Addressing Procrastination and Improving Task Completion Efficiency through Agent-Based Interventions

Ethan Beaird
University of Tulsa
Tulsa, OK, USA
ethan-beaird@utulsa.edu

Selim Karaoğlu
University of Tulsa
Tulsa, OK, USA
sek6301@utulsa.edu

Feyza Merve Hafızoğlu
İstanbul Commerce University
İstanbul, Turkey
fmhafizoglu@ticaret.edu.tr

Sandip Sen
University of Tulsa
Tulsa, OK, USA
sandip-sen@utulsa.edu

ABSTRACT

Procrastination, i.e., irrational delay, seriously and increasingly affects people's daily and professional lives in today's society where social media and easy access to entertainment options are plentiful. Psychology literature offers various types of interventions developed to reduce an individual's level of procrastination; however, only a limited number of people experiencing procrastination have access to such interventions. Leveraging agent technology as even a partial remedy to this widespread public health problem can be both highly beneficial and equitable due to its ubiquitous nature. In this study, we develop a model of procrastination on task completion and two levels of agent-based interventions to assist individuals in overcoming procrastination. The effects of agent interventions on procrastination are evaluated through an extensive set of controlled experiments with participants recruited from Amazon Mechanical Turk. The agent engages the user using instances of given task types to develop a shared awareness of user preferences and capabilities. This preference model is then used both to choose effective interventions as well as measure and reward subsequent user performance. We collect and use both task completion metric data and survey data to assess individuals' perceptions of procrastination, task completion satisfaction, and the usefulness of agent support. Our data analysis indicates that using agent-based interventions can effectively help people reduce procrastination.

KEYWORDS

human-agent interaction, procrastination, agent intervention

ACM Reference Format:

Ethan Beaird, Selim Karaoğlu, Feyza Merve Hafızoğlu, and Sandip Sen. 2024. Addressing Procrastination and Improving Task Completion Efficiency through Agent-Based Interventions. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 9 pages.

1 INTRODUCTION

Procrastination, or irrational delay in completing tasks of importance at hand [29], is caused by failure of self-regulation [27] and lack of self-efficacy [15] and has increasingly posed significant challenges to citizens of our digitally connected world [9]. The disparity between people's intentions to complete important tasks and their tendency to procrastinate illustrates an inclination to irrationally prioritize immediate gratification, i.e. short-term reward over long-term benefits. People tend to procrastinate on necessary tasks and chores that are perceived to be negative, unpleasant, or challenging [33]. For instance, students may delay studying for exams, people may procrastinate filing their tax returns, and individuals often put off visiting the doctor when necessary.

Social media companies thrive on their ability to reward people with instant gratification, and omnipresent communication platforms result in citizens spending considerable time and energy [19] in initiating and responding to messages and calls that can easily divert attention from pending tasks at hand [2]. Furthermore, the widespread availability of smartphones and other mobile devices and constant access to entertainment streaming platforms supported by fast, wireless internet access provides a plethora of options and incentives for users to procrastinate.

Considering its negative consequences on individuals' health [17, 28], well-being [7, 23, 28], and task accomplishments [14], eliminating or reducing procrastination is a challenging but potentially significantly beneficial endeavor. Several studies in the psychology literature [32] focus on the stability of procrastination: whether people can change— or at least reduce— this detrimental behavior. A range of interventions, although not widely available, have been developed to reduce procrastination levels, such as self-regulation [1, 13, 26], cognitive-behavioral therapy [15, 24], and social support [25], among others. However, these interventions are not scientifically proven as perfect remedies, and only a minority of those who suffer from procrastination have access to such interventions due to time constraints, financial limitations, or other resource barriers. Leveraging agent technology to address this pervasive issue can prove highly effective, as these automated assistants, serving as personal aids, can be deployed widely at minimal costs, thus helping reduce the disparity in access to procrastination interventions.

Despite the prevalence and significant impact of procrastination in our society, the majority of previous studies focus only on academic procrastination [11, 27, 34]. Furthermore, most studies rely on surveys rather than actual task performance data. We evaluate two key research questions in the context of agent research:

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

- Can agent intervention mechanisms be developed to help people overcome or significantly reduce procrastination?
- Is there a significant difference in the effect between different levels of agent interventions?

In this study, we develop a model of procrastination and agent-based interventions to support individuals in overcoming procrastination. As part of our empirical approach, we devise five diverse task types that span a range of basic skills. We conjecture that the likelihood and level of procrastination on these task types will differ among individuals. To assess the procrastination tendency of individuals for a certain task, we use their preferences among the task types, i.e. an individual's likelihood to procrastinate on a task type and their preference for that task type are correlated [20].

We conduct experiments with human subjects to simulate procrastination and design utility functions based both on the participants' preferences for different task types and their performance levels while completing those tasks such that completing less preferred tasks with lower performance can still yield higher utilities than highly preferred tasks completed with higher performance. Each participant is asked to perform two sets of task instances selected from the given task types. The first set of tasks is completed without any external assistance, while the second set of tasks is completed with the aid of one of two types of agent interventions. The first type, termed "low level," involves minimally intrusive interventions. Conversely, the second type, termed "high level," entails highly noticeable and apparent interventions. Throughout the study, we collect both task completion metric data and survey data to assess individuals' perceptions of procrastination and agent interventions as well as satisfaction in task completion. Our data analysis indicates that the use of agent-based interventions plays a significant role in helping individuals reduce procrastination.

The remainder of the paper is structured as follows. We first present related work. Following, the procrastination model that is considered in this research is outlined, and our empirical methodology is explained. The results of our experiments are presented in the subsequent section, where we also discuss the empirical findings. The paper concludes with a summary and directions for future research.

2 RELATED WORK

Procrastination behavior is an increasingly serious problem in our professional lives [11, 27, 34] as well as daily lives [16]. There is a relatively small body of empirical research in the area of procrastination of which the majority involves survey data [9, 10, 16, 27]. Furthermore, most research on procrastination has been carried out in academic domains with college students [27, 34]. Recently, there is a growing body of literature that is concerned with adults' procrastination in life-domains [5, 9, 10]. The indisputable effect of online entertainment and other online offerings has contributed significantly to the rise of procrastination behavior among adults [9].

Procrastination is a complex psychological phenomenon. Numerous factors influencing this troubling behavior are addressed in social literature such as age [5], gender [36], personality [18, 31], mood [7], environment [6], and nature of task [33]. Furthermore, it is very well acknowledged that procrastination causes lower levels of health [17, 28], well-being [7, 23, 28], and achievements, i.e., performance [14].

To effectively address procrastination, an individual's intention-action gap should be reduced. Various intervention mechanisms have been identified as helpful [32]. Schouwenburg *et al.* [1] offer three general categories of interventions: 1) Training self-regulatory skills; 2) building self-efficacy; and 3) organizing social support. Training *self-regulatory skills* involves deploying our cognitive, emotional, and behavioral resources to achieve our goals [4]. Self-regulation failure is associated with a lack of self-determination, planning, and prioritizing tasks as well as a failure to concentrate on tasks and shield from distractions [1, 26]. Stimulus-control techniques (i.e. eliminating the effects of distractions), techniques of goal definition, and time management techniques [13] can be classified in this category. Building *self-efficacy* aims to replace negative, unproductive, and inhibiting thoughts with positive, productive, and motivated thoughts [15, 24]. Peer support is centered around the recognition that many other people also struggle with procrastination, forming a sense of solidarity. Similarly, *social support* has also been identified as an effective means of reducing procrastination [25]. Additionally, more recent interventions, like reading self-help books and engaging in training sessions, can be considered as a fourth category of intervention mechanisms.

Recently, several studies have suggested that technology-based interventions can be useful in overcoming irrational delay. CatAlyst [3] supports distracted workers by generating a continuation of interrupted work through the use of a generative model for workers to resume tasks. Zavaleta *et al.* [35] showed that email interventions by instructors can reduce the delay of students starting their online homework. De Vries *et al.* [8] found that experts sending motivational messages through a digital medium could be more motivating in the earliest stages of behavior change, while peer-designed messages in a digital medium could be more motivating in the later stages. GanttBot [22] is a chatbot that is developed using conversational agents with several abilities: reminding students about landmarks, informing tutors when interventions are needed, and the ability to learn from previous interactions. StudiCare [21] is a digital coach rooted in internet and mobile-based cognitive behavioral therapy techniques, helping guide students to achieve their academic goals.

The present study extends existing research on procrastination as follows:

Time: Address procrastination in a short-term context with artificial tasks rather than a long-term context [21, 28, 34, 35]
Data: Consider not just survey data such as in [5, 6] but also actual task performance data.
Task: Task domain involves daily life tasks rather than a specific domain such as in [14, 18, 21, 22, 26, 27, 35]
Users: Recruit adults rather than college students exclusively [7]
Agents: Make use of cost-effective and ubiquitous agent intervention mechanisms rather than time and location-dependent techniques [15, 24]

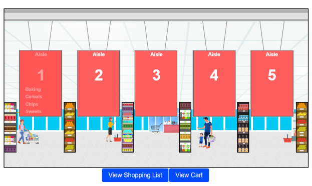
To the best of our knowledge, this is the first empirical study on general life procrastination that consists of agent-based interventions, artificial daily tasks, task performance data, and survey data that is presented to the intelligent agents research community.

3 PROCRASTINATION MODEL

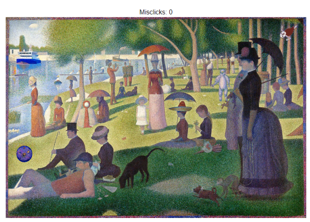
This section presents the task types, the utility function model, and the agent interventions in our study.

3.1 Task Domain

Our task domain comprises of five different task types inspired by and designed to emulate a wide range of common daily activities. These tasks have varying degrees of satisfaction in completion, closely mirroring real-life situations. We present five categories of tasks of which our tasks belong to— *Pattern Recognition, Comprehension, Memory, Generative,* and *Computational*.

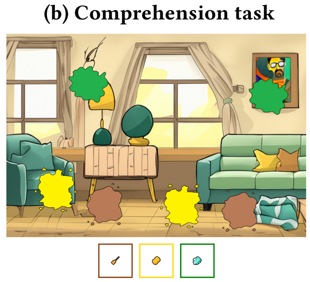


(a) Pattern Recognition task

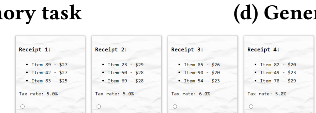


(b) Comprehension task

(c) Memory task



(d) Generative task



(e) Computational task

Figure 1: Different task type instances

Pattern Recognition tasks involve recognizing and identifying recurring structures and objects, such as symbols, shapes, and other patterns. An example of such a task is searching for items on a shopping list at a grocery store. *Comprehension* tasks require individuals to carefully understand and critically interpret information in new and unfamiliar environments, such as recognizing irregularities and potential dangers in social scenarios. *Memory* tasks focus on retaining and recalling prior information; these tasks often require an individual to reflect upon and remember specific details that they have encountered earlier in their day or even week. Examples of memory tasks include remembering the order of steps to make coffee and the correct locations to put away household items. *Generative* tasks involve producing some form of content or otherwise altering the state of one's environment, like cleaning/rearranging one's room or coloring a picture. *Computational* tasks encompass mathematical and algorithmic processes used in daily life such as mental math in receipts, counting one's personal effects, and balancing monetary income and expenditure.

Grocery Shopping (Pattern Recognition) task (see Figure 1a) immerses users in a simulated shopping environment within a grocery store. Users are equipped with a shopping list, a shopping cart, and access to five shopping aisles. The shopping list provides a detailed inventory of items, specifying both the name and quantity of products to be added to the user's cart. Users may view the categories of items present in each aisle by hovering over the signs above each aisle and can enter each aisle by clicking on the signs. Once inside an aisle, users can use a shopping cart scroller at the bottom of the page to move left and right throughout the aisle. Each aisle has numerous items that the user can click on and add any number of to their cart. The challenge here lies in identifying the correct aisles to enter and recognizing the items on the shopping list; users must match the names of the required items on their shopping list with the corresponding images of the items within the shopping aisles and add these items to their shopping cart. The score received on this task is determined by how many of the required items on the shopping list are added with their correct quantities.

Scavenger Hunt (Comprehension) task (see Figure 1b) engages users with cluttered paintings teeming with numerous figures and objects. Within each painting, there lurk five subtle anomalies— discrepancies or incongruencies that do not necessarily belong to the scene. Users are expected to inspect the artwork and pinpoint the locations of these anomalies by clicking. In order to avoid random clicks, the number of misclicks is limited to five clicks, i.e., clicks that are not anomalies, before the task is automatically terminated. The score received on this task is determined by the number of anomalies found by the user.

Toy Boxes (Memory) task (see Figure 1c) begins by showing the users four toy boxes arranged horizontally. Following this, numerous images of toys are displayed above each box, indicating in which box each toy is stored. This display lasts for several seconds, during which users are tasked with memorizing the correct pairing of toys and boxes. Once the toy images disappear, a single toy is displayed, prompting the user to identify and select the appropriate box to which it belongs. The process is repeated until all toys have been shown. The score on this task is determined by the user's accuracy in selecting the correct box for each displayed toy.

Cleaning (Generative) task (see Figure 1d) is realized within a simulated room environment that is in a state of disarray. The scene depicts a messy living room with several interactive spill spots that users must tidy up by using the appropriate cleaning tools. Users may choose from a selection of three cleaning tools: the broom for handling brown spills, the sponge for handling yellow spills, and the soap for handling green spills. Each tool is indicated by a colored outline around its selection box. Once users have selected the appropriate tool for a particular spill, users may clean that spill by repeatedly clicking the spill until it is gone. Throughout the cleaning process, several state transitions and cleaning gestures are animated to provide users with visual feedback for ease of use. The score received on this task is determined by both the number of spills successfully cleaned by the user and their progress in cleaning the remaining spills.

Receipt Computation (Computational) task (see Figure 1e) presents users with multiple itemized receipts featuring randomly generated prices for each item. Additionally, each receipt includes a tax rate percentage at the bottom. Users are required to perform calculations, either mentally or by employing external tools, to compute the total cost for each receipt by summing the item prices and applying the respective tax rate. Subsequently, they must identify and select the receipt with the lowest total cost. The score received on this task is determined by how accurately they choose the receipt with the closest cost to the actual lowest one.

3.2 Utility Function

In this study, we adopt the following definition of procrastination: "to voluntarily delay an intended course of action despite expecting to be worse off for the delay" [29]. Our objective in this section is to successfully engender procrastination within an artificial task domain environment.

To effectively model and characterize procrastination within an artificial task domain environment, it is essential to create conditions that allow individuals to procrastinate during their completion of tasks. Milgram et al.[20] states that "Procrastination is found to be greater in tasks that are regarded as unpleasant". Accordingly, procrastination in our study corresponds to the users' selection of their preferred tasks that are pleasant to them with lower utility despite receiving less total compensation by doing so. Users are limited to selecting only a subset of available tasks and forgoing the utility of any unselected tasks; this is congruent to individuals delaying the completion of certain tasks and thereby "missing" a deadline in a real-world scenario.

To facilitate this, each user completes initial demo instances of each task type within an artificial domain, and we record their performance scores. Additionally, each user is asked to rank their preference for each task type. The value of task i's completion, v_i, by a user is computed as follows:

$$v_i = w \cdot 10 \cdot s_i + (1 - w) \cdot p_i, i \in \{1, 2, 3, 4, 5\} \tag{1}$$

where $s_i \in [0, 1]$ is the user's score in the initial demo, $p_i \in [1, 10]$ is the preference ranking, and $w = 0.3$ is the weighting of s_i for task type i, chosen to emphasize the user's task preference while still considering the user's ability to perform that task. Then we can construct the utility function over the task types as follows:

$$U(v_i) = \frac{1}{1 + e^{0.625 \cdot v_i - 2\pi}}. \tag{2}$$

This utility function is designed to discourage tasks that an individual might naturally prioritize by penalizing tasks that align with their preferences and abilities.

Figure 2 plots the utility curve of a task; the utility decreases as either preference ranking p or score in the initial task demo s increases (the constants are chosen to obtain a concave function; the exact coefficient values are not critical). As such, tasks that a user prefers tend to receive lower utility scores than tasks that they do not prefer. Since the utility of a task is displayed to users and directly correlates to the amount of compensation that a user receives for completing that task, users may be inclined to avoid tasks that they

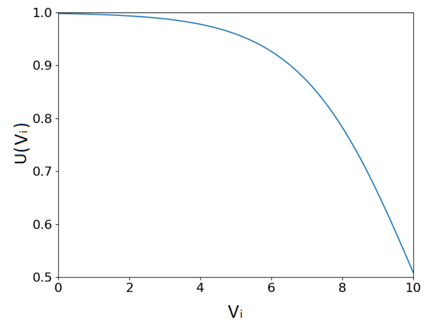


Figure 2: Utility function

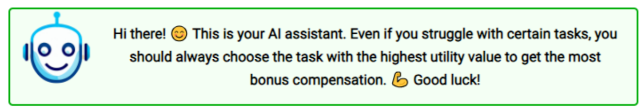


Figure 3: Motivational guide intervention

performed poorly in during their initial task demos. To address this, our utility function also takes into account their performance in the initial task demo and rewards lower performance. This approach encourages individuals to consider tasks where their performance may be subpar, thereby promoting a more balanced task selection process.

3.3 Agent Interventions

We have developed several agent interventions to engage with the user during the study. These techniques encompass proactive interventions (i.e., prior to a user's task selection and completion), reactive interventions (i.e., after a user has chosen and completed a task), and ongoing interventions (i.e., while a user is actively completing a task).

Motivational Guide is a proactive intervention where the agent encourages the user to choose higher utility tasks through a text message. This intervention is subtle and non-intrusive to gently influence a user's decision-making.

Dynamic Utility Highlighting is a proactive intervention where the agent emphasizes the tasks with the highest utility. In this approach, the agent adds "recommendation" popups to the tasks with the highest utility and highlights each task by surrounding it with a green outline. Simultaneously, the other lower utility tasks are surrounded by a red outline. Additionally, all of the tasks are sorted based on their utility in descending order. This intervention is designed to be more noticeable to the user than the "Motivational Guide."

Medal Rewards is a reactive intervention which introduces a "medals box", where users can see the medals that they have earned after successfully completing each task. Medals earned are dependent on both the user's task selection and performance. Users will only receive bronze medals from tasks that are not the highest utility available regardless of their performance within those tasks. However, for high utility tasks, users have the potential to receive

Figure 4: Dynamic utility highlighting intervention

both silver and gold medals for adequate performance. After each task is completed, an appropriate medal is added to the medals box by the agent, accompanied by another reactive agent intervention, which we call *Reactive Motivation*. During this intervention, the agent leaves an encouraging comment tailored to the user's performance in the previous task.

Figure 5: Medal rewards & reactive motivation intervention

Intratask Encouragement is an ongoing intervention which adds agent-driven motivation within each task that the user completes. After thirty seconds have elapsed in any given task, the agent appears with a popup and provides positive feedback and encouragement relevant to the specific task that the user is completing.

4 EMPIRICAL METHODOLOGY

This section presents our hypotheses and a detailed description of our experimental setup, including survey and metric data.

Hypothesis 1. Agent intervention mechanisms can be developed to help people overcome or significantly reduce procrastination.

Hypothesis 2. Low and high level agent interventions may be equally effective in reducing procrastination.

4.1 Experimental Setup

The two experimental conditions based on the levels of intensity of agent interventions are as follows:

Low: The low level group includes the *Motivational Guide* and *Dynamic Utility Highlighting* interventions, providing users with involved but minimally intrusive guidance.

High: The high level group encompasses all agent intervention techniques: *Motivational Guide*, *Dynamic Utility Highlighting*, *Medal Rewards*, *Reactive Motivation*, and *Intratask Encouragement*, offering a large suite of interventions to support users.

The empirical study consists of the following steps:

• *Preference learning:* The user completes a series of demo tasks for each of the five task types in the study. Before each demo, they are presented with a detailed instructional video. The user ranks their preferences for each task in a survey.

• *Phase 1:* The user is provided with ten task instances, with two instances of each of the five task types. They are asked to choose and complete any five of these task instances. They have the flexibility to choose how they distribute their completion across the categories, allowing them to decide based on their preferences. The user completes an inter-phase survey about their procrastination and satisfaction in completing tasks during Phase 1.

• *Phase 2:* The user again encounters two new instances of each task type, just like in Phase 1. However, this time, an agent and agent interventions are introduced. Upon completing Phase 2, the user is directed to a final survey which covers satisfaction in completing tasks, interaction with the agent, and overall procrastination.

The user has one minute to complete each task and is compensated proportional to overall utility gained in the game.

Survey: The study includes an inter-phase survey and a final survey where the items are measured on a 5-point Likert scale from "Strongly disagree" to "Strongly agree".

In order to assess the user's satisfaction with their task completion performance and their perceived procrastination, the inter-phase survey, conducted after Phase 1, and the final survey, conducted after Phase 2, consist of the following items:

- **Q1:** I liked the outcome of today's task list
- **Q2:** I feel satisfied with my performance in the previous tasks
- **Q3:** My accomplishments today give me a feeling of satisfaction
- **Q4:** I chose tasks that I liked regardless of utility

We test for a significant increase in mean survey responses for Q1-3 and a significant decrease in mean survey responses for question 4 in section 5.

The final survey serves a dual purpose. It continues to gauge the user's satisfaction with their task completion performance and perceived procrastination, but it also measures the user's perception of the newly introduced agent interventions. Furthermore, more items are included to further gauge the user's perceived level of procrastination. The items related to the perception of agent interventions include:

- **Q5:** I was satisfied with the experience of interacting with the AI assistant to complete tasks
- **Q6:** The AI assistant provided me helpful guidance

The items that are adapted from the Pure Procrastination Scale [30] include:

- **Q7:** I did not choose the optimal tasks
- **Q8:** Not doing high-utility tasks has undermined my performance

Metrics: Selection score and *selection-performance score* metrics are used in our analysis. In order to analyze the impact of agent interventions on a user's task selection and allocation in a particular phase, we define *optimal utility* as:

$$O = 2 \cdot U(t_1) + 2 \cdot U(t_2) + 1 \cdot U(t_3)$$

where t_1, t_2, and t_3 are the first, second, and third highest utility scores.

We define *user utility* as: $\mu = \sum_T U(t_i)$, where T corresponds to the tasks selected by the user. We also define *user performance utility* as:

$$\mu_p = \sum_T U(t_i) \cdot \text{score}(t_i)$$

Then, *selection score* is the proportion of *user utility* to *optimal utility* μ/O and *selection-performance score* is the proportion of *user performance utility* to *optimal utility* μ_p/O. *Selection score* measures users' efficiency in task selection by comparing their accumulated utility from chosen tasks to the *optimal utility* achievable by selecting the highest utility tasks. A higher *selection score* indicates more effective task selection. Similarly, the *selection-performance* score evaluates users by also considering their performance in each selected task.

The data collected from Phases 1 and 2 are highly relational and dependent; the only change in the experimental process is the inclusion of agent interventions. Therefore, we employ a one-sided paired t-test on *selection score* and *selection-performance score* to test for statistical significance. Furthermore, both one-sided paired t-test and non-parametric one-sided Wilcoxon signed-rank test with Pratt zero-handling are employed to test for statistical significance on survey data.

Participants: We recruited 116 participants through Amazon Mechanical Turk. Attention checks, i.e., basic questions to ensure participants were engaged, were included. No data of participants was eliminated due to insufficient attention. There were 48 participants in each intervention group. Approximately 46% of the participants were female. The age distribution was as follows: 18-24 years, 7.7%; 25-34 years, 44.8%; 35-44 years, 21.6%; 45-54 years, 13.7%; 55-64 years, 11.2%; and 65 years or older, 1%. The distribution of education levels was as follows: high school degree, 7.7%; some college experience, 1%; associate's degree, 1.7%; bachelor's degree, 44.8%; and graduate degree, 44.8%. The ethnicity distribution was as follows: White, 97.4%; Native-American, 1.6%; and African-American, 1%.

5 RESULTS

This section presents the analysis of experimental data collected during both Phase 1 and 2. We examine the impact of agent interventions on the *selection score* and *selection-performance score*, specifically focusing on the relative proportions of procrastination-free optimal task selections. Additionally, we examine the survey data from users.

5.1 Selection Score Analysis

Figure 6 shows the summary data on selection scores for both conditions during Phase 1 and 2. Comparing Phase 1 and 2 (within group), in the low intervention condition, the *selection score* in Phase 2 ($M = 0.905, SD = 0.110$) is significantly ($p < 0.001$) higher than the *selection score* in Phase 1 ($M = 0.839, SD = 0.100$). Likewise, in the high intervention condition, the *selection score* is significantly ($p < 0.005$) higher in Phase 2 ($M = 0.900, SD = 0.095$) compared to Phase 1 ($M = 0.849, SD = 0.093$) in Phase 1. These results indicate that both low and high level agent interventions lead to a significant

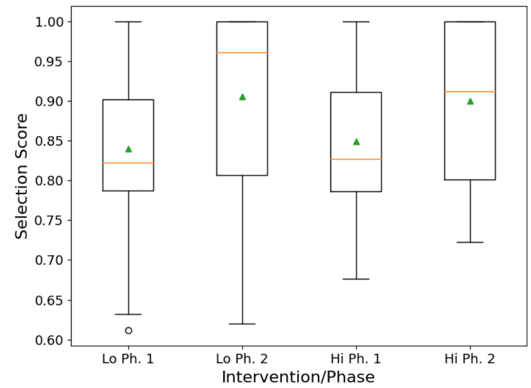


Figure 6: Selection scores

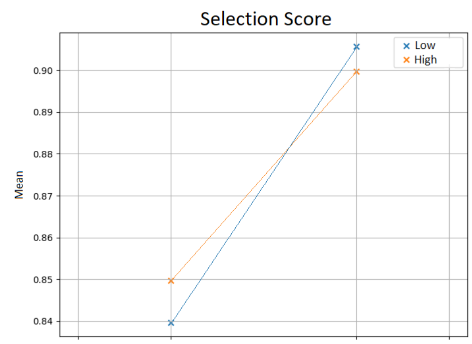


Figure 7: Comparison of selection scores

increase in the selection of the less preferred, i.e., procrastinated, tasks.

To compare the effects of low and high level agent interventions on selection score, t-tests were used. Comparing the two conditions (between groups), no significant difference in selection score is found between the low and high interventions as seen in Figure 7.

5.2 Selection-Performance Score Analysis

Figure 8 shows the summary data on selection-performance scores for both conditions during Phase 1 and 2. Comparing Phase 1 and 2 (within group), in the low intervention condition, the *selection-performance score* in Phase 2 ($M = 0.481, SD = 0.213$) is significantly ($p < 0.05$) greater than the *selection score* in Phase 1 ($M = 0.440, SD = 0.200$). Similarly, in the high intervention condition, the *selection-performance score* significantly ($p < 0.05$) higher in Phase 2 ($M = 0.505, SD = 0.218$) compared to Phase 1 ($M = 0.454, SD = 0.218$) in Phase 1. These results indicate that both low and high level agent interventions lead to a significant increase in the selection of the procrastinated tasks.

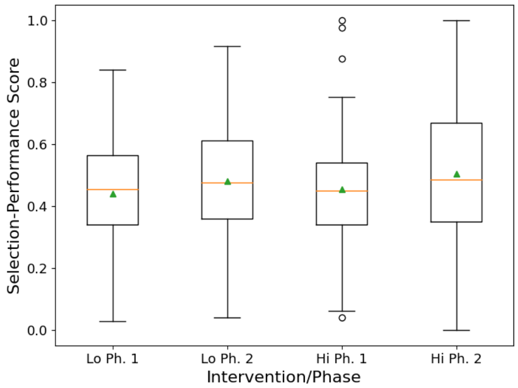


Figure 8: Selection-performance scores

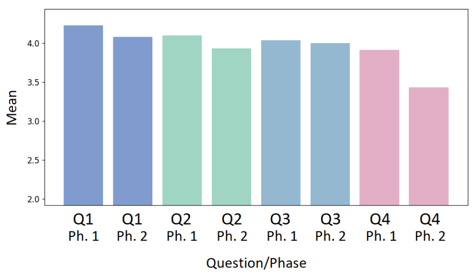


Figure 10: Low intervention condition survey results

Figure 9: Comparison of selection-performance scores

To compare the effects of low and high level agent interventions on selection-performance score, t-tests were used. Comparing the two conditions (between groups), no significant difference in selection-performance score is found between the low and high interventions as seen in Figure 9.

5.3 Survey Data Analysis

Figure 10 presents the average responses to survey items Q1, Q2, Q3, and Q4 in Phases 1 and 2 for the low intervention condition. Average response to Q1 (related to perceived outcome of tasks) was

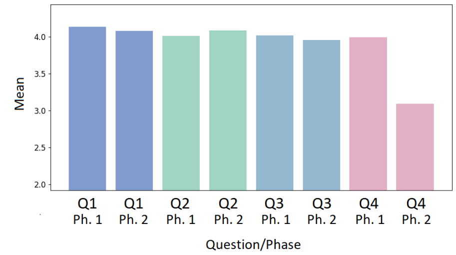


Figure 11: High intervention condition survey results

($M = 4.229, SD = 0.509$) in the inter-phase survey (at the end of Phase 1) and ($M = 4.083, SD = 0.571$) in the final survey (at the end of Phase 2). Average response to Q2 (related to satisfaction in task performance) was ($M = 4.104, SD = 0.742$) in the inter-phase survey and ($M = 3.937, SD = 0.851$) in the final survey. Average response to Q3 (related to satisfaction in users' accomplishment) was ($M = 4.042, SD = 0.705$) in the inter-phase survey and ($M = 4.0, SD = 0.707$) in the final survey. Average response to Q4 (related to users' selection of tasks they prefer without considering utility) was ($M = 3.916, SD = 0.837$) in the inter-phase survey and ($M = 3.437, SD = 1.018$) in the final survey.

To compare the difference in average responses to survey items between Phase 1 and 2, a one-sided Wilcoxon signed-rank test and a one-sided paired t-test were used. There was a significant ($p < 0.05$) difference in average responses to Q4 between Phase 1 and 2, i.e., the users paid more attention to utilities while choosing tasks. However, no significant differences were found for the items Q1, Q2, and Q3.

Figure 11 presents the average responses to survey items Q1, Q2, Q3, and Q4 in Phases 1 and 2 for the high intervention condition. Average response to Q1 (related to perceived outcome of tasks) was ($M = 4.125, SD = 0.725$) in the inter-phase survey (at the end of Phase 1) and ($M = 4.083, SD = 0.671$) in the final survey (at the end of Phase 2). Average response to Q2 (related to satisfaction in task performance) was ($M = 4.000, SD = 0.866$) in the inter-phase survey and ($M = 4.083, SD = 0.731$) in the final survey. Average response to Q3 (related to satisfaction in users' accomplishment) was ($M = 4.020, SD = 0.558$) in the inter-phase survey and ($M = 3.916, SD = 0.759$) in the final survey. Average response to Q4 (related to users' selection of tasks they prefer without considering utility) was ($M = 3.958, SD = 0.888$) in the inter-phase survey and ($M = 3.125, SD = 1.111$) in the final survey.

To compare the difference in average responses to survey items between Phase 1 and 2, a one-sided Wilcoxon signed-rank test and a one-sided paired t-test were used. There was a significant ($p < 0.05$) difference in average responses to Q4 between Phase 1 and 2, i.e., the users paid more attention to utilities while choosing tasks. However, no significant differences were found for the items Q1, Q2, and Q3.

We examine the impact of low and high-level agent interventions on users' perceived procrastination and their perception of agent interventions by conducting t-tests on the mean scores of items Q5, Q6, Q7, and Q8 during Phase 2. The results indicate that there is no significant difference between the low and high interventions for these survey items.

6 DISCUSSION

Prior studies have shown that the intervention techniques can be helpful in reducing an individual's procrastination tendencies [3, 21, 22, 32, 35]. This study investigates the role of agents in controlling and managing procrastination by testing various agent intervention techniques in an artificial empirical environment.

Individuals procrastinate only when actions are voluntarily delayed despite expecting to be worse off for the delay [29]. In the context of our task domain, to engender procrastination, the utility of a task is designed to be inversely proportional to the user's preference for that task. Results show that the users did not consistently select tasks that offered the highest utility. Many users preferred to perform tasks that they enjoyed rather than gain more utility. These irrational task selections are voluntary, and by doing so, they forgo and eliminate potential gained utility, akin to a form of *irrational delay*. This phenomenon is congruent with real-life procrastination, where many individuals often choose tasks that they find more temporarily enjoyable instead of tasks that would ultimately make them better off.

Our findings suggest that our developed agent intervention techniques can significantly impact a user's procrastination tendencies during task selection and completion (see Figures 6 and 8). This effect is evident through two metrics we employed: *selection score* and *selection-performance score*. Notably, these metrics show statistically significant improvements in both the low and high intervention conditions, where multiple intervention techniques are concurrently present. Additionally, survey data suggests that users perceive themselves as procrastinating at a lower rate when agent interventions are present in both conditions (see Figures 10 and 11). **Hypothesis 1** is thereby supported.

Furthermore, while both the low and high interventions effectively reduce procrastination tendencies, our analysis reveals no statistically significant difference between the two conditions in achieving this outcome. **Hypothesis 2** is thereby supported. Further work is needed to analyze the effects of different levels of interventions on procrastination. Additionally, it is crucial to acknowledge potential limitations in translating our findings within our gamified task domain to real-life tasks. It is possible that the online gamified representations of our tasks could significantly influence human behavior and the effectiveness of agent interventions since gamification often increases the quality of work in task performance [12]. Further work is needed to analyze gamification's effects on task selection.

7 CONCLUSION & FUTURE WORK

In conclusion, our study addresses the pervasiveness of procrastination in a highly connected and overstimulated modern era through the use of novel agent intervention techniques. While many psychological interventions and schools of thought already exist to address procrastination tendencies, these methods are often inequitable and not accessible to all. In response, we have developed a model of procrastination through the use of an artificial empirical environment as well as two agent intervention conditions composed of numerous agent intervention techniques aimed at addressing and mitigating procrastination tendencies in individuals. Our findings suggest that user procrastination can be effectively modeled and that agent interventions can provide useful tools for helping individuals reduce their procrastination tendencies.

Our future research direction is centered on the development of adaptive agents that can learn the struggles of users and assist them appropriately in the context of real-life tasks. These agents will be equipped with a range of skills aimed at helping people overcome individual struggles. For instance, agents can support people by dividing complex tasks into smaller sub-tasks. This area of research represents a promising avenue for innovation in procrastination research and enhancing procrastination interventions through agents.

ACKNOWLEDGMENTS

We thank The University of Tulsa, and the Tandy School of Computer Science in particular, for funding this work.

REFERENCES

[1] 2004. *Counseling the procrastinator in academic settings.* American Psychological Association.

[2] Hayat A.A., Kojuri J, and Amini M. 2020. Academic procrastination of medical students: The role of Internet addiction. *Journal of Advances in Medical Education & Professionalism* 8, 2 (2020), 83–89.

[3] Riku Arakawa, Hiromu Yakura, and Masataka Goto. 2023. CatAlyst: Domain-Extensible Intervention for Preventing Task Procrastination Using Large Generative Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, Article 157, 19 pages.

[4] Roy F. Baumeister and Todd F. Heatherton. 1996. Self-Regulation Failure: An Overview. *Psychological Inquiry* 7, 1 (1996), 1–15.

[5] M.E. Beutel, E.M. Klein, S. Aufenanger, E. Brähler, M. Dreier, K.W. Müller, O. Quiring, L. Reinecke, G. Schmutzer, B. Stark, and K. Wölfling. 2016. Procrastination, distress and life satisfaction across the age range–a German representative community study. *PloS one* 11, 2 (2016), 1–12.

[6] Ernest Bielinis, Jenni Simkin, Pasi Puttonen, and Liisa Tyrväinen. 2020. Effect of Viewing Video Representation of the Urban Environment and Forest Environment on Mood and Level of Procrastination. *International Journal of Environmental Research and Public Health* 17, 14 (2020).

[7] K. Constantin, M. M. English, and D. Mazmanian. 2018. Anxiety, Depression, and Procrastination Among Students: Rumination Plays a Larger Mediating Role than Worry. *Journal of Rational-Emotive & Cognitive-Behavior Therapy* 36 (2018), 15–27.

[8] Roelof A. J. de Vries, Cristina Zaga, Franciszka Bayer, Constance H. C. Drossaert, Khiet P. Truong, and Vanessa Evers. 2017. Experts Get Me Started, Peers Keep Me Going: Comparing Crowd- versus Expert-Designed Motivational Text Messages for Exercise Behavior Change. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare* (Barcelona, Spain) *(PervasiveHealth '17)*. Association for Computing Machinery, New York, NY, USA, 155–162.

[9] J. Geng, L. Han, F. Gao, M. Jou, and C.C. Huang. 2018. Internet addiction and procrastination among Chinese young adults: A moderated mediation model. *Computers in Human Behavior* 84 (2018), 320–333.

[10] Meirav Hen and Marina Goroshit. 2018. General and Life-Domain Procrastination in Highly Educated Adults in Israel. *Frontiers in Psychology* 9 (2018), 1173.

[11] Meirav Hen, Marina Goroshit, and Stav Viengarten. 2021. How decisional and general procrastination relate to procrastination at work: An investigation of office and non-office workers. *Personality and Individual Differences* 172 (2021), 1–6.

[12] Changiz Hosseini, Oda Humlung, Asle Fagerstrøm, and Moutaz Haddara. 2022. An experimental study on the effects of gamification on task performance. *Procedia Computer Science* 196 (2022), 999–1006.

[13] Alexander Häfner, Verena Oberst, and Armin Stock. 2014. Avoiding procrastination through time management: an experimental intervention study. *Educational Studies* 40, 3 (2014), 352–360.

[14] Kyung Ryung Kim and Eun Hee Seo. 2015. The relationship between procrastination and academic performance: A meta-analysis. *Personality and Individual Differences* 82 (2015), 26–33.

[15] Ann Krispenz, Cassandra Gort, Leonie Schültke, and Oliver Dickhäuser. 2019. How to Reduce Test Anxiety and Academic Procrastination Through Inquiry of Cognitive Appraisals: A Pilot Study Investigating the Role of Academic Self-Efficacy. *Frontiers in Psychology* 10 (2019).

[16] F.M. Kroese, D.T. De Ridder, C. Evers, and M.A. Adriaanse. 2014. Bedtime procrastination: introducing a new area of procrastination. *Frontiers in psychology* (2014).

[17] F. M. Kroese and D. T. de Ridder. 2016. Health behaviour procrastination: a novel reasoned route towards self-regulatory failure. *Health psychology review* 10, 3 (2016), 313–325.

[18] Tajana Ljubin-Golub, Ema Petričević, and Daria Rovan. 2019. The role of personality in motivational regulation and academic procrastination. *Educational Psychology* 39, 4 (2019), 550–568.

[19] Raquel Lozano-Blasco, Alberto Quilez Robres, and Alberto Soto Sánchez. 2022. Internet addiction in young adults: A meta-analysis and systematic review. *Computers in Human Behavior* 130 (2022), 107201.

[20] Norman A. Milgram, Barry Sroloff, and Michael Rosenbaum. 1988. The procrastination of everyday life. *Journal of Research in Personality* 22, 2 (1988), 197–212.

[21] Agnes Mutter, A.M. Küchler, A. R. Idrees, F. Kählke, Y. Terhorst1, and H. Baumeister1. 2023. StudiCare procrastination - Randomized controlled non-inferiority trial of a persuasive design-optimized internet- and mobile-based intervention with digital coach targeting procrastination in college students. *BMC Psychology* 11, 273 (2023), 1–17.

[22] Juanan Pereira and Óscar Díaz. 2021. Struggling to Keep Tabs on Capstone Projects: A Chatbot to Tackle Student Procrastination. *ACM Transactions on Computing Education* 22, 1, Article 4 (oct 2021), 22 pages.

[23] Timothy A. Pychyl and Fuschia M. Sirois. 2016. Chapter 8 - Procrastination, Emotion Regulation, and Well-Being. In *Procrastination, Health, and Well-Being*, Fuschia M. Sirois and Timothy A. Pychyl (Eds.). Academic Press, San Diego, 163–188.

[24] A. Rozental, E. Forsell, A. Svensson, G. Andersson, and P. Carlbring. 2015. Internet-based cognitive—behavior therapy for procrastination: A randomized controlled trial. *Journal of Consulting and Clinical Psychology* 83, 4 (2015), 808.

[25] Winda Liftiana Sari and Fuadah Fakhruddiana. 2019. Internal Locus of Control, Social Support and Academic Procrastination among Students in Completing the Thesis. *International Journal of Evaluation and Research in Education* 8, 2 (2019), 363–368.

[26] C. Senecal, E. Julien, and F. Guay. 2003. Role conflict and academic procrastination: A self-determination perspective. *European Journal of Social Psychology* 33 (2003), 135–145.

[27] Caroline Senecal, Richard Koestner, and Robert J. Vallerand. 1995. Self-Regulation and Academic Procrastination. *The Journal of Social Psychology* 135, 5 (1995), 607–619.

[28] Fuschia M. Sirois, Michelle L. Melia-Gordon, and Timothy A Pychyl. 2003. "I'll look after my health, later": an investigation of procrastination and health. *Personality and Individual Differences* 35, 5 (2003), 1167–1184.

[29] P. Steel. 2007. The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin* 133, 1 (2007), 65–94.

[30] Piers Steel. 2010. Arousal, avoidant and decisional procrastinators: Do they exist? *Personality and Individual Differences* 48, 8 (2010), 926–934. https://doi.org/10.1016/j.paid.2010.02.025

[31] Piers Steel and Katrin B Klingsieck. 2016. Academic Procrastination: Psychological Antecedents Revisited. *Australian Psychologist* 51, 1 (2016), 36–46.

[32] Wendelien van Eerde and Katrin B. Klingsieck. 2018. Overcoming procrastination? A meta-analysis of intervention studies. *Educational Research Review* 25 (2018), 73–85.

[33] Lena M. Wieland, Johannes D. Hoppe, Anett Wolgast, and Ulrich W. Ebner-Priemer. 2022. Task ambiguity and academic procrastination: An experience sampling approach. *Learning and Instruction* 81 (2022), 101595.

[34] Shlomo Zacks and Meirav Hen. 2018. Academic interventions for academic procrastination: A review of the literature. *Journal of Prevention & Intervention in the Community* 46, 2 (2018), 117–130.

[35] Angela Zavaleta Bernuy, Ziwen Han, Hammad Shaikh, Qi Yin Zheng, Lisa-Angelique Lim, Anna Rafferty, Andrew Petersen, and Joseph Jay Williams. 2022. How Can Email Interventions Increase Students' Completion of Online Homework? A Case Study Using A/B Comparisons. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (Online, USA) *(LAK22)*. Association for Computing Machinery, New York, NY, USA, 107–118.

[36] M. Zhou. 2020. Gender differences in procrastination: The role of personality traits. *Current Psychology* 39 (2020), 1445–1453.