randomized so that majority of the time, turns were selected from the same task until it was complete, but sometimes, a random deviation from the current task to a new task would occur (i.e., a task switch). The number of task switches in the algorithm was based upon a post-hoc calculation of the average number of task switches in the sample (1.76 times or approximately 6% of the time), excluding those in the grouped turns condition. To briefly summarize, this algorithm selected tasks at random and task switched approximately 6% of the time. This was applied to the ungrouped task turn condition only. The

original algorithm—which selected tasks at random without task switching—was applied to the grouped turns condition. As in the previous experiment, if a participant performed worse than chance (by these metrics) on 7 or more out of the 9 rounds, their data was excluded from the analysis.

Exclusions disproportionately affected participants in ungrouped turns group, regardless of task length condition (N = 83 excluded compared to N = 37 for the by-task group; see Table 4)³. This may be explained by differences in the structure of each experimental group. Each element added to the game complicates the process and may overload cognitive systems. As a result, when task turns were ungrouped and task lengths varied, performance dramatically worsened (even after removing participants whose performance was worse than chance).

Table 4

Number of Participants Excluded from Analysis by Condition

	Grouped	Ungrouped	Total
Uniform Lengths	<i>N</i> = 18	<i>N</i> = 37	<i>N</i> = 55
Varied Lengths	<i>N</i> = 19	<i>N</i> = 46	<i>N</i> = 65
Total	<i>N</i> = 37	<i>N</i> = 83	<i>N</i> = 120

Note. Random chance benchmarks were simulated by task for the grouped turns condition and by task with task-switching for the ungrouped turns condition.

Results

We found that both manipulations (segmentation and task comparability) decreased overall performance, as measured by the proportion of points earned. The segmentation

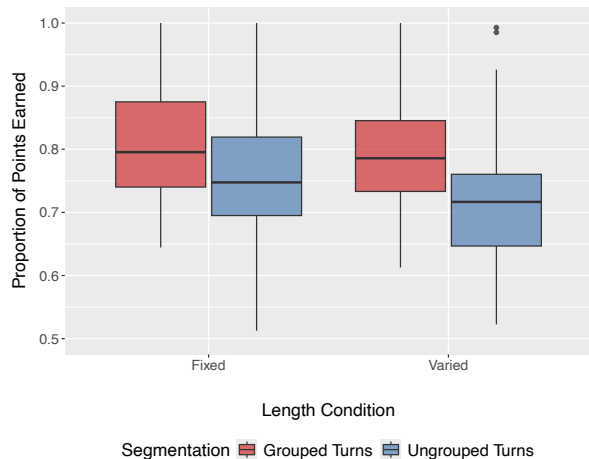
³ Please note that the pattern of results did not change when analyses were run on the full sample (before exclusions).

manipulation did this due to the option to switch tasks, which, in this experiment, was always suboptimal. However, it did not shift the relative weight given to urgency versus importance. Variable task lengths, compared to fixed task lengths, led to decreased performance because participants tended to overemphasize absolute importance relative to task length.

First, we looked at overall performance by analyzing the proportion of points earned across conditions (see Figure 10). A two-way ANOVA revealed a main effect of segmentation condition (grouped vs. ungrouped), $F(1, 479) = 72.54, p < .001, \eta^2 = 0.13$, and task length (fixed vs. varied), $F(1, 479) = 10.27, p = .001, \eta^2 = 0.03$. The interaction was not significant, $F(1, 479) = 1.45, p = .229$. The mean earnings in the grouped turns condition ($M = 0.80, SD = 0.09$) was significantly higher than in the ungrouped condition ($M = 0.72, SD = 0.13$). While the mean earnings in fixed task length conditions ($M = 0.78, SD = 0.12$) was significantly higher than in the varied task length conditions ($M = 0.75, SD = 0.10$). Participants performed best when in the condition where turns were grouped and task lengths were fixed; while they performed worst in the condition where turns were ungrouped and task lengths varied. Overall, performance seemed to suffer as additional game features were introduced.

Figure 10

Proportion of Points Earned by Task Segmentation and Length Conditions



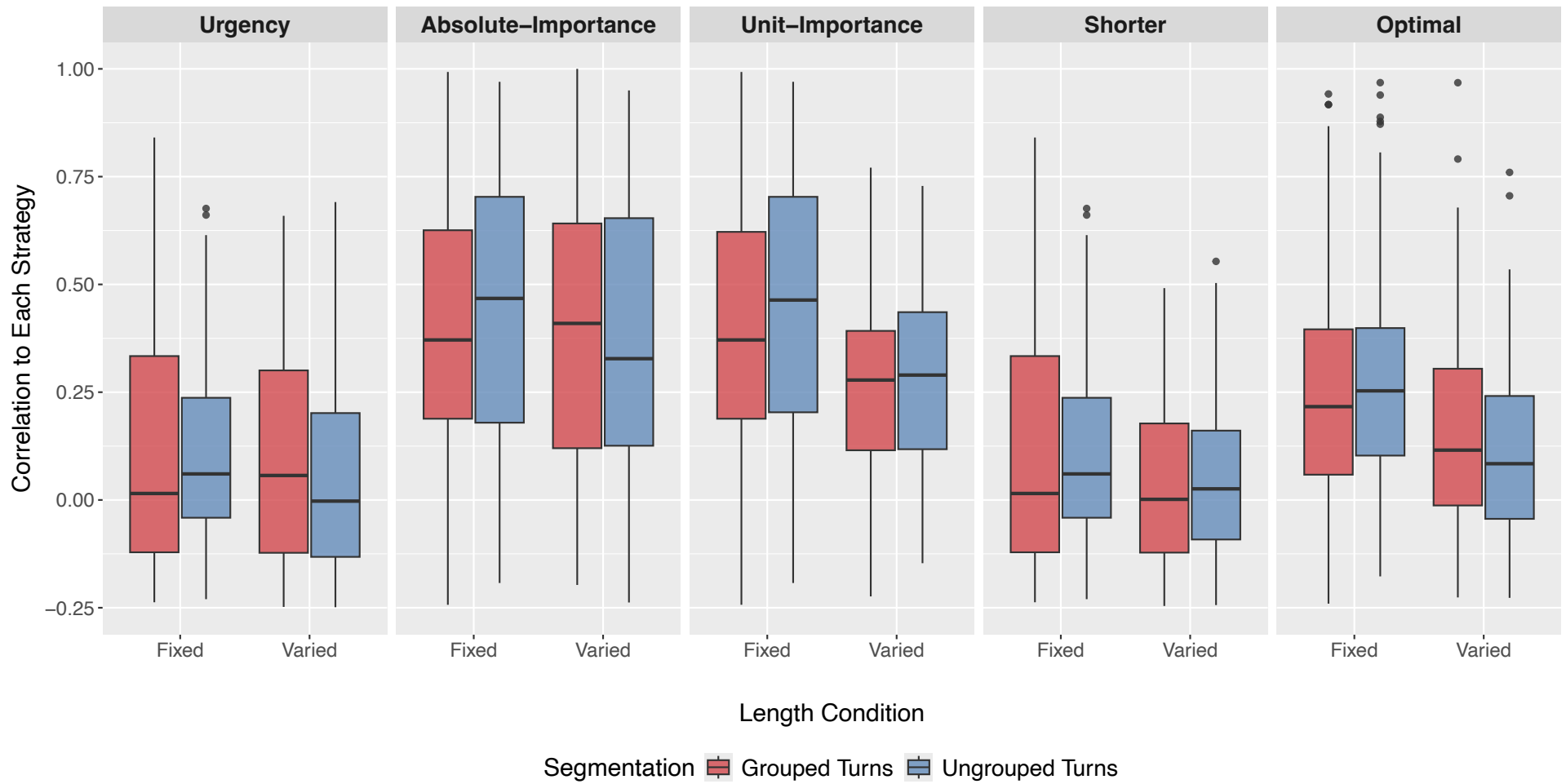
Next, we calculated the correlation between participants' actual rank orders and the five benchmarks⁴ (urgency-first heuristic, absolute-importance-first heuristic, unit-importance first heuristic, shorter-first heuristic, and the optimal order; see Figure 11). Beginning with urgency-first, a two-way ANOVA revealed no significant interaction between segmentation and task length, $F(1, 479) = 1.05, p = .307$. Result also indicated no effect of segmentation on correlation to the urgency-first benchmark, $F(1, 479) = 0.91, p = .341$. However, there was a main effect of task length on correlation to the urgency-first benchmark, $F(1, 479) = 14.62, p < .001, \eta^2 = 0.03$. Participants in the fixed length condition ($M = 0.09, SD = 0.30$) behaved more similar to the urgency-first benchmark compared to those in the varied length condition ($M = -0.01, SD = 0.30$). To summarize, when task lengths were fixed, people tended to prioritize closer deadlines compared to when task lengths varied.

Turning to absolute importance-first, we did not find a significant interaction, $F(1, 479) = 1.99, p = .159$. We also found no effect of task segmentation on the degree to which participants behaved similar to the absolute-importance-first benchmark, $F(1, 479) = 0.67, p = .414$. Similarly, there was no effect of task length on the absolute importance-first strategy, $F(1, 479) = 1.85, p = .175$. With respect to unit-importance, we did not find a significant interaction, $F(1, 479) = 0.38, p = .541$, nor an effect of segmentation, $F(1, 479) = 3.41, p = .066$; however, there was a main effect of task length, $F(1, 479) = 35.98, p < .001, \eta^2 = 0.07$. Participants in the uniform length condition ($M = 0.41, SD = 0.30$) behaved more similar to the unit-importance benchmark than those in the varied length condition ($M = 0.27, SD = 0.22$). It is important to note that the unit-importance and absolute importance strategy are identical in the uniform length condition and cannot be compared.

⁴ There was no main effect of group on priority scores: task segmentation, $F(1, 459) = 0.59, p = .444$, task comparability, $F(1, 459) = 2.83, p = .093$. Priority scores were removed from the following analyses.

Figure 11

Correlation to Each Strategy by Task Segmentation and Length Conditions



Note. Fixed length conditions are identical across (i) absolute-importance and unit-importance panels and (ii) urgency-first and shorter length panels.

However, it is interesting to note that when task lengths varied, the unit importance strategy was not as good a fit compared to the absolute importance strategy.

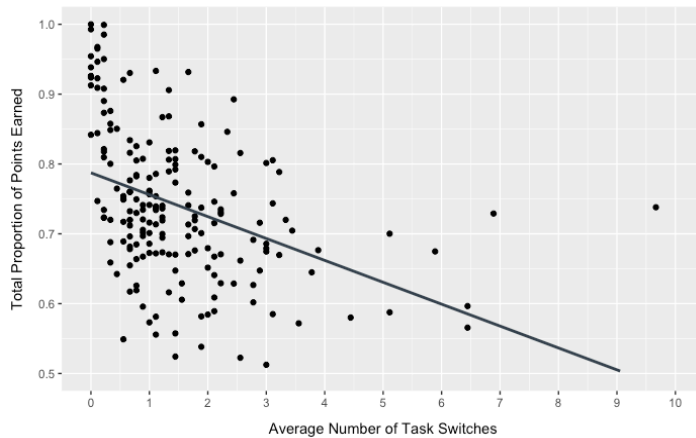
Last, we compared participant behaviour to the optimal order. We found that there was no effect of task segmentation on similarity to the optimal order, $F(1, 479) = 0.50, p = .481$. However, there was an effect of task length, $F(1, 479) = 30.11, p < .001, \eta^2 = 0.06$, and a significant interaction between task segmentation and task length, $F(1, 479) = 5.23, p = .023, \eta^2 = 0.01$. Participants completing tasks in parts (ungrouped condition) and with varied task lengths ($M = 0.09, SD = 0.22$) were further from the optimal order benchmark compared to those selecting whole tasks (grouped condition) with fixed task lengths ($M = 0.23, SD = 0.27$) and those selecting tasks in parts (ungrouped condition) with fixed task lengths ($M = 0.28, SD = 0.27$). To summarize, completing tasks as segmented parts versus unsegmented wholes did not impact the closeness to optimality; however, varying task lengths led to a decline in optimality.

Task-Switching

To further examine what may have caused the difference in performance, we analyzed the relationship between average task switches and performance (see Figure 12). This analysis focused on participants in the ungrouped task turns condition. On average, we found that participants task switched 1.54 times per round. There was no significant difference found in the proportion of task switches across fixed versus variable length conditions, $t(268.23) = 1.14, p = .256$. However, there was a significant negative correlation found between average task switches and proportion of points earned, $r(208) = -0.30, p < .001$. This indicates that participants who task switched more frequently tended to earn a lower proportion of points overall.

Figure 12

Relationship Between Task Switching and Performance



Task Tension Conditions

Next, we ran a series of linear mixed effects models examining how turn segmentation (grouped versus ungrouped), task comparability (fixed lengths versus varied lengths), and urgency-importance tension condition (low tension, moderate tension, and high tension) influenced performance. The models used the following six dependent variables: proportion of points earned, urgency-first rank correlation, importance-first rank correlation, unit-importance first correlation, shortest-first correlation, and optimal rank correlation. For each dependent variable, two models were constructed. In Model 1, fixed effects included segmentation condition, task comparability condition, and tension condition, while Model 2 included an interaction between these variables. A random intercept for each subject was included in all models. Results are summarized in Table 5.

For each dependent variable, the addition of an interaction term did not improve the fit relative to the model without it, and thus each of the following analyses will reflect main effects. We first analyzed the proportion of points earned across conditions.

Table 5*Results of Linear Mixed Models for Experiment 2*

	Points Earned				Cor w/ Urgency				Cor w/ Absolute-Importance			
	<i>Model 1</i>		Model 2		<i>Model 1</i>		Model 2		Model 1		Model 2	
Predictor	<i>B</i>		<i>B</i>		<i>B</i>		<i>B</i>		<i>B</i>		<i>B</i>	
Intercept	0.82***		0.81***		0.05*		0.05		0.41***		0.39***	
LowTen	0.01*		0.02		0.15***		0.11**		-0.03		-0.04	
HighTen	-0.02***		-0.00		-0.06***		-0.05		-0.00		0.04	
Ungrouped	-0.08***		-0.06***		-0.03		-0.03		0.02		0.08	
Varied	-0.03**		-0.01		-0.09***		-0.07		-0.04		-0.01	
LowTen:Ungrouped			0.00				0.07				-0.00	
HighTen:Ungrouped			-0.03				-0.04				-0.06	
LowTen:Varied			0.00				0.05				0.06	
HighTen:Varied			-0.03*				-0.06				-0.04	
Ungrouped:Varied			-0.03				-0.06				-0.05	
LowTen:Ungrouped:Varied			-0.03				-0.05				-0.12	
HighTen:Ungrouped:Varied			0.04				0.13				0.02	
Random Effects	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>
Level 2 (ID)	0.09	0.01	0.09	0.01	0.23	0.05	0.23	0.05	0.27	0.07	0.27	0.07
Residual	0.16	0.03	0.16	0.03	0.50	0.25	0.50	0.25	0.45	0.20	0.45	0.20
ICC	0.25		0.25		0.18		0.18		0.26		0.26	

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$. Moderate tension, grouped task turns, and the fixed task length condition are the reference point. LowTen = low tension condition; HighTen = high tension condition; Ungrouped = ungrouped task turns; Varied = varied task turns. Model of best fit is italicized.

Table 5 Continued (1)*Results of Linear Mixed Models for Experiment 2*

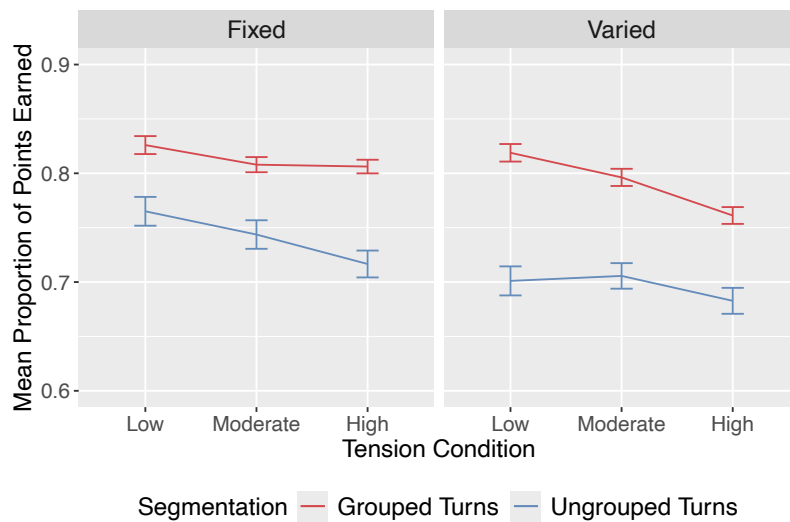
	Cor w/ Unit-Importance		Cor w/ Shortest Length		Cor w/ Optimal							
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 1</i>	<i>Model 2</i>						
Predictor	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>						
Intercept	0.26***	0.24***	-0.09***	-0.06*	0.27***	0.25***						
LowTen	-0.05	-0.02	0.11***	0.06	0.06***	0.45						
HighTen	0.03	0.04	-0.01	-0.04	-0.11***	-0.09**						
Ungrouped	0.04	0.06	0.05	-0.00	-0.01	0.06						
Varied	—	—	—	—	-0.12***	-0.08						
LowTen:Ungrouped		-0.06		0.11*		0.02						
HighTen:Ungrouped		-0.02		0.06		-0.06						
LowTen:Varied		—		—		0.03						
HighTen:Varied		—		—		-0.03						
Ungrouped:Varied		—		—		-0.13*						
LowTen:Ungrouped:Varied		—		—		-0.03						
HighTen:Ungrouped:Varied		—		—		0.10						
Random Effects	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>	<i>SD</i>	<i>Var</i>
Level 2 (ID)	0.15	0.02	0.15	0.02	0.16	0.03	0.17	0.3	0.19	0.04	0.19	0.04
Residual	0.47	0.22	0.47	0.22	0.45	0.20	0.45	0.20	0.48	0.23	0.48	0.23
ICC	0.09	0.09	0.12	0.12	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$. Moderate tension, grouped task turns, and the fixed task length condition are the reference point. LowTen = low tension condition; HighTen = high tension condition; Ungrouped = ungrouped task turns; Varied = varied task turns. Model of best fit is italicized.

Participants performed better in the grouped turn condition than in the ungrouped turn condition ($B = -0.08, p < .001$) and when in the fixed length condition compared to the varied length condition ($B = -0.03, p = .001$). Similar to Experiment 1, performance was better in the low tension condition compared to the moderate tension condition, ($B = 0.01, p = .013$), which was in turn better than the high tension condition, ($B = -0.02, p < .001$).

Figure 13

Proportion of Points Earned by Segmentation, Length, and Tension Conditions



Note. Error bars represent +/-1 standard error of the mean.

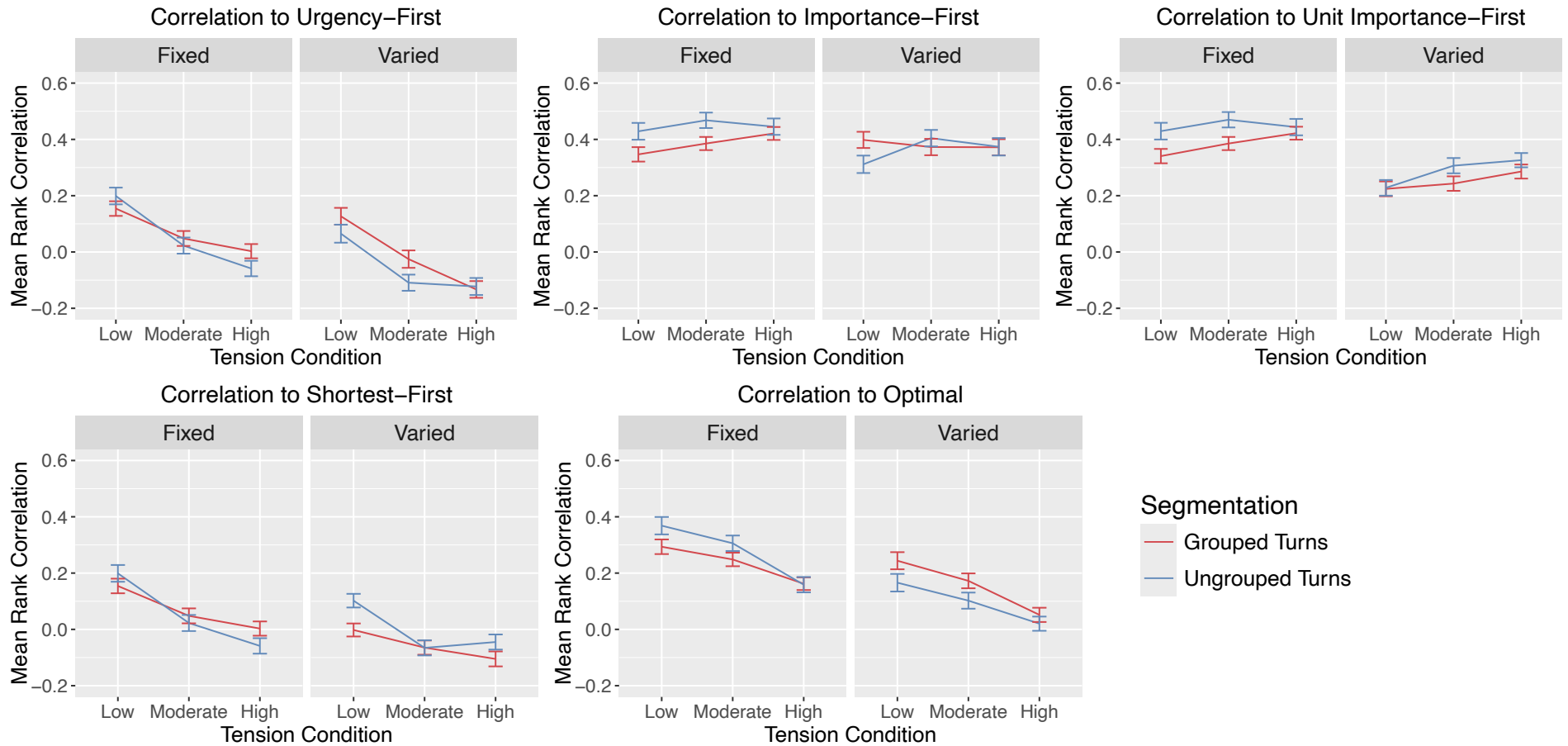
Next, we examined the degree to which behaviour resembled each strategy across urgency-importance (tension) environments (Figure 14). In the case of the urgency, behaviour was more similar to the urgency-first benchmark in the fixed length condition compared to the varied length condition ($B = -0.09, p < .001$). There was no difference in similarity to urgency-first across segmentation conditions ($B = -0.03, p = .252$). As in Experiment 1, the similarity between behaviour and the urgency-first benchmark was higher in the low tension condition than the moderate tension condition ($B = 0.15, p < .001$), which was in turn higher than the high tension condition, ($B = -0.06, p < .001$). When comparing behaviour to the shortest-length

benchmark, we found that it was more closely aligned with this benchmark in the high tension condition than the moderate tension condition ($B = 0.11, p < .001$). The moderate tension condition, in turn, was similar to the low tension condition ($B = -0.01, p = .632$). There was no difference in similarity to urgency-first across grouping conditions ($B = 0.05, p = .057$). Despite the increased alignment with the shortest length heuristic in the high tension condition, the mean hovered slightly above 0, indicating that overall behaviour did not resemble this heuristic overall. The degree to which behaviours resembled the absolute and unit importance-first benchmark did not differ across any condition.

Last, we analyzed how closely behaviour resembled the optimal order. Participant behaviour was more similar to the optimal order in the fixed length condition compared to the varied length condition ($B = -0.12, p < .001$). There was no difference in similarity to the optimal order across the grouping conditions ($B = -0.01, p = .711$). As in Experiment 1, behaviour was most similar to optimal in the low tension condition compared to the moderate tension condition ($B = 0.06, p < .001$), which in turn was more similar than the high tension condition ($B = -0.11, p < .001$).

Figure 14

Correlation to Strategy by Task Segmentation, Length, and Tension Conditions



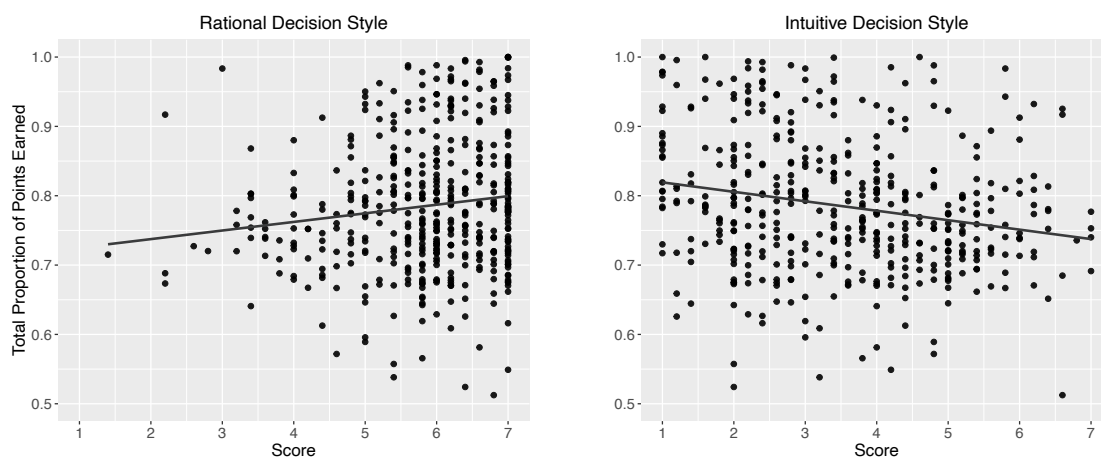
Note. Error bars represent +/-1 standard error of the mean. In uniform length conditions, correlation to importance-first and unit importance-first are identical, as are correlation to smallest-first and urgency-first.

Psychometric Data

Last we examined the relationship between decision-making tendencies (rational, intuitive and maximizing) and performance outcomes (see Table 6). As in Experiment 1, scores were averaged across the 9 trials for each of the dependent variables, including proportion of points earned, urgency-first rank correlation, importance-first rank correlation, unit-importance-first rank correlation, shortest-first rank correlation, and optimal rank correlation. We once again found no relationship between maximization score and any of the six dependent variables. Although, we did find that both subscales of the decisions style scale had significant relationships with proportion of points earned. Like in Experiment 1, rational decision-making was positively correlated with total proportion of points earned, $r(481) = 0.16, p < .001$, while intuitive decision-making was negatively correlated with total proportion of points earned, $r(481) = -0.20, p < .001$.

Figure 15

Relationship Between Proportion of Points Earned and Decision Styles Score in Experiment 2



Compared to Experiment 1, the relationship between the decision styles and benchmarks were less pronounced. Rational decision making was positively correlated with the optimal rank, $r(481) = 0.13, p = .003$, indicating that individuals with a higher tendency for rational decision-

making more closely aligned with the optimal benchmark. Intuitive decision-making was, again, negatively correlated with both the urgency-first strategy, $r(481) = -0.15, p < .001$, and the optimal strategy, $r(481) = -0.22, p < .001$. This suggests that individuals with higher intuitive decision-making scores deviated more from the optimal order. Once again, neither decision style had a significant relationship with the absolute- or unit-importance-first strategy.

Table 6

Correlation Matrix for Experiment 2

Measure	1	2	3	4	5	6	7	8
1. MS	—							
2. DSS-R	0.19***	—						
3. DSS-I	0.14**	-0.40***	—					
4. Urg-First	-0.04	0.05	-0.15***	—				
5. A.Imp-First	-0.02	0.07	0.00	-0.52***	—			
6. U.Imp-First	0.01	0.04	0.00	-0.40***	0.82***	—		
7. Short-First	-0.04	-0.02	-0.11*	0.74***	-0.54***	-0.34***	—	
8. Opt Rank	-0.01	0.13**	-0.22***	0.46***	0.10*	0.15***	0.25***	—
9. Points	-0.06	0.16***	-0.20***	0.50***	-0.10*	-0.13**	0.23***	0.62***

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$. MS = maximization score; DSS-R = rational decision style score; DSS-I = intuitive decision style score; Urg-First = correlation to urgency-first rank order; A.Imp-First = correlation to absolute importance-first rank order; U.Imp-First = correlation to unit importance-first rank order; Short-First = correlation to shortest task first order; Opt Rank = correlation to optimal rank order; Points = proportion of points earned.

Discussion

In Experiment 2, we found that, consistent without our prediction, task segmentation versus unsegmentation did not affect task prioritization strategies, but it did impact performance. The decline in performance was tied to the frequency of task switches, with more switching relating to a lower proportion of points earned. This supports previous research showing the negative effects of task-switching on performance (Monsell, 2003; Kiesel et al., 2010; Wylie & Allport, 2000). While we did not directly examine traditional “switch costs” as the time taken to mentally adjust to a new task (Rubinstein et al., 2001), our findings highlight that the opportunity for frequent task-switching led to suboptimal performance.

Contrary to our predictions, participants did not shift toward selecting shorter tasks despite the performance decrements observed in the varied length condition. In fact, our results demonstrated that the pull of either shorter tasks or urgent tasks were not relevant to participants' task selection choices. Instead, the critical task characteristic driving behavior was, once again, higher point values (i.e., absolute importance). This was confirmed by further analysis, which correlated participants' choices with each simulated order across urgency-importance tension conditions. While participants consistently prioritized importance across different tension environments, they rarely prioritized urgency, except for a slight increase in the low-tension environment, where urgency and importance heavily overlapped. In conditions where urgency and importance were most misaligned, there was a noticeable decline in the prioritization of urgency.

As in Experiment 1, we found that rational decision-making was associated with more optimal urgency and importance trade-offs and greater overall performance, while intuitive decision-making was linked to less optimal trade-offs and worse overall performance. Intuitive

decision-making was also associated with a reduced alignment to the urgency-first strategy, indicating that urgency was not being prioritized. However, the significant correlation between the optimal order and the urgency-first benchmarks complicates the interpretation, making it difficult to determine whether participants were relying on an optimal strategy that coincidentally aligned with the urgency-first strategy, or vice versa. This is a limitation that will be addressed in future studies.

These findings suggest that extrinsic features, such as whether tasks are completed in parts or in one continuous push, can significantly affect performance outcomes. Individuals aiming for improved task management should be mindful of the negative impact of frequent task-switching and the potential bias to focus on task importance over urgency. Adapting strategies to minimize unnecessary task-switching and emphasizing clear prioritization based on urgency and unit-importance, rather than absolute importance alone, can lead to more effective task management and better productivity.

General Discussion

The goal of the current research was to explore how people prioritize tasks across different contexts. We aimed to investigate situations where importance and urgency are in competition with each other and how varying environments might change task management strategies. In Experiment 1, we tested how high schedulability, versus low schedulability, influenced urgency versus importance trade-off weights and optimality. While both groups prioritized importance over urgency, in high schedulability conditions, participants better balanced urgency and importance and consistently performed close to optimal. Those in low schedulability conditions tended to give more weight to importance, and as a result, strayed further from optimality in environments where overprioritizing importance was most detrimental.

In Experiment 2, we investigated the role of task segmentation (completing tasks in one go versus in parts) and task comparability (fixed lengths versus varied lengths). Across all experimental groups, participants gave significantly more weight to absolute importance, while by and large neglecting urgency (and task length when relevant). When tasks were divided into parts, compared to when they completed as a whole, performance declined due to the option to switch tasks, which was always suboptimal in this experiment. When task lengths varied, performance decreased due to an overprioritization of absolute importance and a neglect of importance relative to task length and urgency.

The overprioritization of importance was further demonstrated across environments where the relationship between urgency and importance varied. When the urgency and importance were in high tension with each other, participants often sacrificed urgency for

importance, earning a lesser proportion of points compared to environments where urgency and importance were in moderate or low tension.

Across both experiments, we found that rational decision-making was related to optimality in terms of urgency and importance trade-offs and proportion of points earned. In contrast, intuitive decision-making was related to suboptimal trade-offs and fewer proportion of points earned. The link between both decision styles and performance differed across schedulability conditions. When participants scheduled tasks, relative to when they did not, the relationships between (i) rationality and optimality and (ii) intuition and suboptimality were more subdued. In other words, when creating a schedule, task selection was more resistant to influence by the decision styles, whereas selecting tasks without a schedule allowed decision styles to have a larger impact on decision making.

Relationship to Prior Findings

Contrary to previous literature, we did not discover a tendency for people to select urgent tasks over important ones (Zhu et al., 2018). In fact, we found the opposite where people tended to prioritize important tasks over urgent ones, to the detriment of overall performance. This deviation from past findings may be explained by two key differences in the methods of the current experiment compared to that of previous work. First, their method focused on a single task choice in a simplified decision environment. Within this task choice, there were minimal differences in outcome magnitude (3 Heresy's Kisses versus 5 Heresy's Kisses or 12 cents versus 16 cents) between the two tasks. It is possible that the minimal differences in importance were not enough to engage participants, making them more susceptible to the influence of spurious urgency. While in the current paradigm, point values represented a larger range of values (approximately 50 - 450) and translated into a real payout of \$0.00 to \$1.00. The larger range of

task importance and urgency in our experiments may be one explanation as to why the mere urgency effect did not replicate here.

Second, their method may have fostered a greater sense of urgency. Our paradigm was much more complex, featuring multiple task lists and a series of choices within each. Our task management game included a range of tasks, some without immediate deadlines, and an emphasis on the fact that not every task could be completed. This complexity may have diluted any sense of urgency tied to any single task, as participants had to focus on managing multiple tasks and prioritizing within a broader context. Thus, any sense of tension or discomfort typical brought on by urgency was absent (Sonnentag & Frese, 2003), when compared to the more immediate and direct context of the Zhu et al. study. Future research might tease apart which methodological difference produces the different results. The answer will help us to identify the boundary conditions of both Zhu et al.'s mere urgency effect and the excessive prioritization of importance we found in the current studies.

One possible explanation for the tendency to overprioritize importance in our experiments is the way the goal was framed. Participants were instructed to focus on earning as many points as possible, which was incentivized by a bonus payment proportional to points earned. The emphasis on maximizing points may have shifted attention toward task importance, at the expense of urgency. It is possible that participants valued immediate rewards more than future gains, leading them to choose tasks that offered a higher reward right away, rather than considering the total reward for the round. Participants' task management decisions were therefore made with a focus on immediate rewards at the local level rather than optimizing for greater success at the global level.

An explanation for the absence of a general strategy in varied length conditions is the complexity of the current paradigm. In contrast to previous studies, such as those by Rusou et al. (2020), which focused on simpler tasks with straightforward point values, our paradigm involved up to five columns of information. Of these, three columns contained data required for assessing optimal solutions. This complexity may have contributed to a degree of choice overload (Chernev et al., 2015; Simon, 1955). When faced with complex choice sets, individuals can experience cognitive overload, leading to difficulties in effective trade-offs and decreased decision-making quality (Ceschi et al., 2017; Scheibehenne et al., 2010). Given the relationship between lower expectations of success and reduced engagement in goal pursuit (Kappes & Oettingen, 2014), it is possible that participants experienced a decrease in motivation as the game increased in complexity. This might explain the added noise in the varied task length conditions and general emphasize on task importance at the expense of other task characteristics. With this in mind, future directions of this research can benefit from simplifying the paradigm to focus on more manageable components, which should provide clearer insights.

Limitations

One key limitation of the present work was the lack of compatibility in the structure of schedulability conditions. To address the disproportionate number of exclusions across the planning and selecting conditions, we will improve the compatibility between conditions in a follow-up study. In the planning condition, we will provide a clear signal to participants that there is an error in their task order (e.g., planning to complete a task at time point 20 that had a deadline at time point 10) *before* submission by (i) adding a column that indicates the timepoint of each slot and (ii) highlighting valid task placements in green and “impossible” task placements in grey. These adjustments will more accurately mirror the design of the selection

condition, and hopefully enhance participants' understanding of the task. As a result, it may reduce any possible frustration caused by receiving “invalid task order” messages without an explicit explanation. By providing clearer guidance, we hope to decrease the number of participants excluded from the planning condition.

Another key limitation of the current experiment is the significant correlation between the optimal order and the urgency-first benchmark. This correlation makes it difficult to determine whether participants were using the urgency-first strategy, which happened to align with the optimal order, or vice versa. Future studies will aim to address this issue by adjusting the distributions of deadlines and point values to ensure that optimal orders are proportionally related to both importance-first and urgency-first benchmarks. This approach will prevent the optimal order from inherently favouring the urgency-first approach, and as a result, will clarify whether participants rely on an urgency-first strategy in certain cases or if outcomes are simply a result of the overlap between optimality and the urgency-first.

Conclusions

Despite its limitations, the current study offers novel insights into task management tendencies in complex environments. We found that people manage tasks in suboptimal ways, overprioritizing task importance while neglecting urgency in the process. This research highlights the need for strategies that help individuals better balance urgency and importance, particularly in environments with high complexity and variable task demands. For tool developers, these insights can help inform the development of more effective task management aids and applications. For individuals, adopting this strategy can enhance productivity and efficiency in work, school, and everyday life. Ultimately, this work aims to help people make more informed choices about how to allocate their time and efforts.

References

- Aeon, B., Faber, A., & Panaccio, A. (2021). Does time management work? A meta-analysis. *PloS One*, *16*(1), 1-20.
- Allen, D. (2001). *Getting Things Done: The Art of Stress-Free Productivity*. Penguin Books.
- Altmann, E. M. (2004). Advance Preparation in Task Switching: What Work Is Being Done? *Psychological Science*, *15*(9), 616-622.
- Arabacı, G., & Parris, B. A. (2020). Inattention and task switching performance: the role of predictability, working memory load and goal neglect. *Psychological Research*, *84*, 2090-2110.
- Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, *13*(3), 219-224.
- Azmi, Z. R. M., Bakar, K. A., Abdullah, A. H., Shamsir, M. S., & Manan, W. N. W. (2011). Performance comparison of priority rule scheduling algorithms using different inter arrival time jobs in grid environment. *International Journal of Grid and Distributed Computing*, *4*(3), 61-70.
- Bannert, M. (2002). Managing cognitive load—recent trends in cognitive load theory. *Learning and Instruction*, *12*(1), 139-146.
- Baron, R.S. (1986) Distraction-Conflict Theory: Progress and Problems. *Advances in Experimental Social Psychology*, *19*, 1-39.
- Ben Alla, H., Ben Alla, S., Ezzati, A., & Touhafi, A. (2021). A novel multiclass priority algorithm for task scheduling in cloud computing. *The Journal of Supercomputing*, *77*(10), 11514-11555.

- Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica, 47*(2), 89-104.
- Białaszek, W., Ostaszewski, P., Green, L., & Myerson, J. (2019). On Four Types of Devaluation of Outcomes Due to Their Costs: Delay, Probability, Effort, and Social Discounting. *The Psychological Record, 69*(3), 415-424.
- Bolino, M. C., & Turnley, W. H. (2005). The Personal Costs of Citizenship Behavior: The Relationship Between Individual Initiative and Role Overload, Job Stress, and Work-Family Conflict. *Journal of Applied Psychology, 90*(4), 740–748.
- Bugayenko, Y., Bakare, A., Cheverda, A., Farina, M., Kruglov, A., Plaksin, Y., Pedrycz, W., & Succi, G. (2023). Prioritizing tasks in software development: A systematic literature review. *PLoS One, 18*.
- Ceschi, A., Demerouti, E., Sartori, R., & Weller, J. (2017). Decision-making processes in the workplace: How exhaustion, lack of resources and job demands impair them and affect performance. *Frontiers in Psychology, 8*, 313.
- Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology, 25*(2), 333-358.
- Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis, 34*(1), 101-122.
- Da Silva Cezar, B. G., & Maçada, A. C. G. (2023). Cognitive overload, anxiety, cognitive fatigue, avoidance behavior and data literacy in big data environments. *Information Processing & Management, 60*(6), 103482.
- Dai, B., Ali, A., & Wang, H. (2020). Exploring information avoidance intention of social media users: A cognition–affect–conation perspective. *Internet Research, 30*(5), 1455-1478.

- Dhar, R., & Nowlis, S. M. (1999). The Effect of Time Pressure on Consumer Choice Deferral. *Journal of Consumer Research*, 25(4), 369-384.
- Ferrari, J. R. (2001). Procrastination as self-regulation failure of performance: effects of cognitive load, self-awareness, and time limits on 'working best under pressure'. *European Journal of Personality*, 15(5), 391-406.
- Garhammer, M. (2002). Pace of life and enjoyment of life. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being*, 3(3), 217-256.
- Giurge, L. M., Whillans, A. V., & West, C. (2020). Why time poverty matters for individuals, organisations and nations. *Nature Human Behaviour*, 4(10), 993-1003.
- Gleick, J. (1999). *Faster: The Acceleration of Just About Everything*. Pantheon Books.
- Häfner, A., & Stock, A. (2010). Time management training and perceived control of time at work. *The Journal of Psychology*, 144(5), 429-447.
- Holmes, M. (2022). Time Scarcity and Student Performance: Instructional Strategies for Busy Adult Online Students. *Canadian Journal of Learning and Technology*, 48(3), 1-14.
- Hu, D., Bhuiyan, M. M., Lim, S., Wiese, J., & Lee, S. W. (2024, June). *Unpacking Task Management Tools, Values, and Worker Dynamics* [Paper presentation]. Conference: CHIWORK 2024: Annual Symposium on Human-Computer Interaction for Work, Newcastle, UK.
- Husain, M., & Roiser, J. P. (2018). Neuroscience of apathy and anhedonia: a transdiagnostic approach. *Nature Reviews Neuroscience*, 19(8), 470-484.
- Ilies, R., Dimotakis, N., & De Pater, I. E. (2010). Psychological and physiological reactions to high workloads: Implications for well-being. *Personnel Psychology*, 63(2), 407-436.

- Job, V., & Brandstätter, V. (2009). Get a taste of your goals: promoting motive-goal congruence through affect-focus goal fantasy. *Journal of Personality, 77*(5), 1527-1559.
- Jyothi, N.S., & Parkavi, A. (2016). A study on task management system. *2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS)*, 1-6.
- Kahol, K., Satava, R. M., Ferrara, J., & Smith, M. L. (2009). Effect of short-term pretrial practice on surgical proficiency in simulated environments: a randomized trial of the "preoperative warm-up" effect. *Journal of the American College of Surgeons, 208*(2), 255-268.
- Kappes, A., & Oettingen, G. (2014). The emergence of goal pursuit: Mental contrasting connects future and reality. *Journal of Experimental Social Psychology, 54*, 25-39.
- Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching--a review. *Psychological Bulletin, 136*(5), 849-874.
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology, 139*(4), 665-682.
- Levine, R. V., & Norenzayan, A. (1999). The pace of life in 31 countries. *Journal of Cross-Cultural Psychology, 30*(2), 178-205.
- Li, W., Zhang, H., Zhang, C., Luo, J., Wang, H., Wu, H., Zhu, Y., Cui, H., Wang, J., Li, H., Zhu, Z., Xu, Y., & Li, C. (2021). The Prevalence of Psychological Status During the COVID-19 Epidemic in China: A Systemic Review and Meta-Analysis. *Frontiers in Psychology, 12*, 614964.

- Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology, 75*, 5-18.
- Liberman, N., & Trope, Y. (2014). Traversing psychological distance. *Trends in Cognitive Sciences, 18*(7), 364-369.
- Liefooghe, B., Barrouillet, P., Vandierendonck, A., & Camos, V. (2008). Working memory costs of task switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 478-494.
- Liu, Y., Yang, Y., Bai, X., Chen, Y., & Mo, L. (2022). Do Immediate External Rewards Really Enhance Intrinsic Motivation? *Frontiers in Psychology, 13*, 853879.
- MacDonald, W. (2003). The impact of job demands and workload on stress and fatigue. *Australian Psychologist, 38*(2), 102-117.
- Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options. *Psychological Science, 8*(6), 423-429.
- Microsoft. (2008, January 14). *Survey shows increasing worldwide reliance on to-do lists*.
- Monsell S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134–140.
- Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. Times Books/Henry Holt and Co.
- Naveed, M. A., & Anwar, M. A. (2020). Towards information anxiety and beyond. *Webology, 17*(1).
- Nussbaum, S., Trope, Y., & Liberman, N. (2003). Creeping dispositionism: the temporal dynamics of behavior prediction. *Journal of Personality and Social Psychology, 84*, 485-497.

- O'Donoghue, T., & Rabin, M. (2000). The economics of immediate gratification. *Journal of Behavioral Decision Making*, 13(2), 233-250.
- Odum A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior*, 96(3), 427-439.
- Oettingen, G., & Gollwitzer, P. M. (2010). *Strategies of setting and implementing goals: Mental contrasting and implementation intentions*. Konstanzer Online-Publikations-System.
- Rabin, L. A., Fogel, J., & Nutter-Upham, K. E. (2011). Academic procrastination in college students: The role of self-reported executive function. *Journal of Clinical and Experimental Neuropsychology*, 33(3), 344-357.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology*, 124, 207-231.
- Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of experimental psychology. Human Perception and Performance*, 27(4), 763-797.
- Rusou, Z., Amar, M., & Ayal, S. (2020). The psychology of task management: The smaller tasks trap. *Judgment and Decision Making*, 15(4), 586-599.
- Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, 37(3), 409-425.
- Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: an effort-reduction framework. *Psychological Bulletin*, 134(2), 207.
- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1), 99-118.

- Sonnentag, S., & Frese, M. (2013). Stress in organizations. In N. W. Schmitt, S. Highhouse, & I. B. Weiner (Eds.), *Handbook of psychology: Industrial and organizational psychology* (2nd ed., pp. 560–592). John Wiley & Sons, Inc.
- Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin*, *133*(1), 65-94.
- Svenson, O., Edland, A., & Slovic, P. (1990). Choices and judgments of incompletely described decision alternatives under time pressure. *Acta Psychologica*, *75*(2), 153-169.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.
- Task Management Software Market Size, Share, and Trends Analysis. (2023). *Regional Research Reports*.
- Thomas, M., & Tsai, C. I. (2012). Psychological distance and subjective experience: How distancing reduces the feeling of difficulty. *Journal of Consumer Research*, *39*(2), 324-340.
- Waugh, C. E., Thompson, R. J., & Gotlib, I. H. (2011). Flexible emotional responsiveness in trait resilience. *Emotion*, *11*(5), 1059-1067.
- Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective and Behavioral Neuroscience*, *15*(2), 395-415.
- Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PloS One*, *8*(7).
- Wylie, G., & Allport, A. (2000). Task switching and the measurement of “switch costs.” *Psychological Research*, *63*, 212-233.

- Zhihong, H., Johnson, S.G.B., Brown, G.D.A. (2024). Diversion Aversion: A Simple Model for Cognitive Efficiency in Task Sequencing. *Manuscript in Preparation*.
- Zhu, M., Yang, Y., & Hsee, C. K. (2018). The mere urgency effect. *Journal of Consumer Research*, 45, 673-690.
- Zumbach, J., & Mohraz, M. (2008). Cognitive load in hypermedia reading comprehension: Influence of text type and linearity. *Computers in Human Behavior*, 24(3), 875-887.