Measuring Mental Effort with Visual estimation tasks

A thesis submitted in fulfilment of the requirements

for the degree of Master of Science (by Research)

by

Samarth Mehrotra

20128408

DEPARTMENT OF COGNITIVE SCIENCE

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

December 2023

Certificate

It is certified that the work contained in this thesis entitled **"Measuring Mental Effort"** by **Samarth Mehrotra** has been carried out under my supervision and that it has not been submitted elsewhere for a degree.

Digitally signed by Nisheeth Srivastava
DN: cn=Nisheeth Srivastava, o=IIT Kanpur, ou=CGS, email=nsrivast@iitk.ac.in, c=IN
Date: 2023.12.11 12:40:28 +05'30'

Nisheeth Srivastava

Professor

Department of Cognitive Science

Indian Institute of Technology Kanpur

Declaration

This is to certify that the thesis titled **"Measuring Mental Effort"** has been authored by me. It presents the research conducted by me under the supervision of **Nisheeth Srivastava**.

To the best of my knowledge, it is an original work, both in terms of research content and narrative, and has not been submitted elsewhere, in part or in full, for a degree. Further, due credit has been attributed to the relevant state-of-the-art collaborations with appropriate citations and acknowledgements, in line with established norms and practices.

Samarth

Samarth Mehrotra

Roll No. 20128408

CGS Department

Indian Institute of Technology Kanpur

Abstract

Name of the student: **Samarth Mehrotra** Roll No: **20128408**

Degree for which submitted: **MS (R)** Department: **CGS Department**

Thesis title: **Measuring Mental Effort with Visual Estimation Tasks**

Thesis supervisor: **Prof. Nisheeth Srivastava**

Month and year of thesis submission: **December 2023**

The experienced effort of the mind has its behavioural manifestations in instances like mental exhaustion. It is also intuitive to believe that humans take considerations of this said mental effort into account when making decisions. However, the expansion of how much we know about mental effort has been limited owing to the absence of direct measurements of this psychological construct. Existing measurements of mental effort are limited in that they either construct a fallacious relation of mental effort with proxies, provide at most ordinal measures, or fail to compare mental effort at different levels. The 'econometric problem of mental effort research' identifies the inability to make the amount of effort a feature of the measurement task alone as the critical reason for the failure of robust measurement of mental effort. In this thesis, we present a new measurement task that utilizes the operation of 1000 ms visual attention as the unit of mental effort, whose ability requirements have empirically been shown to not vary significantly between individuals. We use this unitary cognitive operation to construct a visual estimation task, using eye-tracking, that enables direct measurement of mental effort. This takes the choice of performing the same task in the same way but with different mental effort allocations, making the measurement a feature of the target task alone. We also report results from a simple experiment conducted using this task, which reproduces existing findings of costly effort-aversion, and also demonstrates adaptive adjustment of mental effort.

In the second part of the thesis, we utilize this measurement method to progress along the traditional economics line of decision-making research. We construct a meaningful scale of mental effort level and measure its utility in time. The particular choice of this unit of utility was taken, as applying effort towards any target precludes doing anything else in that duration. We find that the shape of the utility curve thus measured can be best described as linear.

Acknowledgements

I believe my existence to be particularly social in characteristic and my motivations tend towards the same end. I would not have been able to reach a point in my pursuit of growing up where I work on and write this thesis without my friends back at home and the friends that I made at IIT Kanpur. They know who I am talking about. I would like to thank them for allowing me to make mistakes. Especially Prof. Nisheeth.

Thank you to Prof. Srinivasan and Kaushik for getting me into philosophical readings. It stayed with me while I left.

I apologise to everyone whose intentions were misattributed as malicious by me.

कठिन था लेकिन बुरा नहीं।

Contents

List of Figures

Abbreviations

DST	Demand Selection Task
Cog-ED	Cognitive Effort Discounting
RT	Reaction Time
LMM	Linear Mixed Model

Symbols

EG	Expected Gain function over several hints
P	Condition probability of a successful hit on a number of hints
R	Reward
C	Cost
n	Number of hints
Φ	Probability density function of hints
Σ	Covariance matrix for probability density function
V	Relative decision value
d	Rate of integration of value
θ	Decision bias
v	Subjective value of the attended choice
ϵ	Random decision noise
α	Learning Rate
γ	Exploitation parameter
Q	Quality value of the option
L_x	Upper limit of parameter x

List of Publications

Publications from Thesis

1. Measuring the time utility of mental effort [1]
2. Selecting between visuomotor lotteries to measure mental effort in risky decisions. [2]

Dedicated to the past.

Chapter 1

Introduction

1.1 Motivation

The primary decision-making research, whether in economics or psychology, concerns itself with the case of valued options. Economics popularized the concept of utility and expectations in their attempt to theorize about how a choice is made between options of different value and uncertainty [2] by rational economic agents. Axioms of dominance and transitivity were established, which codified the foremost intuitions about the choice of the highest preferred option and the presence of transitivity between choices. Any agent acting contrary to these actions was considered irrational. In a domain outside, psychology was converging to these results and concluding that the expected utility theorem of economics is a fair approximation of how humans make decisions [3]. To their surprise, Kahneman and Travesky published their now-popular work on systematic biases in human decision-making that make humans appear rather irrational, which in turn gave birth to the prospect theory [4]. This began the hunt for decision-making heuristics in the field of psychology to catalogue the errors and biases in the human decision-making process. With abundant research being done on suboptimal decision-making in terms of not picking the highest value option under various environmental parameters, only a scarce amount of attention has been given to the phenomenon of decision difficulty (and sometimes even abandoning deciding at all) that humans experience when the number of options increases substantially.

With decision difficulty, the first thing to notice while understanding it in the context of familiar concepts of utility in decision-making is that the size of the options set is commonly perceived to have a utility of its own. The experience of perceived difficulty with choice when the option set is very large highlights the fact that the said utility can turn into a cost. Sheena Iyengar's research shows that participants in her study were more likely to make a purchase when their choices were limited, as opposed to a large choice set. Participants in the limited option case reported more satisfaction with their choices and performed better in tasks when they had to choose a task from a small pool of choices [5]. A descriptive economic model has characterized the cost associated with increasing option-set size as an increase in opportunity cost, where more choices mean giving up on more things when the final choice is made [6]. As opposed to a stronger claim of the cost being exclusively the opportunity cost of the choice, we purport a weaker claim of the cost being the resources depleted by mental actions made during the decision-making process, in a way similar to the metabolic cost of physical effort. Within this framework, opportunity cost might modulate the cost by utilizing mental actions of making comparisons, but the phenomenon of difficulty while making decisions makes itself feel as effortful and tiring as if depleting a resource pool, even when we are not conscious of alternative processes that we can potentially engage with. The foremost challenge that follows from the objective of studying the cost of mental effort is undoubtedly its measurement. A modest attempt at familiarizing oneself with mental effort literature would immediately introduce the 'econometric problem of mental effort' [7]. It describes the difficulty with measuring mental effort as it is difficult to separate it from the capacity of effort that varies across individuals. The current thesis attempts to walk the reader through our attempt to solve this problem.

In the rest of the current chapter, we characterize mental effort. In Chapter 2, we discuss the literature on measuring mental effort and the problems with these existing methods of measurement. In chapter 3, we describe and test our method of measurement that avoids the existing problems, followed by utilizing it to characterize time valued as a function of mental effort in chapter 4. We conclude our research in the final chapter (5).

1.2 Characterizing Mental Effort

1.2.1 Avoidance of cognitive demand

Biological beings show a tendency to follow the principle of energy conservation, which honours their limited metabolic resources. When it comes to physical effort, it is an indisputable phenomenological experience that we would prefer to minimize our effort if the utility of the effort remains unchanged. The intuitive appeal that limited metabolic resources encourage this conservative behavioral preference extends itself to the mental domain too. Modern workers do get tired and have a preference for minimizing work in their information-service vocations that are non-physical.

The preference for minimizing effort, specifically in mental terms, can be found in literature as early as the 1980s. Since then, human beings have been called 'lazy organisms' [8] and 'cognitive misers' [9], which is indicative of this preference, but only in 2010 did any empirical evidence for this be brought forward by Wouter Kool et al. [10]. Participants were presented with two decks, and a draw from either one revealed a task that needed to be performed. One of the decks demanded more effort as it resulted in more frequent changes in the tasks that needed to be performed (more task-switching demands). In this case, it was observed that participants converged to form a preference for the less effortful deck. Similar results were shown for other cognitively demanding tasks like reading passages in an abnormal orientation [11]. This aversion towards mental effort propped up its 'costly' conception [7]—something that we wish to minimize in the decision process. Two theoretical accounts try to explain this phenomenon of aversion. Each of them talks about a different limited resource that mental effort utilizes.

1.2.2 Depletable resource account

One of the accounts talks about the resource being limited in total quantity and is thus depletable, [12], similar to the depletion of metabolic resources in physical effort. Support for this theory comes from the popular literature on 'Ego Depletion', which is our experience of limited willpower resources. Ego depletion literature shows results of deteriorating performance where vigilance, executive function, and such other mental actions are required [13**?** , 14]. In addition, performance deterioration is also observed when two different tasks are performed in a sequence that requires cognitive control, suggesting that different tasks utilize a common resource [15]. A possibility of this shared common resource possibly being metabolic comes from the literature where hypoglycemia and fasting affect performance in cognitive tasks [16]. Moreover, between two ego depletion tasks, performance depletion can be reduced if a glucose drink is administered to the participant in between the tasks [17], and it is interesting to note that in this case, a placebo doesn't work [18]. Although this suggests glucose is a critical resource in all mental tasks, it is also observed that blood glucose levels don't change much under high cognitive load tasks [12]. Extrapolating from animal models, it is claimed that the brain uses glycogen stored in astrocytes as the energy reserve instead of glucose directly from the bloodstreams, thus high cognitive activity is not reflected in blood glucose levels [19].

A question of whether all or only some mental activities utilize the limited metabolic resource naturally arises. After all, mental effort in many cases is equated with cognitive control, and automatic and controlled human information processing literature talks about how automatic processes that one acquires through habituation and practice feel effortless [20]. Although this might suggest that only those processes that are not automatic require effort, the authors of automatic and controlled human information processing dismiss the idea of automatic processes being limitless in terms of not requiring resources. They believe that although in their initial task, automatic processing was not put under substantial stress, under such conditions, there will be interference in the automatic processing [21]. Thus, this suggests that all mental activity requires resources as one would expect, albeit on a different scale.

A problem with using a limited central resource bottleneck to explain a preference for conserving energy is that it misses out on explaining why humans increase their effort under conditions of reward [22].

1.2.3 Limited capacity account

The other account assumes a comparator system that values the current task being performed against other potential tasks that one can do, as it has a limited capacity to do tasks concurrently. The *opportunity cost* of the task encoded by the brain is what mental effort is [6] and it is claimed that there is no putatively depletable resource. As there is always something that we would rather do, for example, leisurely activities, we always experience mental effort. This account does a really good job at explaining the effect of incentives as they reduce the opportunity cost of the current task and also why the same mental activity can feel differently effortful in different contexts, as the context determines the value of the next best alternative. When the reward is increased for a task, its opportunity cost goes down, and it is possible to exert more effort. The assumption of a resource allocator explains the phenomenon of effort modulations under rewards. It is interesting to note that in a special case of opportunity cost models, the author claims that when we engage in a task, we commit to an exploitative nature, and the opportunity cost, in this case, is the forgoing opportunity cost to explore, which has served our genes well in our evolutionary history [23].

1.2.4 Resources with a comparator

Although the opportunity cost account explains the role of rewards and context well, it does not account for the effects of glucose on cognitive functionality. Similarly, even when we concede the presence of depletable resources, there needs to be a comparator system that allocates these resources and accounts for behavioural observations in the face of incentives.

Thus, we believe that an account that accommodates both ideas is most suitable in light of all the evidence. One such theory talks about metabolic resources being assigned by accounting for both the rewards and the remaining resource reserve, like a marathon runner would [19]. Thus, we believe that mental effort is the amount of resources allocated by this comparator system to the task.

1.2.5 Phenomenology of effort

The same set of mental operations or mental activities can have a different phenomenology under different circumstances and contexts. The most obvious example is fatigue, where based on how 'exhausted' we are, we feel more aversion towards doing the same set of mental operations. For example, reading a passage of similar length with a similar comprehension difficulty might feel more effortful at night when you are more tired. It is also the case that this phenomenon of burnout is manipulable if rewards are increased [24]. Effort 'feels' different under different conditions.

Another facet of effort phenomenology that presents itself as an opponent of the ubiquitous nature of effort aversion is reflected in the instances of mentally effortful activities like reading books, playing games, and other activities that are intrinsically motivating. Activities of this nature have their mention in the 'need for cognition' literature in psychology [25], where humans have an affinity towards cognition and have learned industriousness [26], where they have learned a baseline effort level in our industrious economy, a drop below which is experienced as a disutility.

Also, as discussed before, there are effects of automaticity [20] where substantial activities don't require effortful vigilance, control, and attention and have a phenomenology of tasks 'flowing' through with practice [27].

We are very careful and deliberate in characterizing mental effort as a resource used and differentiating it from the *sense of effort* that accompanies it, like other authors [19, 28], and by illustrating the cases above, we wish for the reader to appreciate that at the same effort level, the sense associated with it can be different. In any sense whatsoever, we will not be suggesting that the sense of effort is epiphenomenal, must be ignored while studying mental effort, and is not crucial to the process of decision-making when mental effort factors into it. Christie and Schrater claim that the comparator system signals the current resource depletion rate against the benefits of the actions being taken through the subjective phenomenology of effort, and thus it is the output of the comparator system. In agreement with this proposition, we believe that the sense of effort resides over the actual physiological mental effort. Thus, it is necessary that if we want to talk about the sense of effort, we only talk about it concerning the actual current resources being used, and that is what we wish to measure.

1.3 Operationalization

A recent review paper with contributions from well-known heavyweights of mental effort research makes an interesting proposition of conceiving mental effort in terms of information processing instead of subjective terms, to make it a subject of scientific study' [29]. It is a useful conception of mental effort if we accept the assumption that information processing of different sorts is precisely what uses resources while making decisions. It must be noted that our particular interest is in the mental effort that partakes in the decision-making process. Then, using this conception, they define mental effort as something that mediates between the information processing capacity and the fidelity of current information processing operations, as reflected in the performance levels. The problem with this definition is that, firstly, it uses another poorly operationalized concept of information processing capacity—we have no measure of an individual's entire or specific information processing capacity. Second, we do not believe that the fidelity of current information processing operations can be represented by performance levels. It can very well be the case that one performs poorly exactly because they are suboptimal operations that utilize more resources. In this case, more mental effort will be applied at a poorer performance level.

Using our characterization of mental effort, we provide a better operationalization. We agree with the information processing conception of mental effort provided and believe that different mental operations that process information at different levels are effortful or use resources at different levels. The mental effort put into a task is thus the sum of all mental operations applied to the task. This avoids the problem faced by the previous definition, as mental operations in principle have a clear characterization.

In the following chapter, we talk about how we develop a measurement of mental effort using this operationalization.

Chapter 2

Measurement of Mental Effort

We begin by describing and criticizing the current methods of measuring mental effort. We then follow up by describing and arguing for our method of mental effort measurement under the characterization given hitherto.

Our proposal characterizes mental effort as the utilization of a depletable metabolic resource. A direct measurement of the utilization of these resources through neuro-imaging methods seems obvious, but on its own, neural activity doesn't have any validity, and it is only as valid as its behavioural counterpart. The currently employed behavioural methods have fundamental limitations. We discuss the behavioural measures of mental effort that have been employed to date.

2.1 Revealed Preferences

As discussed before, one of the seminal pieces of empirical literature on mental effort is the avoidance of cognitive demand, which describes a general tendency of humans to minimize mental effort. It uses revealed preferences to measure mental effort as it gives individuals a choice between two tasks with varying amounts of cognitive demand in what is called the demand selection task (DST). Kool et al.[10], participants are given a choice between drawing cards from two separate decks, and an individual card will ask them to perform a single task from a fixed set of two: a) Parity judgment of a single number OR b) Magnitude judgment of a single number against 7. One of the decks has a higher chance of switching the task from what it was in the previous draw, thus asking for more task-switching operations, hence being more cognitively demanding. As a testament to the

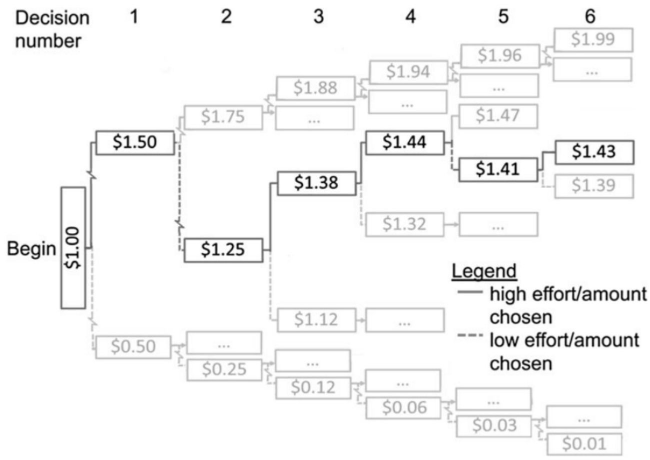

FIGURE 2.1: Economic indifference between levels of n-back task.[1]

general phenomenon of avoidance of cognitive demand, it is observed that people choose the less demanding deck more. This method is limited to allocating ordinal mental effort values to tasks and does not quantify the amount of mental effort exerted.

This paradigm is exploited in a more advanced economic indifference approach that uses an n-back task to put a load on working memory[1]. For a string of letters being recited, the participant has to respond if the current letter is the same as the letter *n iterations back*. For example, for the string 'ABBCAD**C**' in a 3-back task, participants are supposed to respond when the last letter C is being read as the same as the letter 3 iterations back. In the usual setting, the 1-back task is used as a base case, and participants are given a monetary reward on successful completion of the task. Then the reward of a higher-level task, say a 3-back task, is negotiated with the participant, and an amount is tried to be achieved at which the participant is indifferent about doing any of the tasks (Fig. 2.1). This gives us a framework for finding the subjective value of doing a higher task relative to a base case and thus measuring effort in terms of an explicit metric. However, the method discussed is not entirely sound. Participants may perform lower n-back tasks algorithmically differently from higher n-back tasks. Lieder and Griffiths show results of people selecting strategies to solve problems in a resource-rational way, i.e., they change their approach to a problem if it asks for more effort[30]. This would mean that higher effort in n-back tasks might not scale linearly. This would render the comparison between the 1-back and 2-back tasks different from the 2-back and 3-back tasks, with the difference being unknown and there being an unmeasured effort component. Thus, this method is not useful both within and across individuals.

The problem with the revealed preferences measurement of mental effort is that although it tells which task required more mental effort, it does not tell us how 'much' more mental effort it required, a problem of quantification. Although the Cog-ED paradigm seems to solve this problem by providing a subjective value evaluation of different effort levels, it fails to construct meaningful comparisons between different levels of tasks, something that we will discuss later.

2.2 Performance Measures

A prominent empirical literature within the mental effort domain is on performance deterioration, as discussed above. Performance indexed by different metrics, like reaction times (RTs), is used as a proxy of mental effort as it reflects the effort applied, under the assumption that participants are performing at capacity level. Lieder's model of strategy selection uses RTs[30]. RTs act as costs in their model of selecting strategies under meta-reasoning, with longer RTs meaning higher costs. In other words, in his model, RTs act as the cost that one must pay to perform a particular strategy. The problem with this measurement is that it is not always the case that people try to perform at the optimal level. Consider a case where an individual on a task decides to do the task as actively as they can, thus completing the task faster, and then on another iteration do it lazily, taking a longer time. As both iterations of performing the task required the same mental operations and thus the same amount of effort, RTs don't reflect accurate effort measurements. In this way, performance indexes of mental effort fail.

2.3 Current proposal

Kool and Botvinick talk about the econometric problem of mental effort as the difficulty of making mental effort the property of the target task and separate it from an individual's capabilities [7]. As discussed above, participants can choose a different algorithm to solve the same problem if given the chance to do so. Different amounts of effort can be put into a task, and in the absence of a method to compare or control the use of different algorithms, the variable effort in a task would not be measured. In addition, it is possible, even with the same algorithm, to use less effort through better abilities that are reflected by metrics like the g-factor or IQ. If individuals can process the same information with different efficiency, mental effort measures cannot be comparable across individuals.

Thus, we look for the following features in our task used to measure mental effort to solve the econometric problem:

1) Restrict the use of multiple algorithms to complete the task.

2) The operations used in the algorithm don't have significant inter-subjective variability.

In the following chapter, we discuss our design that follows these criteria.

Chapter 3

Using Visual Estimation Task to Measure Mental Effort

In the current chapter, we explain how the principles of information sampling help bridge the gap between current mental effort methods. We then use these methods to develop and test our method for mental effort measurement.

3.1 Information sampling behavior

3.1.1 Decisions from experience

The domain of optimal information sampling within psychology studies how much information humans look for before making a decision. In the popular study of decisions from experience, participants are allowed to simulate outcomes from two lotteries as much as they want before making a final draw for real monetary payoff[31]. It was observed that participants only made a median of 15 simulated draws across both decks, which is not statistically significant as this small sample distribution of the rewards is not representative of the population distribution. The sampling distribution of rare events when the sample size is small is left-skewed, so under small samples, rare events are always underrepresented (Fig. 3.1). A reason for the observed under-sampling behaviour, when in principle participants can sample as much information, can be associated with the general tendency of humans to be cognitive misers, i.e., they don't want to work hard in general to gather information. The original author of the decision from experience study talks about reasons

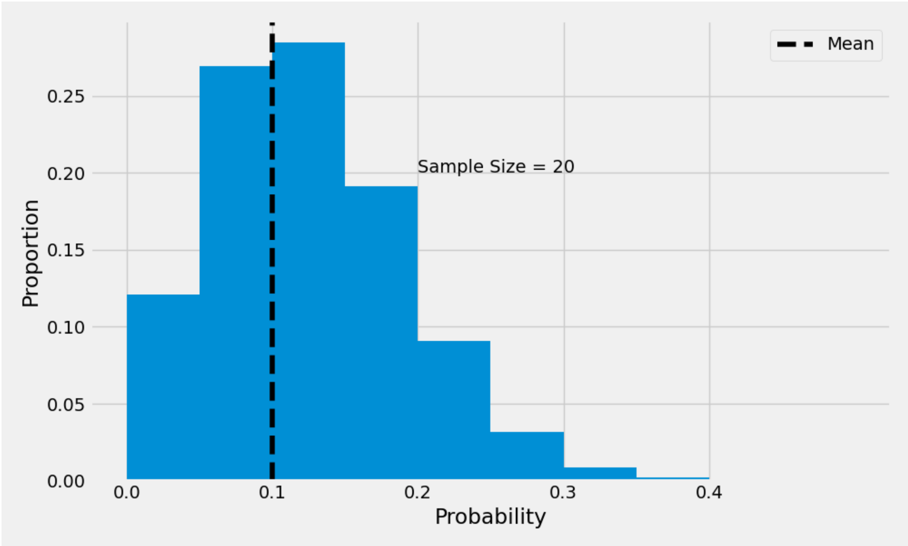

FIGURE 3.1: Representation of rare events in small samples

like opportunity cost and limited working memory for this small sampling[32], which are also characteristics of mental effort as operationalized by us.

3.1.2 Visuomotor lottery

Another potential reason for undersampling behaviour could be the absence of explicit sampling costs, as pointed out by Juni et al.[33]. In the lottery simulation task, there is no clear cost for not taking enough samples and no definition for 'enough samples' for that matter. An optimal strategy can only be ascertained by maximizing the outcome of the trade-off between rewards and costs. Hence, in the absence of an optimal strategy, the undersampling result can very well be observed if participants are using a satisfying strategy, where they take a decision as soon as a criterion is fulfilled. For example, a decision is taken when there is a significant difference between two lotteries[32].

Interestingly, when Juni et al. ran an experiment involving information sampling in a risky visual estimation task with an explicit cost of sampling, they found out that participants overestimated the optimal amount. Participants were required to estimate the location of a hidden circular target. For this, they can sequentially ask for point hints that hint at the location of the centre of the circle with some accuracy. The hints keep adding up on the screen to form a cluster. Essentially, through these hints, participants sample the mean location of the hidden target from a bivariate Gaussian distribution centred on the true location of the hidden circle and can get themselves a cluster of these dot-like hints. The information is sampled through these generated hints, which act as explicit indicators. Participants can sample as many hints before guessing by touching the touch-screen display where they guess the circular target to be. On a successful guess, they receive a reward. To make sampling costs explicit, every sampled hint reduces the reward by a fixed amount. Thus, the hint sampled had the benefit of greater confidence and the cost of reduced reward associated with it. With a cost and reward in place, an optimal trade-off was mathematically amenable and provided the explicit optimal strategy. They discuss that the likely reason for oversampling behaviour is risk-aversive preference, where participants are willing to sample more information even if it means a suboptimal decision.

3.2 Mental effort as information sampling

We discussed operationalizing mental effort as a volume of information processing. In a case where a task exclusively involves information sampling, the measurement of the participant's information sampling tells us how much mental effort was exerted. Although it is still possible that the information sampling process involves different mental operations that sample different information that we can't compare by incompetent methods, A way to mitigate this issue is to ensure that the task constitutes identifiable and comparable mental operations throughout the entire process. In the lottery sampling experiment, every sample required the same operation of observing the outcome of the lottery, while in the visual estimation task, the operation is of paying attention to the cluster of points. Thus, theoretically, the visual attention paid to these tasks is a sufficient measure of mental effort applied.

For the design of our method, we use the visual estimation task for the reasons discussed below.

3.3 Experiment 1

3.3.1 Overview

To confirm our measurement, we wished to put information sampling tasks in a demand selection paradigm. The demand selection paradigm puts two random task generators in parallel that the participant has to draw from and perform the task during the experiment duration. One of the generators has a high mental effort demand by switching the tasks more frequently, thus demanding that the participant make an effortful switch between the

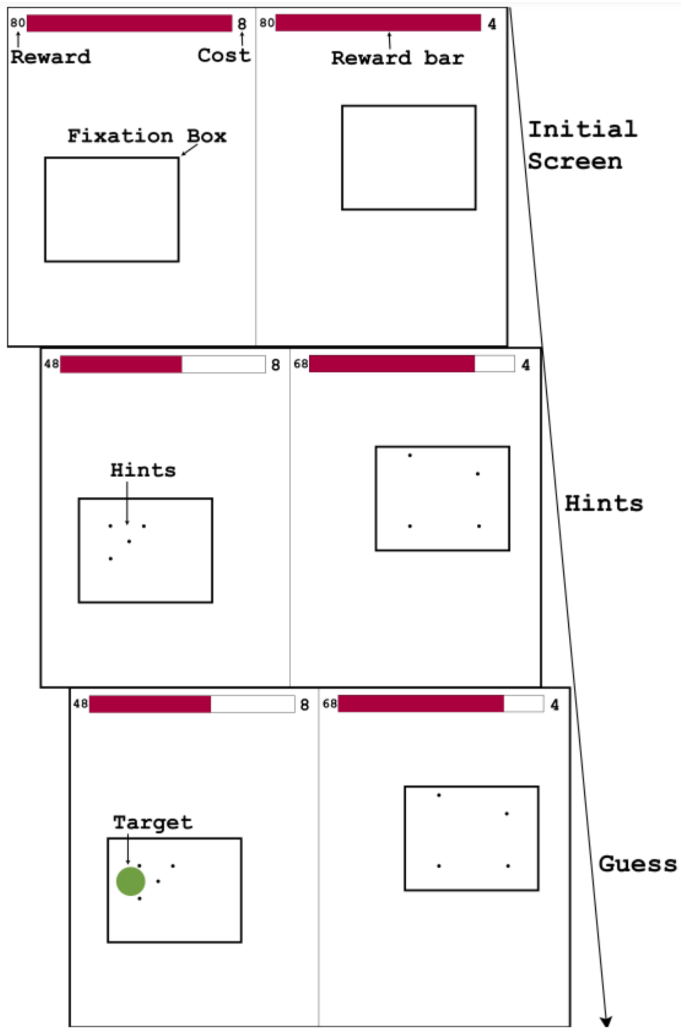

FIGURE 3.2: Sequence of experimental stimuli.

nature of the tasks that they have to perform. Similarly, we wished to give participants a choice between two visual estimation tasks similar to those of Juni et al.[33]. This was done to test that a difference in effort demand for a visual estimation task is reflected in choice and number of hints (measurement of effort), i.e., an easy task has fewer hints generated or less effort put towards it.

Any information sampling task would have sufficed and had the feature of modulating the effort demand. The reason for the choice of the visual estimation task was because it provides a normative value of the number of samples, or, in other words, a normative value of mental effort. This was useful as the previous demand selection tasks only provided information that the higher demand task is chosen less often[10]. It was interesting to ask whether choosing the less demanding option would change how much effort is being put

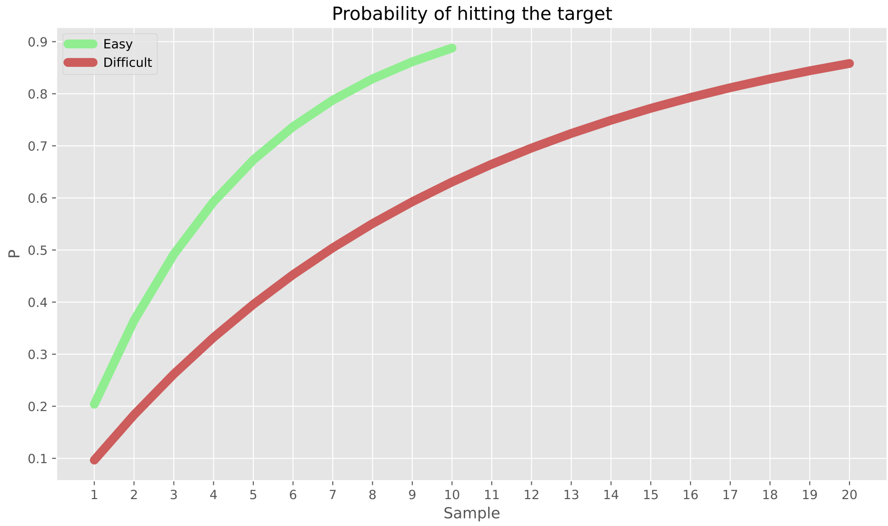

FIGURE 3.3: Probability of hitting the target.

into the more demanding task. For that, we needed a normative measure to enable us to comment on whether participants oversampled or undersampled.

This setup also solves the problems faced by the current mental effort measurement methods. First, it tells *how much* effort was put into the task in terms of the number of hints generated. Second, however, you perform on the task or how much time you take on the task, the number of hints is independent of these two measures; thus, how much work was done is reflected in the number of hints generated.

The only problem was then that attention, which was the mental operation employed, and hint generation was not as tightly coupled as it might seem at first. In the original task by Juni et al., hints were generated through key presses. Now the participants could draw hints without paying attention to each hint being generated and just press the key a couple of times to generate a bunch of hints and then pay attention once. In this case, multiple hints correspond to a single attentional action. We used eye-tracking equipment to ensure attention and hint coupling. The details of the experiment follow.

Likelihood of Success One of the estimation tasks was more difficult to guess than the other one, as the hints were less accurate. This task requires additional hints over those of its counterpart to provide as much confidence about the location of the hidden circle.

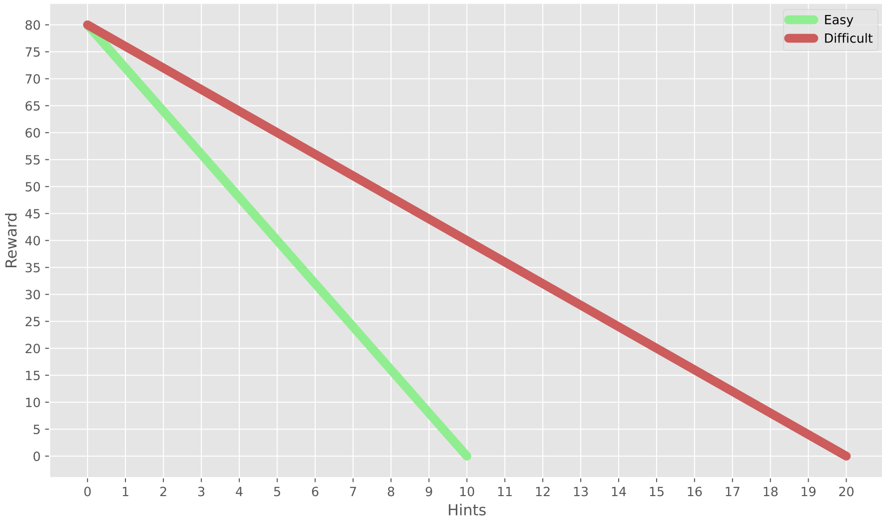

FIGURE 3.4: Potential reward.

Rewards and costs Having different probabilities of success at the same effort level would have made the choices imbalanced, so changes in the reward structure were made to adjust for this. Both of the lotteries started with a fixed amount that the participant would get indicated by the red horizontal bar on top of each estimation task. Each hint reduced the bar and thus the reward by a fixed amount, which was greater for the easy lottery. This ensured that the marginal benefit of confidence in the success that the easy lottery has at the same number of hints as the difficult one will be compensated by giving extra cost. On a successful guess, the participant would receive the reward indicated by the bar and would get no reward on an unsuccessful guess.

Designing Expectation-Matched Lotteries To ensure one of the lottery draws required more effort and was not different on any other account, it was ensured that when their respective optimal number of hints were drawn, the same reward could be expected. We use the same approach used by Juni et al. to calculate the expected reward or the expected gain as a function of the number of samples. $(EG(n))$ as follows:

$$EG(n) = P[hit/n](R - nC),$$

where R is the initial reward, C is the constant redacted from reward with each sample and $P[hit/n]$ or probability of the centre of the hidden circle being in the centre of n samples is:

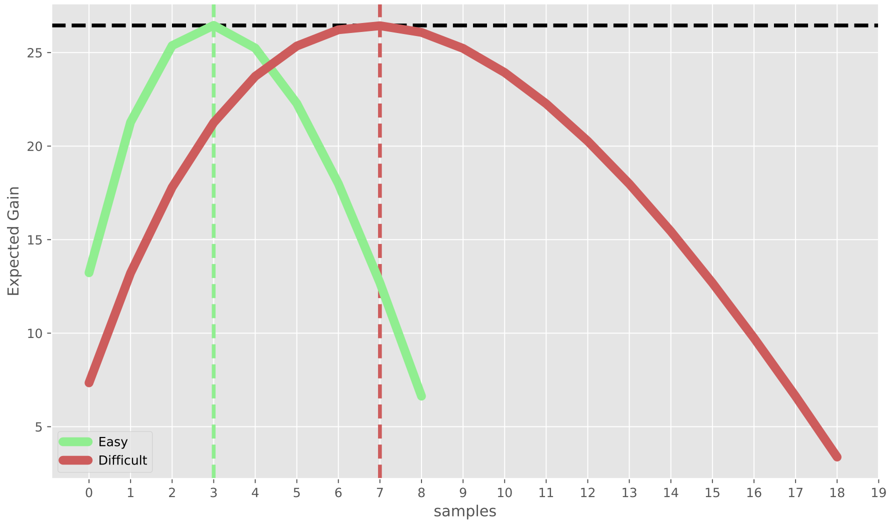

FIGURE 3.5: Expected gain.

$$P[hit/n] = \int\int_T \phi(0, \Sigma) dx dy,$$

where T is the area of the hidden circle and ϕ is the probability density function of the multivariate Gaussian from which the samples are drawn with Σ as co-variance matrix

$$\Sigma(n) = \begin{bmatrix} \sigma^2(n) & 0 \\ 0 & \sigma^2(n) \end{bmatrix}$$

Coupling attention and hint generation To ensure that one hint corresponds to a single unit of effort, that is, the work done by paying attention to the cluster of hints on the screen for target location estimation, we used eye tracking equipment. The participant needed to bring their gaze inside a square located inside the task screen allocated to a particular estimation task, which also contained the circle. This was the actual target area for the location of the hidden circle. When they did this, it indicated that they wanted a new hint, and a hint would be generated within one second of holding their gaze inside the box. This way, every time a new hint was generated, attention was already inside the target area. In the seminal work of controlled and automatic human information processing, it is noted that there is no significant difference between people's ability to focus visual attention for 1000ms[20]. Thus, this effortful work was comparable within and across individuals, which is unlikely with other mental effort paradigms that load other cognitive abilities like memory capacity[1] and task switching[10].

Attention and Preference Eye-tracking studies of multi-option choice frequently use a drift-diffusion framework for modelling the process of attending to different options and accumulating information about the value of options by sampling[34]. The option that is currently selected gets its relative value updated with a bias, and attention keeps switching until a threshold is reached in favour of any of the options.

$$V_t = V_{t-1} + d\theta v + \epsilon$$

Here V is the relative decision value, with a positive value in favour of one object and a negative value in favour of the other in the case of two option choices. d is the rate of integration of value, θ is the bias, and v is the subjective value of the attended choice, with ϵ as random noise. In this formalism, while it is easy to explain why people select options they attend to more, the choice of which item people will attend to as a function of prior preferences is not yet characterized[35]. Reward-based attention capture may cause observers to keep bringing their attention back to the option that they feel is more valuable.

Given the high visibility of the interaction between attention and preference afforded by our task, test the reward-based attention capture hypothesis, estimating participants' probability of fixating on recently rewarded options.

To this end, we fit a variation of the strategy selection model first proposed by [**?**], which tries to model strategy use in repeated choice settings like ours.

Between two choice i and j he model predicts the probability of sampling i more than j as,

$$P(s_i, t|c_i, t-1) = P_{repeat} + (1 - P_{repeat}) \times Softmax(i)$$

This equation describes the probability of sampling item 'i' (s_i), when the same item 'i' was chosen and rewarded at time t-1(c_i) as a combination of P_{repeat} or probability of repeating the choice, and a $Softmax$ function which is:

$$Softmax(i,t) = \frac{e^{\gamma Q_{i,t}}}{e^{\gamma Q_{i,t}} + e^{\gamma Q_{j,t}}},$$

where γ is the exploitation parameter, with higher values leading to the better quality option being chosen more frequently. The quality of the option itself is defined as,

$$Q_{i,t} = Q_{i,t-1} + \alpha[r_t - Q_{i,t-1}],$$

where α is the learning rate and r_t is the reward received. Quality summarizes which option has been historically more rewarding for a participant up to the current trial.

If the last rewarded item is not sampled more, the model incorporates this information as,

$$P(s_j, t|c_i, t-1) = 1 - P(s_i, t|c_i, t-1).$$

In case the previous choice is not rewarded the sampling prediction is simply,

$$P(s_i, t|c_i, t-1) = Softmax(i)$$

,

To fit the model to data, we estimated parameters that maximized the likelihood of a participant,

$$L = \prod_{trials} P$$

3.3.2 Experiment

Participants Thirty university students participated in the experiment for monetary compensation. The rate of compensation included a fixed base rate plus a variable performance-dependent component. All participants reported perfect eyesight. Data was collected with approval from the university's IRB.

Apparatus The experiment was displayed on a 1920x1080 pixel screen split vertically to display two lotteries in parallel in a dark room. A standard PC mouse was used to click and guess the position of the target. An Eyelink 1000 eye tracker was used to record gaze data at 1000Hz. A head mount was used to fix the position of the head. The PsychoPy Python library was used to create the stimuli and Pylink was used to integrate the eye tracker.

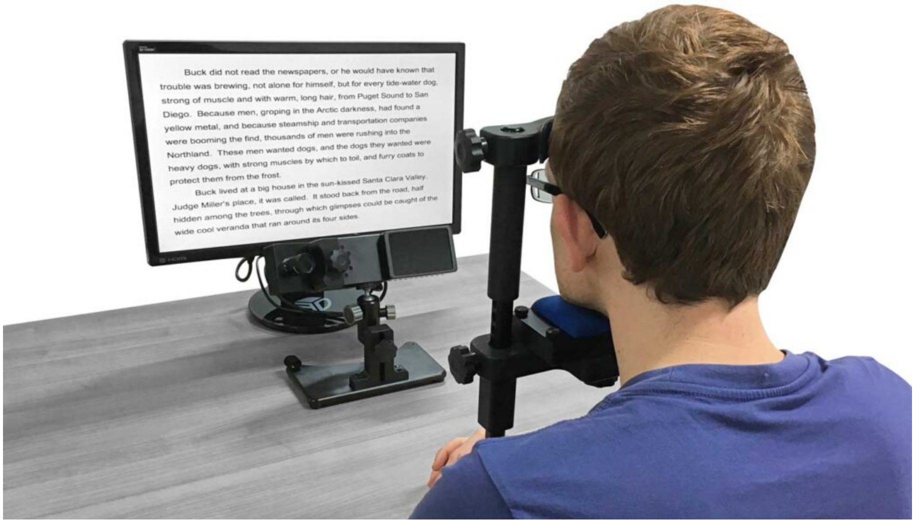

FIGURE 3.6: Eyelink 1000 plus.

Stimuli A circle of radius 24 pixels was used as the target. The screen was split vertically into two equal sides. Each side displayed a bar on top, with a number on the left side of the bar indicating how much potential reward the participant could earn concerning the corresponding lottery, which was 80, to begin with for both lotteries. Hints were dots with a 4-pixel radius. The standard deviation used for one of the multivariate Gaussian was 40 pixels (easy lottery), and it was 60 pixels for the other one (hard lottery). The standard deviations were not revealed to the participants, but the lottery with the higher SD had 4 written to the right of the reward bar, which indicated the reward reduction value with each hint. Similarly, 8 was the constant reduction rate for the other reward. No reward was received for the wrong guess. Participants were allowed to draw hints until the reward bar went to zero, in which case the trial would terminate with a zero reward. The fixation box inside which the participants needed to foveate for a minimum of 1000 meters was a rectangle with 1/4 the height and width of the screen.

Design Each participant went through a block of validation trials to learn to fixate on the displays to generate samples. In these trials, they had to make 90% successful fixations out of 30 total attempts to proceed to the main experiment, as detailed below. All participants completed the validation block. They then did 10 practice trials of selecting between visuomotor lotteries before doing 50 main ones. The left-right position of the lotteries was randomized on each trial.

Procedure The participants were shown the instructions and a play-through of how the experiment would look like and calibrated the eye-tracker. They then had to go through validation trials in which they had to get used to the fixation process. For 30 times, a box appeared on one side of the split screen and the participants had to bring their gaze inside that side and fixate inside the box for 1 second in a limit of 5 seconds in total. If they were successful, the box turned green, otherwise, it turned red. The box appeared in orange colour to inform the participants that they had to 'bring' their gaze inside that side of the screen or in other words, they needed to look outside the area of the corresponding lottery to turn the box black again and enable another hint to be drawn. The 'bring the gaze inside' step in our procedure made sure that the fixations they make while estimating the location of the target don't result in drawing an indiscriminately large number of hints.

After validation, practice and main trials began, with participants free to sample hints for either of the targets at any time. If they drew a hint for a given target, the corresponding box turned orange, indicating the need to bring their gaze inside to draw another hint. Participants were told to use the left mouse button to guess the position of the circle whenever they wanted to. On a correct guess, the circle appeared in green colour and red otherwise. The points won for the particular trial were indicated afterwards, and the next trial began.

Before the beginning of the validation block and before every main trial, participants were given the option to re-calibrate the eye tracker. This was done so that they had the option to take a rest by dismounting their head between any two trials.

3.3.3 Results

The analysis mainly includes using split-half analysis and linear mixed models. In a split-half analysis, the main trials of a participant are split into a first and second half along the sequential axis along which the trials occur. It is used to confirm that the results obtained along all the trials are not obtained because of dependence or auto-correlation of trials. In linear mixed models (LMM), the slope and intercept as random effects of participants. This ensures the variability of individual participants is modelled within the linear relationship that we are interested in through the slope and intercept.

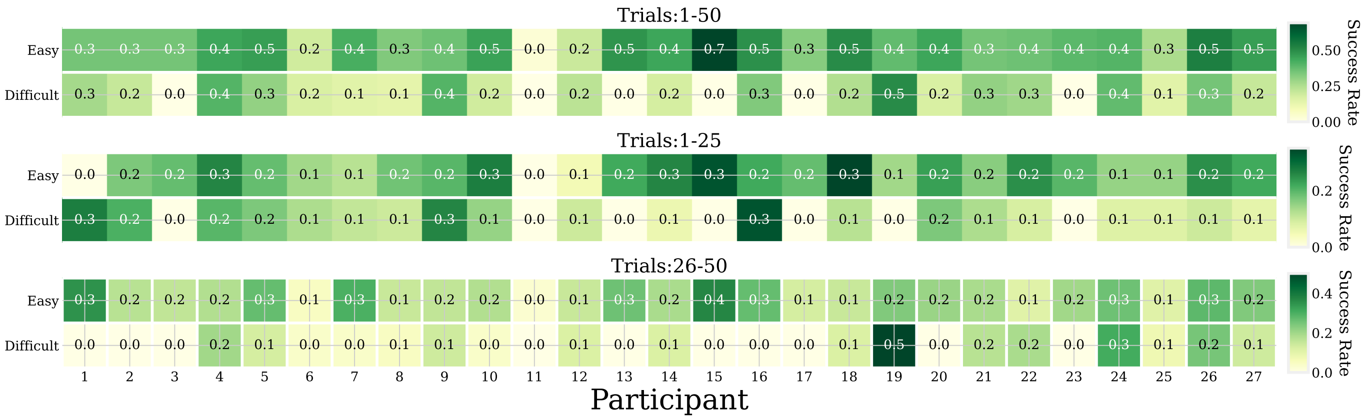

FIGURE 3.7: Choosing an easier task results in more success.

3.3.3.1 Baseline Task Performance

As our first analysis, we check whether participants' performance matched baseline expectations of having been sensitive to information-theoretic task characteristics. In particular, we expected participants to have a lower success rate for the difficult task, to have drawn more samples for the difficult task, and to have seen greater success in hitting the target after having drawn more samples. Two of these three baseline expectations hold in our participants. The success rate of a participant, measured by the number of successful guesses over selected tasks, for easy tasks was significantly greater than for difficult tasks. Figure 3.7, top:

$$wilcoxon\ test\ z(26) = 20.5, p < 5e - 06$$

This result was consistent in a split-half analysis. Figure 3.7, middle and bottom:

$$1^{st}\ Split : wilcoxon\ test\ z(26) = 57, p = .002$$
$$2^{nd}\ Split : wilcoxon\ test\ z(26) = 49, p = .0003$$

The success rate increases significantly with an increasing number of samples as the slope of the best linear mixed model fit was significantly positive. Figure 3.8, top:

$$slope = 0.05,\ z = 4.09,\ p = 0.00$$

This result held for a split half analysis as well:(slope = 0.058, z = 5.6, p = 0.00 and slope = 0.025, z = 3.29, p = 0.001). Figure 3.8, bottom lef and bottom right:

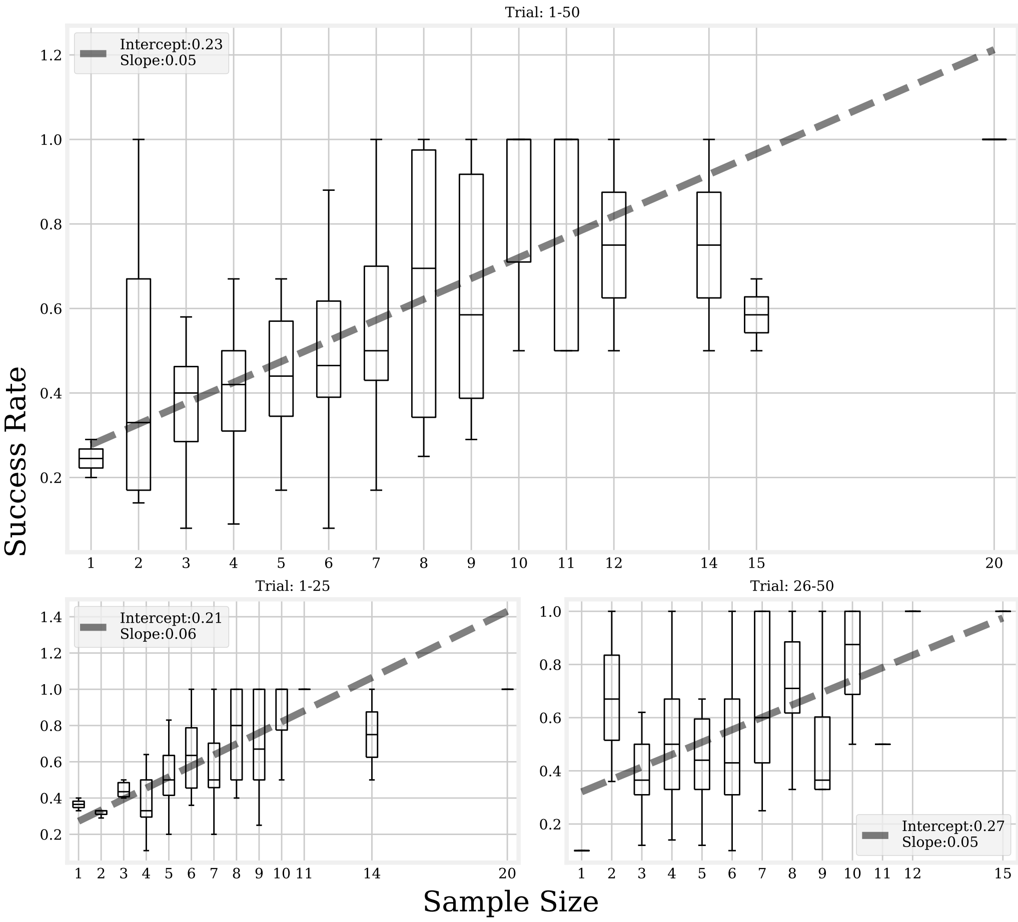

FIGURE 3.8: Success increases with sample size.

$$1^{st}\ Split : slope = 0.06,\ \ z = 9.8, p = 0.00$$
$$2^{nd}\ Split : slope = 0.05\ z = 4.1, p = 0.00$$

Another linear mixed model for choice versus sample size data suggests that the sample size for hard choice was significantly higher than that of easy choice. Figure 3.9, right:

$$slope = -2.1,\ \ z = -6.1,\ \ p = 0.00$$

It too was consistent with a split half analysis. Figure 3.9, left and center:

$$1^{st}\ Split : slope = -2.3,\ \ z = -6.4, p = 0.00$$
$$2^{nd}\ Split : slope = -1.9\ z = -4.1, p = 0.00$$

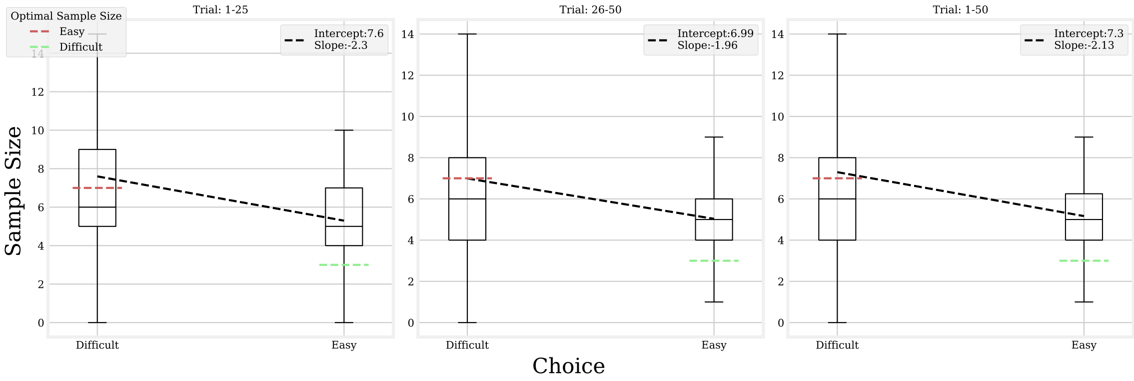

FIGURE 3.9: Difficult task requires more sample.

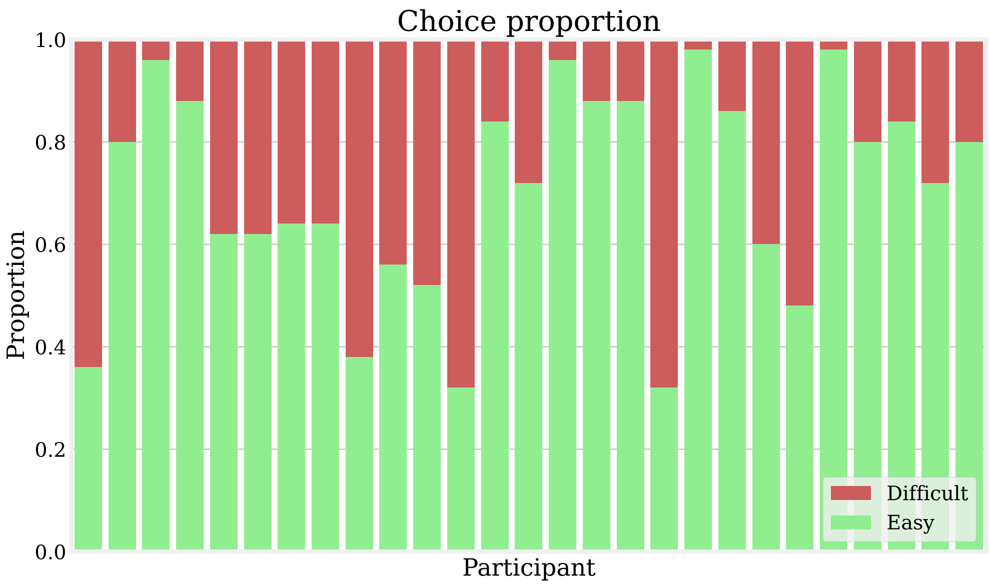

FIGURE 3.10: Preference for easier task.

This result which was not expected will be explored further later. Otherwise, participants' responses were sensitive to the information processing demands of the task.

3.3.3.2 Preference for easier

Figure 3.10 shows that most participants had a preference for the easy lottery, represented by the green part of the bar. Across all participants, this preference was significant with one sample proportion:

$$proportions\ test\ z(49) = 2.86, p = 0.00$$

This result was consistent in a split-half analysis:

$$1^{st}\ Split : proportions\ test\ z(49) = 1.73, p = .04$$
$$2^{nd}\ Split : proportions\ test\ z(49) = 2.32, p = .0$$

This revealed preference, prima facie, can be interpreted as avoidance of cognitive demand [10]. Effort aversion is also evident in participants' behaviour across time in our experiment. Across all participants, the correlation between the trial number and total samples in that trial was significant:

$$spearman\ r(26) = -0.06, p = 0.023$$

This negative correlation means that participants drew fewer samples on later trials across the entire range of the experiment. It must be noted that although negative, the correlation was not consistent in either of the split halves:

$$1^{st}\ Split : spearman\ r(26) = -0.04, p = 0.27$$
$$2^{nd}\ Split : spearman\ r(26) = -0.05, p = 0.12.$$

3.3.3.3 Mental effort in sampling lotteries

In Figure 3.11, red bars represent the average number of samples taken for the harder lottery and green bars show the average samples taken for the easy lottery by all participants. The green and red vertical dotted lines show the optimal number of samples for both lotteries respectively - 7 and 3. As can be seen in figure 3.9, fitting a linear mixed model with participant again being the random effect tells us that they systematically oversampled both the hard and easy lottery - 7.3 and 5.16, respectively - we find significant oversampling for both the hard and easy lottery as seen in Juni et al.[33], in comparison with their respective theoretical norms.

$$Slope = -2.13, z = -6.1, p = 0.00$$
$$Intercept = 7.3, z = 12.9, p = 0.00$$

This was also consistent in the split-half analysis:

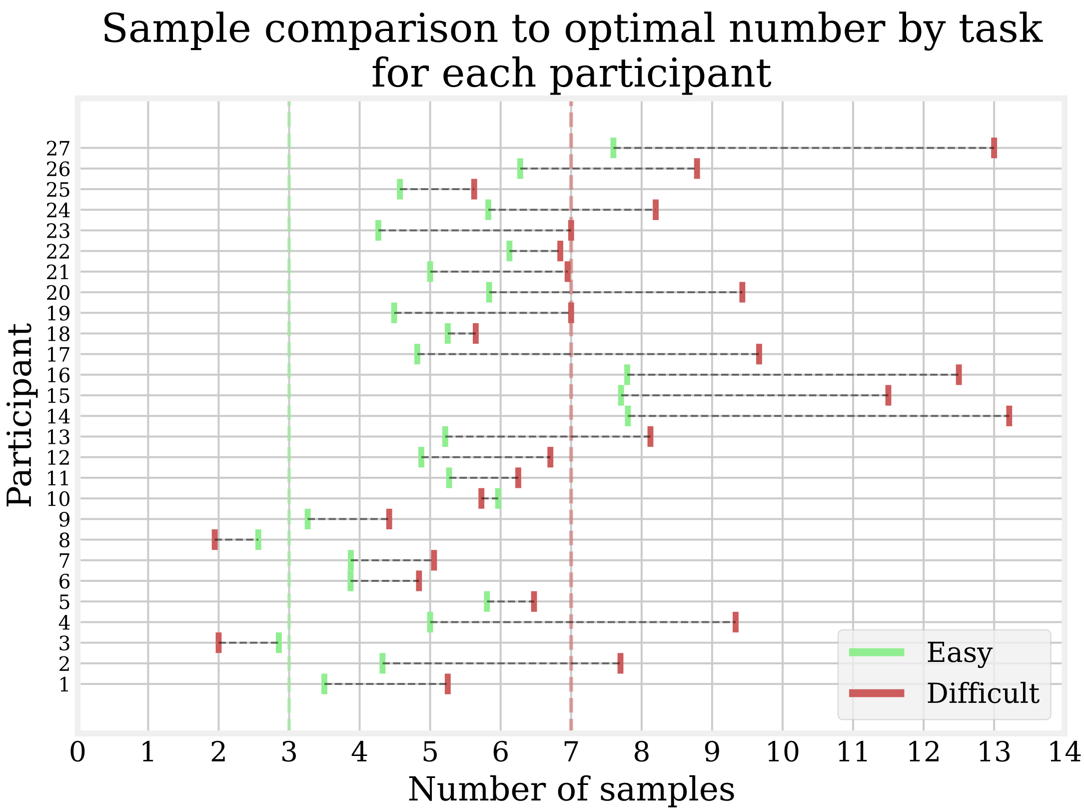

FIGURE 3.11: Oversampling in both tasks.

$$
\begin{aligned}
&1^{st}\ Split: \\
&\qquad Slope = -2.3, z = -6.4, p = 0.00 \\
&\qquad Intercept = 7.6, z = 14.2, p = 0.00 \\
&2^{nd}\ Split: \\
&\qquad Slope = -1.96, z = -4.1, p = 0.00 \\
&\qquad Intercept = 6.99, z = 9.9, p = 0.00
\end{aligned}
$$

We also tested if participants were responsive to success and failure, by adjusting their sampling effort. For this, we measure the average change in the number of samples in the choice following success and failure. Our findings are illustrated in figure 3.13. Fitting an LMM to both the entire data and split for both the lotteries separately, we only find significant results for the easy task's entire and first half split data. This suggests that, at least at the cohort level, participants responded resource rationally to the task design for

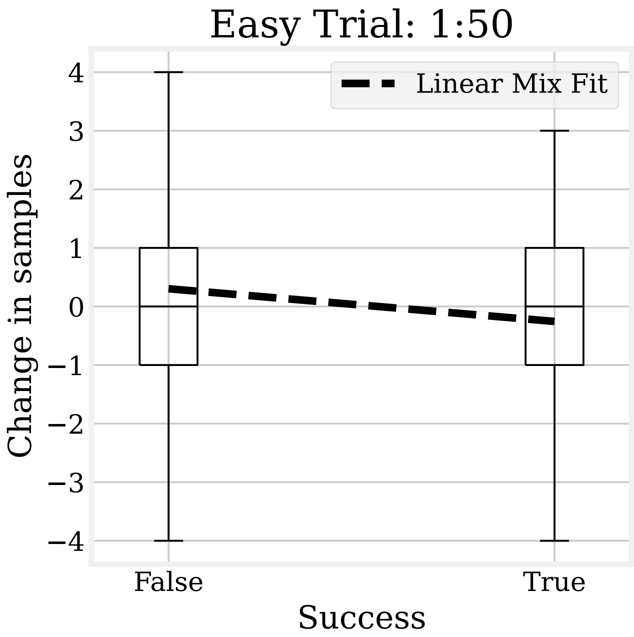

FIGURE 3.12: Probability of attending more to the choice rewarded in the last iteration.

the easier task, calibrating effort towards the minimum value needed to produce success, but this adaptive calibration died down in the second half and did not happen at all for the harder lottery. Participants increased the sampling post-failure (+0.3) and decreased it on success (-0.5). Given the intrinsic difficulty of the harder lottery and the systematic under-sampling seen for it, it is unsurprising that participants were not able to calibrate effort for it. The LMM results for easy lottery for all the trials are given:

$$Slope = -0.5, z = -4.2, p = 0.00$$
$$Intercept = 0.3, z = 3.5, p = 0.00$$

3.3.3.4 Attention and Preference

Figure 3.13 shows that the probability of fixating more on the rewarding option is significantly lower than 0.5 at the cohort level:

$$one\ sample\ proportion\ z(49) = 0.8, p = 0.2.$$

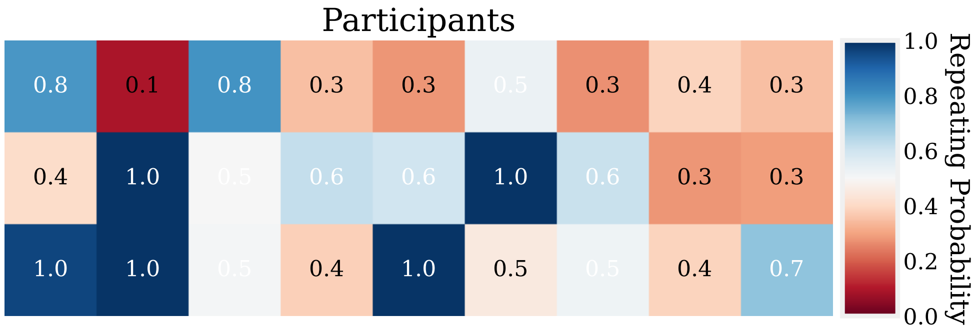

FIGURE 3.13: Probability of attending more to the choice rewarded in the last iteration.

The current findings suggest that, whereas reward-based attention capture appears to not dominate the decision of what to attend to next in our task, it could explain some participants' behaviour quite well.

3.3.4 Discussion

The author of Decisions from Experience lists multiple reasons for observing small samples[32]. The reasons were discussed within the framework of the task used in the paradigm, i.e., the lottery simulation task that was discussed above, but they can be generalized to other tasks and should be checked for:

1. **Memory demands**: In the task, there were two lotteries, and both had two possible outcomes: high and low reward. Every simulation that occurred needed to be counted. In the usual case, they might rely on their sense of how many times an event occurred, but with a larger number of occurrences, memory needs to be employed, thus a potential reason for small samples.

2. **Low motivation**: The author posits that as it was a lottery scenario, most of the outcome was chance-based, and there was no clear motivation associated with doing more simulations.

3. **High perceived difference**: The author talks about how small samples have a higher chance of misrepresenting the population. Thus, small samples may lead to larger differences between the outcome values of both choices. This larger difference makes the choice seem easy at this point, and thus the participants are drawn to committing to making an early decision.

In our task, none of the above can be the case. All the sampled information was always visible on the screen in the form of point hints; thus, the task had no memory demand to hold information. A wrong guess would result in zero rewards, and every sample substantially added to the success probability; thus, by design, the participants were motivated to draw more samples. Finally, the design of the experiment also ensured that at low samples, the difference in expected gain is not substantial, as can be seen in figure 3.6. In conjunction with the fact that our method tells how much effort was made in the task in unitary terms, with the unit being comparable within and across participants, our method is more robust than any other existing method.

The major conceptual limitation of our task is that we do not account for the possibility of withdrawing covert attention from the task. That is, someone may attend to the task overtly via eye movements while being mentally disengaged. While this possibility does not appear phenomenologically salient in the task in our experience with it, it is certainly theoretically realistic[36]. Future work may empirically measure covert attention shifts in our task by examining on-task microsaccades[37]. Multiple experimental directions also present themselves immediately for future investigation. For instance, combining our task with the economic indifference approach[1] should help estimate effort-cost curves using direct effort measurements. In the following chapter, we use the same paradigm to establish a relationship between effort, demand, and time.

Chapter 4

Time and Effort

4.1 Introduction

The utility function is a fundamental mathematical device in the analysis of microeconomic behavior[38], with the metaphor of utility extending deep into the study of behaviour more generally, with ideas of utility maximization and risk aversion richly informative about people's behaviour in both natural and artificial settings[39, 40].

It has recently been realized that it is possible to meaningfully characterize the phenomenon of mental effort, previously mostly conceptualized in philosophical[41] or biophysical terms[42, 43] in an economic framework resembling utility maximization. We 'invest' our mental effort with things like education and we 'pay' for tasks like taking care of our young ones with mental effort[7], an insight substantiated by empirical results supporting the avoidance of cognitive demand[10]. This line of work has solidified the appropriateness of using the economic metaphor of 'cost' in the study of mental effort - something that we prefer to reduce, given a choice[44].

On this account, positive utility accrued from performing a mental task is offset by a *negative utility or cost* incurred by mental effort. By permitting the operationalization of mental effort in the same mental currency units as monetary preferences or other overt markers of value, such accounts enable econometric methods to be used to characterize the arithmetic relationship between these negative costs and positive utilities.

The shape of this curve has interesting theoretical implications. One possibility is a linear relationship where every additional amount of effort adds a fixed cost, implying that people are willing to continue working harder with better economic incentives. Empirically

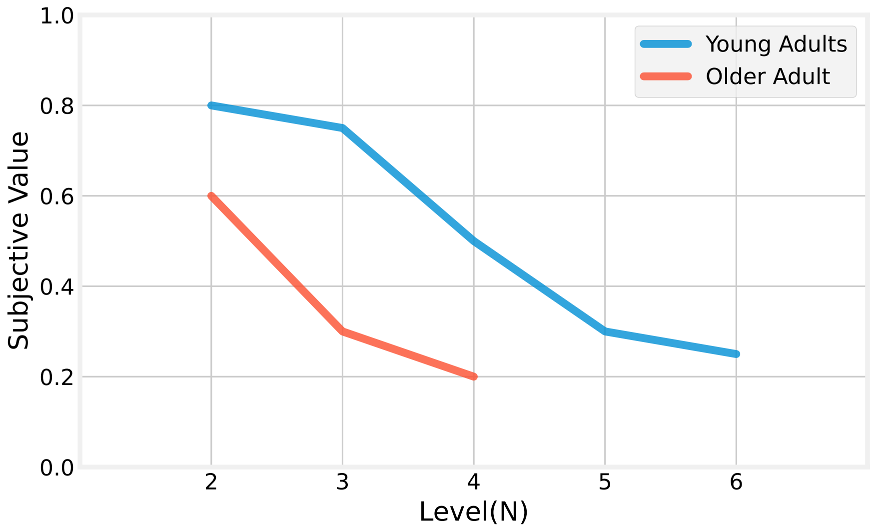

FIGURE 4.1: Subjective Value of Old Adults(OA) and Young Adults(YA). Re-plotted fromwestbrook2013subjective

measured labour-leisure indifference curves, however, are convex in shape[45], where every unit of effort contributes an increasing cost to the equation suggesting there is some 'satisficing' solution for effort level in any given task[46]. A convex relationship where increasing effort has decreasing cost would imply, counterintuitively, that greater effort is attractive rather than aversive.

The cognitive effort discounting (henceforth Cog-ED) experimental paradigm uses monetary reward to measure the utility of mental effort, which they call 'subjective value' with the understanding that mental effort can be measured in the same mental currency as task utility[47]. In this experimental paradigm, an n-back memory task is used to stimulate varying levels of effort, and the participants are given trade offers between doing a base-level task (1-back) at a lower reward or a higher n-back task for a higher reward. Keeping the the higher reward and base level task fixed, the compensation for the lower task is modulated at every individual n-back task (n=2 to 6), until the subject is indifferent between doing the higher and lower n-back task. This obtained indifference value normalised by the higher reward magnitude yields the subjective value, which serves as a proxy for mental effort.

Figure 4.1, re-plotted from [47], shows a monotonically decreasing relationship of subjective value (increasing costs) with increasing mental effort, as assessed using the Cog-ED paradigm. However, attempts to characterise a relationship curve between variables as one of linear, concave or convex, as discussed above, assume the scale of cost and utility measures as linear. One of them being non-linear would change the shape characteristic of the curve. The linear progression on the n-back task level scale might not necessarily translate to linear progression on the effort scale. This was discussed in 2 in the context of Cog-ED being unable to quantify the difference in level of effort between tasks. The issue of possible non-linear response characteristics is rooted in this inability. Thus, econometric analyses of curves obtained using Cog-ED and related methods cannot proceed assuredly. This issue, familiarly known as the 'econometric problem of mental effort'[7] points to two prominent confounds in existing measurements of mental effort: a) individual differences between processing the same information and b) subjective allocation of mental effort; which make its measurement non-trivial.

4.2 Solving the econometric problem

In a nutshell, mental effort while making decisions is conceived in information processing terms to operationalize it as an object capable of quantitative study[29]. But how much information is processed while doing a task also depends on the confounds we describe above. Translated to the n-back memory task, participants can choose to process more information by trying to remember details essential for the completion of the task. Even if they do the task the same way, they display individual differences in how efficiently they process information in each step, with practice enabling a reduction of effort for the same level of performance[48]. Therefore completion of an n-back trial for a fixed n might take different levels of mental effort for different subjects, or even for one subject across different trials.

Further, even if this problem is solved, every additional unit of the independent variable (x-axis in Figure 4.1) should correspond to the same additional unit of the dependent variable (y-axis), if we are concerned with characterising the shape of the curve. In the simplest case, scales should be linear. It is uncertain whether the expected mental effort difference between the 1-back and 2-back task is the same as that between the 2-back and 3-back task; certainly, the response dynamics become highly non-linear as participants approach their working memory capacity limits.

Thus, a solution to the econometric problem requires experimenters to use effort tasks with low inter-subject variability, low potential for across-trial learning, and low potential for distraction. To this end, we replaced the n-back task in the Cog-ED paradigm with the visual estimation task in 3[49]. We demonstrate that it provides an approximately linear scale for mental effort.

Additionally, instead of measuring the subjective utility of mental effort in terms of monetary units, we measure it concerning time. The opportunity cost of mental effort is naturally measured in time units since mental effort precludes doing something else at any point in time. Therefore, measuring the utility of mental effort in time units makes more ecological sense than doing so in monetary units. Thus, in this paper, we present a new approach to measuring the time utility of mental effort.

4.2.1 The visual estimation task

The visual estimation task developed in the previous Chapter makes considerable progress towards solving both parts of the econometric problem of constructing a linear scale of measurement of mental effort. The action of generating hints is mapped to visual fixations of 1000ms which are found to be insignificantly variant inter-subjectively[48]. The number of hints sampled provides a simple parametric measure of effort. With an increasing number of hints, the probability of success increases and the potential reward decreases. The expected reward can, in turn, be calculated as a function of the number of hints by multiplying these two measures and the optimal number of hints (Figure 4.2).

For our purposes, we modify this task by removing the cost element from it and imposing repetition of the trial until success is achieved, shifting the nature of cost from intangible 'money' to tangible 'time'. With the number of hints sampled remaining the unit of effort, manipulating the expected number of hints for success would mean manipulating the level of effort, which was the driving force of our study. Measuring tradeable time at the indifference point at multiple such levels of effort maps out a time utility curve for a single participant. To do this, we had to recalculate the expected number of hints in this modified setup.

4.2.2 Calculating optimal effort

The original task set the difficulty or the level of effort required to complete the task by changing the standard deviation of the bi-variate Gaussian distribution from which the location of hints was sampled. This distribution was centred on the hidden circle and thus the point hints that appeared essentially hinted at the probable center of the hidden circle. Generating a few such hints that would stay on the screen would give a sampling distribution of the centre of the hidden circle thus indicating a probable area at which the participant could take a guess.

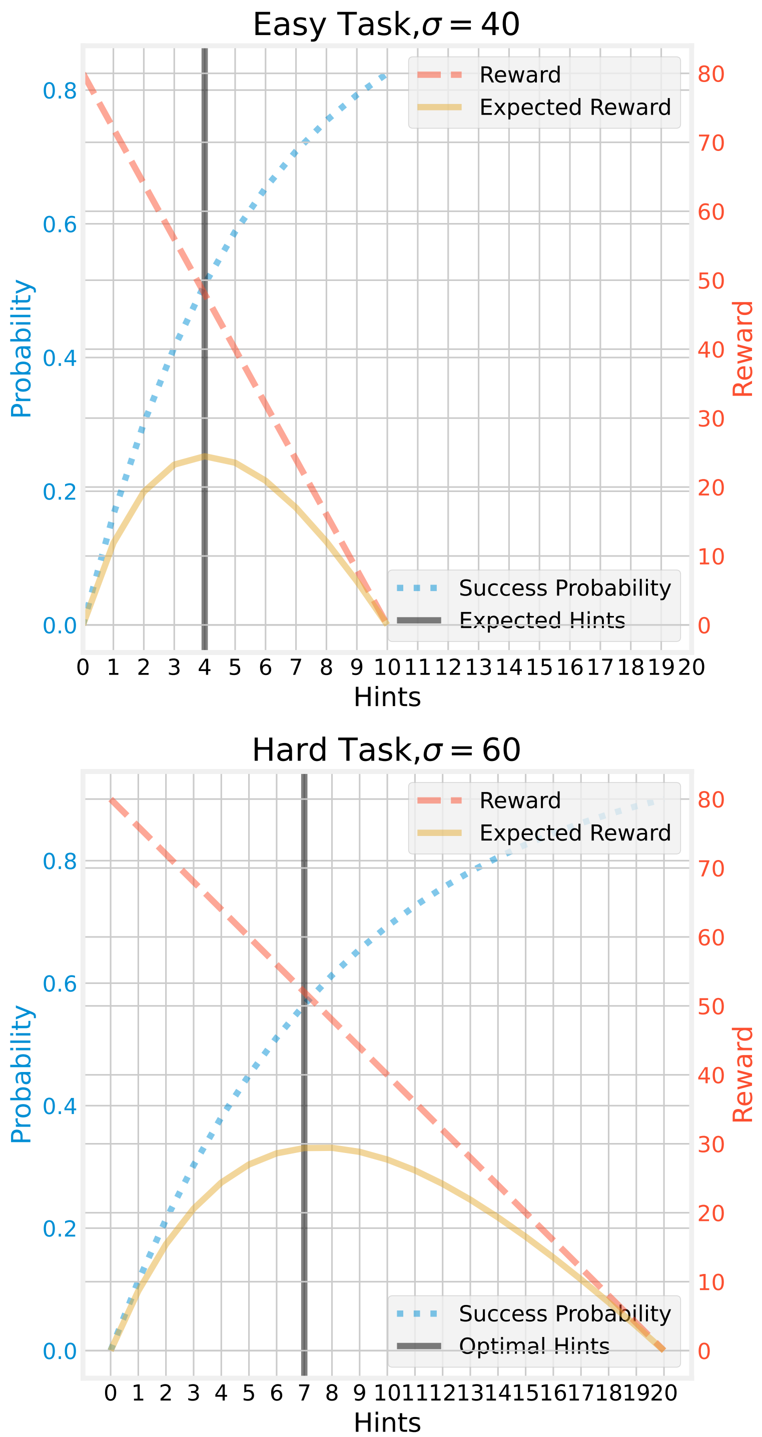

FIGURE 4.2: Expected number of hints for tasks of different difficulty levels. The mode of the success probability curve gives us the optimal number of hints for a given difficulty level.

In a given trial of our modified visual estimation task, the participants had to repeat the task until they became successful. The probability of hitting the target in n hints as a function of σ, the standard deviation was:

$$P[hit/n] = P(\sigma) = \int \int_{T} \phi(0, \Sigma) dx dy,$$

where T is the area of the hidden circle and ϕ is the probability density function of the multivariate Gaussian with Σ as co-variance matrix:

$$\Sigma(n) = \begin{bmatrix} \sigma^2/n & 0 \\ 0 & \sigma^2/n \end{bmatrix}$$

Participants were free to sample as many hints as they wanted before making a guess but errors led to the process restarting. In this process, the expected number of hints for a given effort level (σ) is simply the probability of being successful at the current attempt and not being successful at any previous one weighted by the number of hints taken in respective attempts, summed over all possible number of attempts and hints. This can be specified by the following equation:

$$EH(\sigma) = \frac{1}{L_n} \sum_{n=1}^{L_n} n \sum_{a=1}^{L_a} aP[hit/n] \prod_{0}^{a-1} (1 - P[hit/n]),$$

where $EH(\sigma)$ is the expected hints which is the function of σ, n is the number of hints, a is the number of attempts, L_n is upper limit of number of hints and L_a as the upper limit of number of attempts. An analytical solution of this equation above would consider the summations over infinity but an appropriate numerical approximation estimated from observational data serves our purpose, as we show further below. Modulating the expected hint count as a function of effort level constructs a scale for mental effort, whose properties we also examine in more detail below.

4.2.3 COG-ED with visual estimation

We incorporated the visual estimation task designed above in the cognitive effort discounting paradigm[47] which in our case ascertains an indifference point where the participant would be indifferent between doing the visual estimation task at a particular effort level or simply waiting for a stated duration of time. For a given effort level, a participant was given an option to do the task or wait (wait task) for time t which was sampled from the range 0-40 seconds in factors of 5. In the wait task, participants had to fixate at a fixation cross for the said duration. If participants accepted the time trade, all lower values from the range were discarded from further sampling transitively assuming lower values to also be preferred against a task at the current effort level. Similarly on the rejection of the wait task upper values were discarded. This process was repeated until a single value was left in the range and this value was regarded as the point of indifference - a value at which the participant is approximately indifferent to doing the visual estimation task or the wait task.

4.3 Methods

Here, we describe methodological details for an experiment we conducted using the Cog-ED visual estimation task.

4.3.1 Participants

30 university students (11 females, 19 males; mean age: 22.3) participated in the experiment for monetary compensation (100). All participants reported perfect eyesight. Data was collected with approval from the university's Institutional ethics committee.

4.3.2 Apparatus

The experiment was displayed on a 1920x1080 pixel screen in a dark room. A standard PC mouse was used to click and guess the position of the target. An Eyelink 1000 eye tracker was used to record gaze data at 1000Hz. A head mount was used to fix the position of the head. The PsychoPy Python library was used to create the stimuli and Pylink was used to integrate the eye tracker.

4.3.3 Stimuli

For the visual estimation task, a circle of radius 24 pixels was used as the target. Hints were dots of 4 pixels radius. The standard deviation used for five effort levels was 70 to 110 pixels. The standard deviations were not revealed to the participants, only the 'difficulty level'. At the beginning of every trial, two cards (1/8 of the screen size) were displayed in the middle of either side of the screen, displaying two choices: wait 'T' seconds or do the visual estimation task at 'D' difficulty. A correct guess revealed the circle in green, otherwise red after the participant had responded. In the latter case, we generated the visual estimation task again at the same difficulty level but with a different location of the hidden target. The fixation box inside which the participants needed to foveate for a minimum of 1000 ms was a rectangle with 1/4 the height and width of the screen. For the wait task, the fixation cross was 200 pixels in height and width.

4.3.4 Design

The factor of 'expected hints' with 5 levels indexed the effort-demand level. The dependent variable of 'the point of indifference' was measured in time units. Participants had to do each level once, with the number of fixation tasks and estimation tasks varying within each level, based on participant choices.

4.3.5 Procedure

The participants began with the familiarisation phase where they had to succeed in the visual estimation task at least five times at every effort level. The average time taken to succeed at all levels was equally weighted with each level's time to create a reference point for each level. In the main phase, they were shown all the relevant stimuli and instructions and were told what randomly picked 'difficulty level' they would be going through. A fixation box appeared randomly positioned on the screen, and they had to bring their gaze inside it and fixate for 1000 ms to generate a dot-like hint. They could generate as many hints as they wanted, and when they were confident, they had to point and click with the mouse where they thought the circle was, and then the circle revealed itself in either green or red colour based on whether the guess was correct or not, respectively. If incorrect, the trial started again with the fixation box in a new randomly generated position. They had to repeat this for every difficulty level.

In the main phase, they were given a choice between two options presented on two cards. They could either choose to do nothing for 'T' seconds, in which case they would have to fixate on a cross for that said duration, or they could do the task at 'D' difficulty. For a given difficulty level, this was repeated until the indifference point was found.

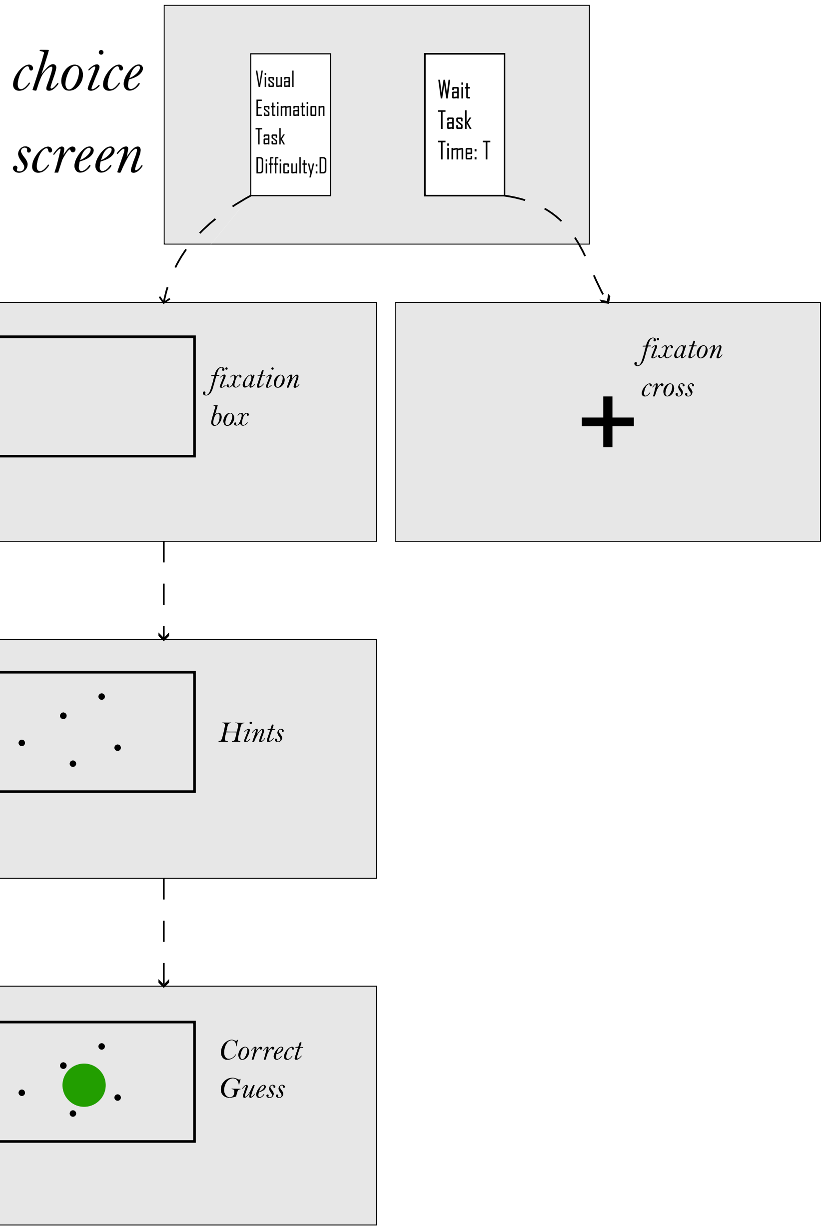

FIGURE 4.3: Illustration of the experiment setup. Upon making a choice the participant either does the visual estimation task or the wait task.

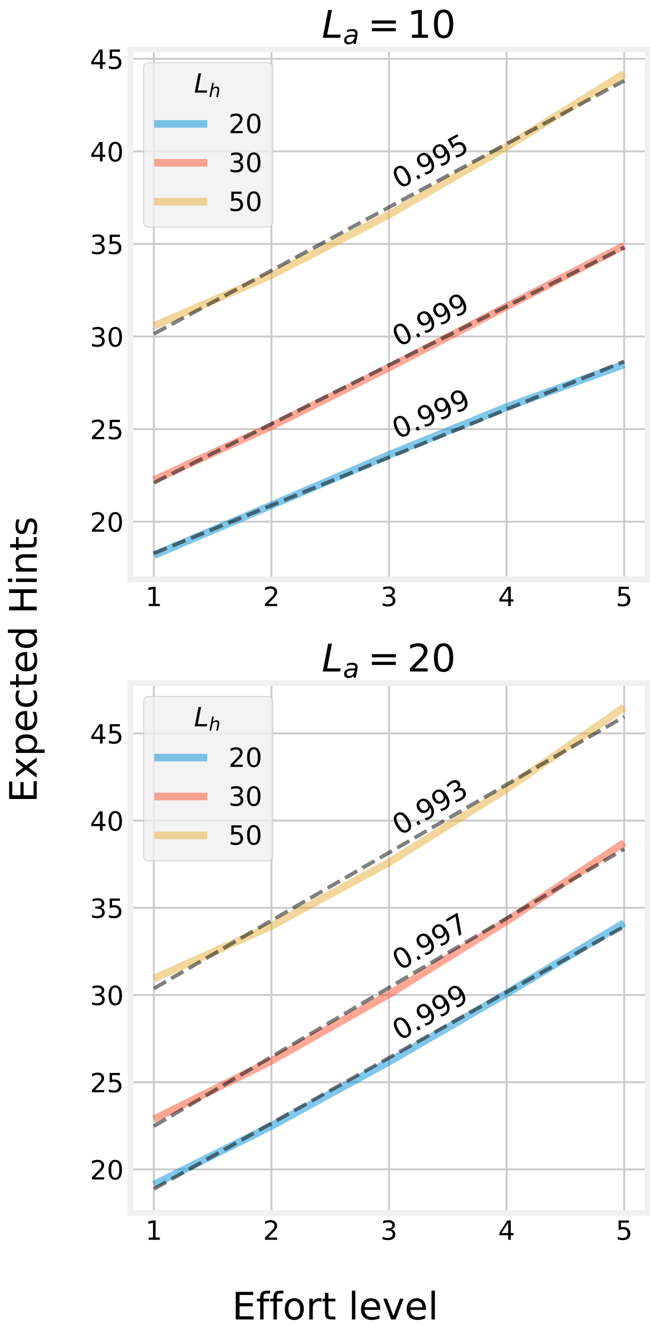

FIGURE 4.4: Linear fit across different limits. Text on the line shows the R^2 values for the best-fit line.

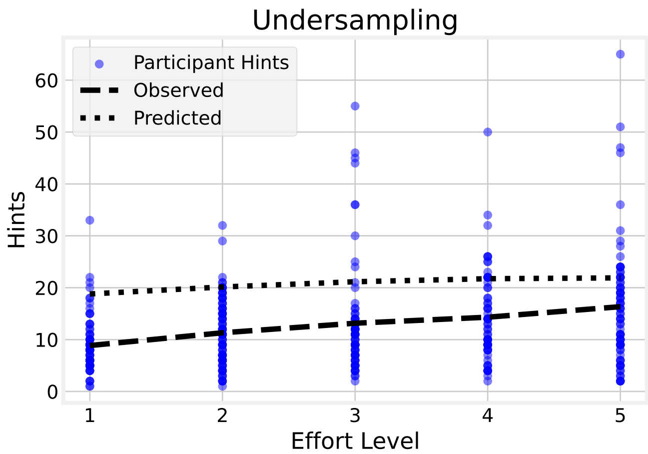

FIGURE 4.5: Observed hints curve compared with empirically estimated expected hints curve.

4.4 Results

4.4.1 An effort scale with linear response characteristics

Figure 4.4 shows the expected hints curve for effort levels for the upper limit of possible hints set at L_h at 20,30 and 50 and the upper limit of possible attempts (failures + 1) L_a set at at 10 and 20. For these limits, varying the standard deviation σ of the generative distribution of hints between 70 to 110 in increments of 10 provides an almost linear scale, indicated by R_2 fit values of around 99 when plotting SD vs expected effort. Thus, we obtain 5 values of the independent variable on a linear scale, with the assurance that they vary linearly with the expected effort at each level.

4.4.2 Respondents undersample hints

Figure 4.5 summarizes participants' sampling behaviour vis-a-vis the expected sampling behaviour estimated using L_h and L_a limits estimated from the data itself. For the participant data L_h=22 at the 90^{th} percentile across all participants, with a maximum of L_h=27 seen for one participant. We set $L_h = 27$ to err on the side of caution. $L_a = 4$ at the 90^{th} percentile, and is set as such to calculate the expected effort curve. As we see in Figure 4.5, participants under-sample across effort levels, in contrast to earlier observations of oversampling seen in [33] and [49]. We reflect upon this incongruity further below.

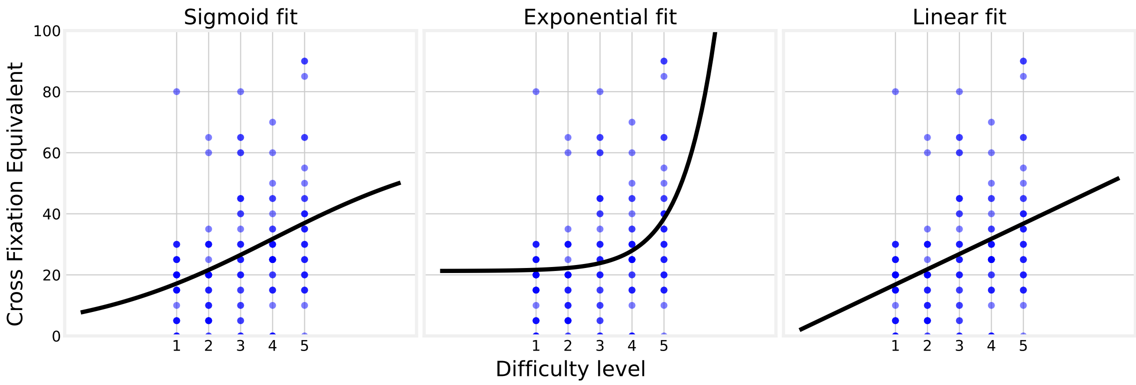

FIGURE 4.6: Three curve families fit to the data.

4.4.3 The time utility curve is approximately linear

The primary goal of this study was to estimate the time utility curve for mental effort. To achieve this, we fit a variety of mathematical functions to the effort level vs several hints of data collected from our participants. The following curve families were fit by minimizing the least square error; best-fit curves are shown in Figure **??**:

1. Linear: $y = ax + b$
2. Convex: An exponential curve: $y = a(1 + b)^{x-c} + d$
3. Concave: A sigmoidal curve: $y = \frac{a}{1+e^{-b(x-c)}} + d$

AICs for model fits are 2800,2800 and 2794 for the convex, concave and linear curves respectively, suggesting that the time utility curve of mental effort, at least on the time and effort scales measured in our experiment, is approximately linear, in accord with similar linear measurements of the money utility curve for mental effort previously documented in [47]. R^2 value for each curve was 0.12. Here we are not concerned with predictions but rather the shape of data, thus low R^2 values don't affect our results.

4.5 Discussion

Here we presented a method for measuring the time utility of mental effort, repurposing a recently proposed mental effort measurement to trade off effort for time instead of effort for money. An experimental evaluation of this method revealed three observations. One is that the method is econometrically reasonable in the sense that it produces linear response characteristics between task difficulty and expected effort level. Two, that respondents under-sample relative to the expected effort level. Three, the measured time utility curve is approximately linear.

Our observation of systematic under-sampling is in contrast with the results of previous Chapter[49] and [33] who have observed consistent oversampling in the same visual estimation task. It is interesting to note that [33] describe the purpose of their study as an effort to measure optimal information sampling behaviour under explicit costs, as a move away from the then en vogue paradigm of information sampling studies without explicit costs, which have reliably demonstrated under-sampling[32]. We find this trend again in our study in the absence of explicit costs, reinforcing the argument made in 2011don. It appears that individuals under-sample evidence when costs are imposed in terms of time, and over-sample when an explicit cost structure in terms of quasi-monetary units is presented, consistent with time being treated as a more fungible good than money, but also with many other theoretical possibilities. A deeper investigation of this difference constitutes an interesting avenue for future work.

Our experimental data are consistent with a linear shape for the time utility of mental effort on time scales and effort scales consistent with lab experiments (up to one minute of moderately effortful activity). That is, in this setting at least, every increasing unit of mental effort has the same utility for individuals if the utility is measured with time, suggesting that neither respondents' time nor effort budgets are constrained by the task we set them.

It is almost certain that this conclusion cannot generalize to longer time scales, wherein diminishing returns are certain to arise, or to other task contexts wherein the shape of the curve may be a function of expertise with the task and other factors[50]. Nonetheless, the time and effort scale of our experiment is entirely consistent with lab-scale experimentation and also seems likely to generalize to important practical use cases, such as the design of computer-human interfaces[51, 52], wherein measurements of time-effort trade-offs are an explicit need, and in pedagogical assessments that relate mental effort with difficulty[53, 54].

Finally, while mental effort may be operationalized in information processing terms to render it better suited for quantitative and computational reasoning[29], it is important to remember that this operationalization is limited, and ignores considerable phenomenological and practical detail about the nature of mental effort, e.g. the fact that the same amount of effort can 'feel' starkly different based on motivation factors[55]. It remains important, therefore, to caveat the discussion about the measurement of mental effort to the specific aspects of it that are amenable to an information-processing metaphor.

Chapter 5

Conclusion

5.1 Measurement of mental effort

The phenomenological effort of the mind is difficult to treat empirically. Like other phenomenological experiences, it does not lend itself to direct scientific observation except through the method of self-reporting which itself has a reputation of not being robust when pursuing accuracy. Despite the fact, that mental effort does manifest in behaviour through commonly experienced phenomena like mental exhaustion. Like its counterpart physical effort, mental effort displays properties of being limited in nature, which can either be because of capacity or quantity. In principle, humans can experience exhaustion because they can't take up too many mental operations together and also because mental operations exhaust limited mental resources. While the limitations of making additional mental operations concurrently have been clearly shown in empirical results like limited working memory[56], empirical treatment of limited resource quantity remains elusive. The econometric problem of mental effort research[7], which explains why empirical studies have been limited, is intuitive to grasp. To measure cognitive resources exhausted by mental operations we require a scale to compare resource utilisation by different types of mental operations within and across individuals.

We discuss how current methods either insufficiently inform about equality judgements between two mental operations[10] or fail to construct a meaningful scale of mental effort[1]. The limitation circles back to the inability to compare mental operations across and within individuals where scaling mental effort up through, for example, memory demands does not scale linearly. We construct a new method of mental effort measurement to remedy this, which utilizes the mental operations of paying visual attention for 1000ms whose demanded ability does not vary significantly both within and across individuals. Each 1000ms visual attention operation progresses the task by one unit and the number of such operations done stands in for the amount of mental effort applied. We utilize this method to measure preferences across differentiated effort levels and consistent with earlier findings, we show that a less effortful option is preferred against a more effortful one. In similarity to the original result, individuals prefer to give up larger rewards to incur less risk but a novel finding that our results suggest is that people are willing to expend more effort for less risk. Another novel finding is that with increasing and decreasing confidence, individuals change effort levels in opposite directions at least with the easier task, consistent with theoretical resource-rationality expectations[57]. This observation suggests that resource-rationality assumptions may have greater ontic significance than pure modelling devices[58].

5.2 Measuring utility of time versus effort

Following the micro-economic nature of mental effort research, we utilize the developed method of measurement to understand the relationship between mental effort and with its opportunity cost, which is measured in time units as applying effort precludes doing anything else in that duration. To construct a meaningful scale of mental effort that scales linearly with our independent variable, we use 1000ms of visual attention as the unit of mental effort in our task. This is done by ensuring no significant variability in any other mental operations besides 1000ms of visual attention is observed and modulating the expected number of these operations required as the independent variable. An experimental evaluation of this method revealed three essential results. First, the method is econometrically reasonable in the sense that it produces linear response characteristics between task difficulty and expected effort level, as such a relationship is mathematically amenable given the inter- and intra-subjective comparability of the unit of effort. Two, that individuals are cognitive misers, confirming intuition, as they consistently put in less effort compared to the nominal level. Three, the measured time utility curve is approximately linear.

Our experimental data are consistent with a linear shape for the time utility of mental effort on time scales and effort scales consistent with lab experiments (up to one minute of moderately effortful activity). That is, in this setting at least, every increasing unit of mental effort has the same utility for individuals if the utility is measured with time, suggesting that neither respondents' time nor effort budgets are constrained by the task we set them.

5.3 Method Limitations

While mental effort may be operationalized in information processing terms to render it better suited for quantitative and computational reasoning[29], it is important to remember that this operationalization is limited, and ignores considerable phenomenological and practical detail about the nature of mental effort, e.g. the fact that the same amount of effort can 'feel' starkly different based on motivation factors[55]. It remains important, therefore, to caveat the discussion about the measurement of mental effort to the specific aspects of it that are amenable to an information-processing metaphor.

Bibliography

[1] A. Westbrook and T. S. Braver, “Cognitive effort: A neuroeconomic approach,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 15, no. 2, pp. 395–415, 2015.

[2] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior (60th Anniversary Commemorative Edition).* Princeton university press, 2007.

[3] W. Edwards, “The theory of decision making.,” *Psychological bulletin*, vol. 51, no. 4, p. 380, 1954.

[4] D. Kahneman and A. Tversky, “Prospect theory: An analysis of decision under risk,” in *Handbook of the fundamentals of financial decision making: Part I*, pp. 99–127, World Scientific, 2013.

[5] S. S. Iyengar and M. R. Lepper, “When choice is demotivating: Can one desire too much of a good thing?,” *Journal of personality and social psychology*, vol. 79, no. 6, p. 995, 2000.

[6] R. Kurzban, A. Duckworth, J. W. Kable, and J. Myers, “An opportunity cost model of subjective effort and task performance,” *Behavioral and brain sciences*, vol. 36, no. 6, pp. 661–679, 2013.

[7] W. Kool and M. Botvinick, “Mental labour,” *Nature human behaviour*, vol. 2, no. 12, pp. 899–908, 2018.

[8] W. J. McGuire, “Attitudes and attitude change,” *The handbook of social psychology*, pp. 233–346, 1985.

[9] S. E. Taylor, “The interface of cognitive and social psychology,” *Cognition, social behavior, and the environment*, vol. 1, pp. 189–211, 1981.

[10] W. Kool, J. T. McGuire, Z. B. Rosen, and M. M. Botvinick, “Decision making and the avoidance of cognitive demand.,” *Journal of experimental psychology: general*, vol. 139, no. 4, p. 665, 2010.

[11] T. L. Dunn, D. J. Lutes, and E. F. Risko, "Metacognitive evaluation in the avoidance of demand.," *Journal of experimental psychology: human perception and performance*, vol. 42, no. 9, p. 1372, 2016.

[12] M. T. Gailliot, R. F. Baumeister, C. N. DeWall, J. K. Maner, E. A. Plant, D. M. Tice, L. E. Brewer, and B. J. Schmeichel, "Self-control relies on glucose as a limited energy source: willpower is more than a metaphor.," *Journal of personality and social psychology*, vol. 92, no. 2, p. 325, 2007.

[13] N. H. Mackworth, "The breakdown of vigilance during prolonged visual search," *Quarterly Journal of Experimental Psychology*, vol. 1, no. 1, pp. 6–21, 1948.

[14] J. S. Warm, G. Matthews, and V. S. Finomore Jr, "Vigilance, workload, and stress," in *Performance under stress*, pp. 131–158, CRC Press, 2018.

[15] M. S. Hagger, C. Wood, C. Stiff, and N. L. Chatzisarantis, "Ego depletion and the strength model of self-control: a meta-analysis.," *Psychological bulletin*, vol. 136, no. 4, p. 495, 2010.

[16] M. L. Evans, A. Pernet, J. Lomas, J. Jones, and S. A. Amiel, "Delay in onset of awareness of acute hypoglycemia and of restoration of cognitive performance during recovery.," *Diabetes Care*, vol. 23, no. 7, pp. 893–897, 2000.

[17] P. Y. Martin and D. Benton, "The influence of a glucose drink on a demanding working memory task," *Physiology & Behavior*, vol. 67, no. 1, pp. 69–74, 1999.

[18] D. O. Kennedy and A. B. Scholey, "Glucose administration, heart rate and cognitive performance: effects of increasing mental effort," *Psychopharmacology*, vol. 149, no. 1, pp. 63–71, 2000.

[19] S. T. Christie and P. Schrater, "Cognitive cost as dynamic allocation of energetic resources," *Frontiers in neuroscience*, vol. 9, p. 289, 2015.

[20] W. Schneider and R. M. Shiffrin, "Controlled and automatic human information processing: I. detection, search, and attention.," *Psychological review*, vol. 84, no. 1, p. 1, 1977.

[21] W. Schneider, S. T. Dumais, and R. M. Shiffrin, "Automatic/control processing and attention.," tech. rep., Illinois Univ Champaign Human Attention Research Lab, 1982.

[22] R. M. Krebs, C. N. Boehler, and M. G. Woldorff, "The influence of reward associations on conflict processing in the stroop task," *Cognition*, vol. 117, no. 3, pp. 341–347, 2010.

[23] M. Inzlicht, B. J. Schmeichel, and C. N. Macrae, "Why self-control seems (but may not be) limited," *Trends in cognitive sciences*, vol. 18, no. 3, pp. 127–133, 2014.

[24] M. A. Boksem, T. F. Meijman, and M. M. Lorist, "Mental fatigue, motivation and action monitoring," *Biological psychology*, vol. 72, no. 2, pp. 123–132, 2006.

[25] R. E. Petty, P. Briñol, C. Loersch, and M. J. McCaslin, "The need for cognition.," 2009.

[26] R. Eisenberger, "Learned industriousness.," *Psychological review*, vol. 99, no. 2, p. 248, 1992.

[27] M. Csikszentmihalyi and K. Rathunde, "The measurement of flow in everyday life: toward a theory of emergent motivation.," 1993.

[28] M. Székely and J. Michael, "The sense of effort: A cost-benefit theory of the phenomenology of mental effort," *Review of Philosophy and Psychology*, vol. 12, no. 4, pp. 889–904, 2021.

[29] A. Shenhav, S. Musslick, F. Lieder, W. Kool, T. L. Griffiths, J. D. Cohen, and M. M. Botvinick, "Toward a rational and mechanistic account of mental effort," *Annual review of neuroscience*, vol. 40, pp. 99–124, 2017.

[30] F. Lieder, D. Plunkett, J. B. Hamrick, S. J. Russell, N. Hay, and T. Griffiths, "Algorithm selection by rational metareasoning as a model of human strategy selection," *Advances in neural information processing systems*, vol. 27, 2014.

[31] R. Hertwig, G. Barron, E. U. Weber, and I. Erev, "Decisions from experience and the effect of rare events in risky choice," *Psychological science*, vol. 15, no. 8, pp. 534–539, 2004.

[32] R. Hertwig and T. J. Pleskac, "Decisions from experience: Why small samples?," *Cognition*, vol. 115, no. 2, pp. 225–237, 2010.

[33] M. Juni, T. Gureckis, and L. Maloney, "Don't stop 'til you get enough: adaptive information sampling in a visuomotor estimation task," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 33, 2011.

[34] I. Krajbich and A. Rangel, "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions," *Proceedings of the National Academy of Sciences*, vol. 108, no. 33, pp. 13852–13857, 2011.

[35] G. Tavares, P. Perona, and A. Rangel, "The attentional drift diffusion model of simple perceptual decision-making," *Frontiers in neuroscience*, vol. 11, p. 468, 2017.

[36] A. R. Hunt and A. Kingstone, "Covert and overt voluntary attention: linked or independent?," *Cognitive Brain Research*, vol. 18, no. 1, pp. 102–105, 2003.

[37] Z. M. Hafed and J. J. Clark, "Microsaccades as an overt measure of covert attention shifts," *Vision research*, vol. 42, no. 22, pp. 2533–2545, 2002.

[38] S. Dominick, "Principles of microeconomics," 2008.

[39] J. Marschak, "Rational behavior, uncertain prospects, and measurable utility," *Econometrica: Journal of the Econometric Society*, pp. 111–141, 1950.

[40] R. D. Luce and H. Raiffa, *Games and decisions: Introduction and critical survey.* Courier Corporation, 1989.

[41] J. Conlisk, "Why bounded rationality?," *Journal of economic literature*, vol. 34, no. 2, pp. 669–700, 1996.

[42] G. Mulder, "The concept and measurement of mental effort," *Energetics and human information processing*, pp. 175–198, 1986.

[43] S. H. Fairclough and K. Houston, "A metabolic measure of mental effort," *Biological psychology*, vol. 66, no. 2, pp. 177–190, 2004.

[44] D. Navon and D. Gopher, "On the economy of the human-processing system.," *Psychological review*, vol. 86, no. 3, p. 214, 1979.

[45] N. G. Mankiw *et al.*, "Introduction to microeconomics," 1997.

[46] B. Schwartz, A. Ward, J. Monterosso, S. Lyubomirsky, K. White, and D. R. Lehman, "Maximizing versus satisficing: happiness is a matter of choice.," *Journal of personality and social psychology*, vol. 83, no. 5, p. 1178, 2002.

[47] A. Westbrook, D. Kester, and T. S. Braver, "What is the subjective cost of cognitive effort? load, trait, and aging effects revealed by economic preference," *PloS one*, vol. 8, no. 7, p. e68210, 2013.

[48] R. M. Shiffrin and W. Schneider, "Controlled and automatic human information processing: Ii. perceptual learning, automatic attending and a general theory.," *Psychological review*, vol. 84, no. 2, p. 127, 1977.

[49] S. Mehrotra and N. Srivastava, "Selecting between visuomotor lotteries to measure mental effort in risky decisions," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 44, 2022.

[50] T. F. Meijman, "Mental fatigue and the efficiency of information processing in relation to work times," *International Journal of Industrial Ergonomics*, vol. 20, no. 1, pp. 31–38, 1997.

[51] S. Zugal, J. Pinggera, H. Reijers, M. Reichert, and B. Weber, "Making the case for measuring mental effort," in *Proceedings of the second edition of the international workshop on experiences and empirical studies in software modelling*, pp. 1–6, 2012.

[52] J. Baumeister, S. Y. Ssin, N. A. ElSayed, J. Dorrian, D. P. Webb, J. A. Walsh, T. M. Simon, A. Irlitti, R. T. Smith, M. Kohler, *et al.*, "Cognitive cost of using augmented reality displays," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 11, pp. 2378–2388, 2017.

[53] A. Srivastava, N. Srivastava, and S. Chandrasekharan, "Measure conceptmapping, not concept-maps: Procedural analysis elucidates stages in students' understanding of biology concepts," in *annual meeting of the American Educational Research Association. Retrieved July*, 2020.

[54] A. Srivastava, N. Srivastava, and S. Chandrasekharan, "Quality of concept maps is affected by map building strategies," *Biochemistry and Molecular Biology Education*, vol. 49, no. 1, pp. 129–139, 2021.

[55] D. Kleiber, R. Larson, and M. Csikszentmihalyi, "The experience of leisure in adolescence," *Journal of leisure research*, vol. 18, no. 3, pp. 169–176, 1986.

[56] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information.," *Psychological review*, vol. 63, no. 2, p. 81, 1956.

[57] F. Lieder and T. L. Griffiths, "Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources," *Behavioral and brain sciences*, vol. 43, 2020.

[58] D. Rahnev, "Resource-rational analysis vs. resource-rational humans," *The Behavioral and brain sciences*, vol. 43, p. e19, 2020.

[59] J. R. Busemeyer and J. T. Townsend, "Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment.," *Psychological review*, vol. 100, no. 3, p. 432, 1993.

[60] B. Schwartz and B. Schwartz, “The paradox of choice: Why more is less,” Ecco New York, 2004.

[61] Versus, “Apple iPhone 14 Pro Max vs Nokia 3310: What is the difference?.”

[62] M. W. Scerbo, “Stress, workload, and boredom in vigilance: A problem and an answer,” in *Stress, workload, and fatigue*, pp. 267–278, CRC Press, 2000.

[63] M. Rescorla, “The computational theory of mind,” 2015.

[64] J. F. Cavanagh and M. J. Frank, “Frontal theta as a mechanism for cognitive control,” *Trends in cognitive sciences*, vol. 18, no. 8, pp. 414–421, 2014.

[65] J. T. McGuire and M. M. Botvinick, “Prefrontal cortex, cognitive control, and the registration of decision costs,” *Proceedings of the national academy of sciences*, vol. 107, no. 17, pp. 7922–7926, 2010.

[66] B. Bruya and Y.-Y. Tang, “Is attention really effort? revisiting daniel kahneman’s influential 1973 book attention and effort,” *Frontiers in psychology*, vol. 9, p. 1133, 2018.

[67] N. Elkins-Brown, B. Saunders, and M. Inzlicht, “Error-related electromyographic activity over the corrugator supercilii is associated with neural performance monitoring,” *Psychophysiology*, vol. 53, no. 2, pp. 159–170, 2016.

[68] S. Frederick, G. Loewenstein, and T. O’donoghue, “Time discounting and time preference: A critical review,” *Journal of economic literature*, vol. 40, no. 2, pp. 351–401, 2002.

[69] G. Loewenstein, T. O’Donoghue, and M. Rabin, “Projection bias in predicting future utility,” *the Quarterly Journal of economics*, vol. 118, no. 4, pp. 1209–1248, 2003.

[70] C. Seidl, “The st. petersburg paradox at 300,” *Journal of Risk and Uncertainty*, vol. 46, pp. 247–264, 2013.