Title
Essays on the neuroeconomics of motivation

Author
Angelides, Nicholas H

Publication Date
2020

Essays on the neuroeconomics of motivation

by

Nicholas Angelides

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Psychology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor Ming Hsu, Co-chair
Professor Richard Ivry, Co-chair
Professor Sheri Johnson
Professor Dana Carney

Fall 2020

Abstract

Essays on the neuroeconomics of motivation

by

Nicholas Angelides

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Ming Hsu, Chair

The ability to initiate goal-directed behaviors related to every day outcomes is a crucial component of motivation. Across clinical and subclinical domains, variability in motivated decision-making accounts for failures to both begin and complete actions towards improved health. When faced with decisions that involve expending our own energetic resources in order to attain rewards, whether for ourselves or others, individuals vary broadly in the behaviors they take, or fail to take, to reach said goals. In order to classify, and potentially provide a mechanistic account of, differences in these cost-benefit processes, an objective investigation into the computations underlying motivated behavior is tantamount. This is particularly relevant to the clinical domain, wherein the broad goal of computational psychiatry is a methodologically quantitative approach to taxonomic classification. The following essays offer an overview of three studies with the aim of providing formal accounts of integrating information from individual differences (chapters 1 and 2) and stereotype content (chapter 3) into mathematical models of motivated behavior. I will begin by discussing a computational approach towards predictive modeling of effort-based decision-making utilizing dimensions of clinical and subclinical approach and avoidant traits. In the first chapter, I will focus on modeling behavior in a community-recruited sample of adults by integrating a number of demographic and individual difference measures into behavioral economic models of effort discounting. In the second, I will focus my analyses on a separate but demographically similar cohort of adults who completed the same experimental paradigm during functional neuroimaging. Here, I will investigate changes in neural activity (inferred by blood-oxygen level dependent (BOLD) signal) responding to different levels of effort-reward tradeoffs. I will utilize the same behavioral modeling processes as the first chapter in order to extract individual difference measures by estimating free parameters specific to individual subjects, and relating them to changes in the neural responses to cost-benefit evaluations. Finally, I will discuss the manner in which evaluations of effort expenditure for other-rewarding outcomes are distinct from those with self-rewarding outcomes, and how social information – or stereotypes – about the recipient of said rewards robustly contributes to effort-related choice processes.

Dedicated to my father, who showed me the meaning of *ἀρετή*,
and the California Lightweight Rowing team, who taught me its practice.

TABLE OF CONTENTS

Acknowledgements

My doctoral journey is crystalized, here in the writing of this dissertation, at a peculiar moment in time. As I write, we find ourselves doing something akin to our best to navigate the precipitous height of a global pandemic, to deconstruct the racial and economic inequity that rests solidly on our shoulders, and to grapple with the ways in which scientific discovery and nuance have never been more important, or more disbelieved. This all, of course, constitutes an exercise in forcing perspective. The vicissitudes of graduate school show us that when Gandhi says, "satisfaction lies in the effort, not in the attainment," he only captures part of the picture (sorry Gandhi). By stepping away and taking stock, and as this thesis will hopefully clarify, we see that without attainment, effort is but a sunk cost; without effort, there is no attainment. So allow me briefly to step away and take stock.

This thesis would not have been possible without the support, love, and understanding of my family and my partner. Thank you all for having the grace to know when to offer advice or just to listen, and for carrying the torch of confidence in me when I could not. I would especially like to thank my partner, Christian, whose unending patience and love made 2500 miles over 5 years feel easy (the fact of which made me appreciate that I do not study temporal discounting). To all seven of my siblings, each of you gave me bravery when I needed it most. To all three of my parents, you are my rock.

Thank you to my friends and colleagues who gave me a home and family when I was so far from my own. I would not have gotten through this experience without being able to share it with some of the best people I've had the luck of meeting. Vinitha and Joe, our stars aligned early and together; you both kept me motivated and sane. To Enitan and Christina, because of the two of you, I felt like I had a home at the end of every long day. Sorry if I woke you up at 4am too many times.

To the Cal Lights: lab work may have fed my mind (and my stomach), but you fed my soul. Always stay ready to move boats.

To Dr. Nick Altemose, twinny, a most patient and brilliant friend and teacher, I could not have gotten through this without you, in too many ways to spell out here. Thank you to Dr. Anna Jenkins, whose presence in lab was at once comforting and inspiring, and who often abetted my faith in myself as a scientist and humanist.

Thank you to my committee members Dr. Rich Ivry, Dr. Dana Carney, and Dr. Sheri Johnson for your feedback, guidance, and insights. I'd particularly like to thank Sheri and the CalM team who graciously put their trust in me and in so doing allowed me a broader view not only of our field, but – crucially – of its impact as well.

Finally, I would like to extend my deepest gratitude to my advisor, Dr. Ming Hsu, whose mentorship, guidance, and faith in me ensured my success. You were always willing to listen to and engage with the ways of thinking I was used to, while also pushing me to think in ways I didn't know I could before. Thank you.

“The real price of every thing, what every thing really costs to the [person] who wants to acquire it, is the toil and trouble of acquiring it.”

-Adam Smith

“Work smart, not hard.”

– a Snapple cap I got one time

Chapter 1

A computational framework for integrating stable trait measures and effort-based decision-making

1.1 Introduction.

Behavioral apathy, or lack of motivation, has a significant societal and personal cost, affecting education, employment performance, and the ability to achieve health-related outcomes. From the bottom up, motivation is made up of a complex set of processes, operating in several domains: motivation to take a course of action, to engage in cognitive effort, or to engage in emotional interaction. The phenomenon underlying motivated behavior is also influenced by various interacting but separate developmental, cultural, and environmental factors. Given these discrete components broadly contributing to motivation, a challenge in studying motivation across individuals is the significant interindividual variability, ranging from healthy individuals who are highly motivated, to clinical populations who suffer from disorders of diminished motivation (Chong et al., 2016[1]). These individual differences (IDs) are widespread and often distinguished through self-report or symptomology based on observation (Treadway et al., 2012[2]; Holt et al., 2003[3]). However, a growing body of literature continues to evidence a set of cognitive and neural mechanisms underlying both individual difference measures and their relationship to motivated behavior (Salamone & Correa, 2012[4]; Chong et al., 2017[5]). Many current examinations of individual differences at both the behavioral and neural levels rely on the use of multi-level linear models, which may offer some insight into emergent behavioral phenomena, but do not bear on possible computational or system-level mechanisms (Patzelt et al., 2017[6]). There exists a need for formal, mechanistic definitions of these individual differences in order to describe how and why the composition of processes leads to specific observable outcomes or behaviors. In an attempt to address this gap, the following chapter will describe a method for 1) integrating subclinical individual difference measures into computational models to determine whether and how they contribute to the behavioral processes underlying motivated decision-making, and 2) estimate free parameters from behavioral data that may be interpreted in terms of cognitive processes. The aim of this process is to derive a collection of parameters from behavior to allow quantitatively precise predictions about future behavior and neural activity, thus helping to identify the contribution of specific individual difference measures beyond, or in conjunction with, self-report.

1.1.2 Effort-based decision-making as proxy for motivation

An important preliminary assumption to clarify lies in the operationalization of motivation using effort-based decision-making tasks. Experimental investigations into the neurobiological and computational mechanisms of motivation have historically focused on observing an organism's goal-directed behavior towards obtaining rewarding outcomes in order to infer motivational states (see, e.g., Cardinal et al, 2006[7]; Salamone 2009[8]). One critical component of that behavior can be defined as the amount of metabolic energy – or effort – an organism is prepared to deploy for a given reward. Like delay, risk, or uncertainty, effort is typically defined as a cost, and has the effect of devaluing associated rewards (Hull, 1943[9]; Salamone & Correa, 2012[4]). Therefore, while physical effort is by no means the sole or most meaningful component of motivation, for the purposes of this thesis, motivation and motivated decision-making will be indexed by choices between effort- and no-effort options in the dual-choice tasks described below.

Motivated behavior in non-human model organisms has been measured with well-validated tasks examining tradeoffs between physical effort (e.g., lever presses, as in Shafiei et al., 2012[10]) and consummatory rewards (e.g., sugary cereal). These studies have allowed us to characterize motivation as that process which facilitates overcoming the cost of an effortful action to achieve the desired outcome. As such, physical effort is definitionally an ecologically valid measure of motivation across species. One particular challenge has been developing techniques that adequately capture the complexities of motivated behavior in humans, when the utility of certain rewarding outcomes can depend upon a variety of individual differences and response costs. More recently, behavioral and fMRI studies have translated such protocols (e.g., dual-alternative and progressive-ratio tasks) to paradigms amenable to investigations in human subjects (Chong et al., 2016[1]). Operationalizing motivation beyond the use of self-report and personality measures is crucial and can be done using various physically or cognitively effortful tasks, including grip strength testing, button-pressing, or difficult puzzles (Hartmann et al., 2013[11]; Treadway et al., 2012[2]; Westbrook & Braver, 2015[12]). How these effortful cost-benefit valuations may shift as a result of variability in stable, trait-like differences between individuals, however, remains unknown.

1.1.3 Effort allocation within and across individuals

The objective characterization of the cost-benefit processes that underlie motivated behavior are particularly important within the clinical domain. Recent work suggests that motivation – the process underlying the way in which organisms make such cost-benefit tradeoffs about metabolic costs – in both health and disease is likely to lie on a continuum (Chong et al., 2016[1]). Theoretical frameworks that offer a quantitative characterization of dimensions regarding motivated behaviors (Ang et al., 2017[13]; Radakovic et al., 2020[14]) may help to facilitate the generation of mechanistic predictions regarding the contributions of individual differences to evaluations of effort and reward. Broadly speaking, this is a major goal of computational psychiatry.

The Behavioral Inhibition System/Behavioral Activation System (BIS/BAS) questionnaire (Carver & White, 1994[15], following Gray, 1987[16]) is a well-validated inventory shown to relate to inter-individual differences in appetitive motivation and approach behavior. The BIS/BAS questionnaire assesses three personality dimensions related to behavioral approach and reward sensitivity (BAS), and one related to behavioral inhibition and anxiety (BIS). The three BAS scales include items probing the persistent pursuit of desired goals (drive scale), the desire for novel rewards and willingness to approach potentially rewarding events (fun-seeking scale), and positive responses to the occurrence or anticipation of reward (reward responsiveness scale). This set of scales is useful for the purposes of these chapters for a number of reasons. First, while not used for clinical diagnoses, links between the BIS/BAS scales and psychopathology (e.g., bipolar mania) are instantiated at both the behavioral and neural levels (Johnson et al., 2003[17]). Second, functional neuroimaging studies have implicated the mesocortical circuitry involved in reward sensitivity reflected by the BIS/BAS subscales (Angelides et al., 2017[18]; Kahnt et al., 2012[19]; Scheres & Sanfey, 2006[20]). Experimental paradigms that elicit behavior using rewarding stimuli reliably activate regions in the subcortical striatum and orbitofrontal/ventromedial prefrontal cortex, both of which are the focus of studies investigating reward coding and motivation (Laricchiuta et al., 2014[21]; Jarbo & Verstynen, 2015[22]).

These traditional approaches to detecting, measuring, and predicting individual differences have been effective in reducing the dimensions of personality to latent constructs, such as the BIS/BAS scales. Such dimensions are typically stable across the lifetime, but are largely agnostic as to the link between cognitive mechanisms subserving behavior, and the behavior itself. For that reason, the following chapter will focus largely on the estimation of individual differences from behavioral data using computational methods, and the relationship between those free parameters and BIS/BAS scores.

Effort discounting models allow us to estimate parameters defining the degree to which individuals discount rewards by physical effort. The data are entered into a model, which produces a set of parameters specific to the behavior of interest (here, effort-based decision-making). Ultimately, in addition to these self-report measures of approach and avoidance, we are able to estimate behavioral indices specific to motivation. By integrating ID measures related to aversive and appetitive motivation into effort discounting models, we may be better able to characterize the underlying relationship between the latent personality constructs and cognitive processes related to motivated behavior. In this chapter, I aim to 1) describe behavioral variability in effort-discounting as a function of broadly used individual difference measures with established links to pathology and verifiable, biological substrates and 2) introduce a mathematically formal method of integrating such measures, in a manner extendable beyond BIS/BAS to true diagnoses, that may help better predict aberrant behavior using free parameter estimation.

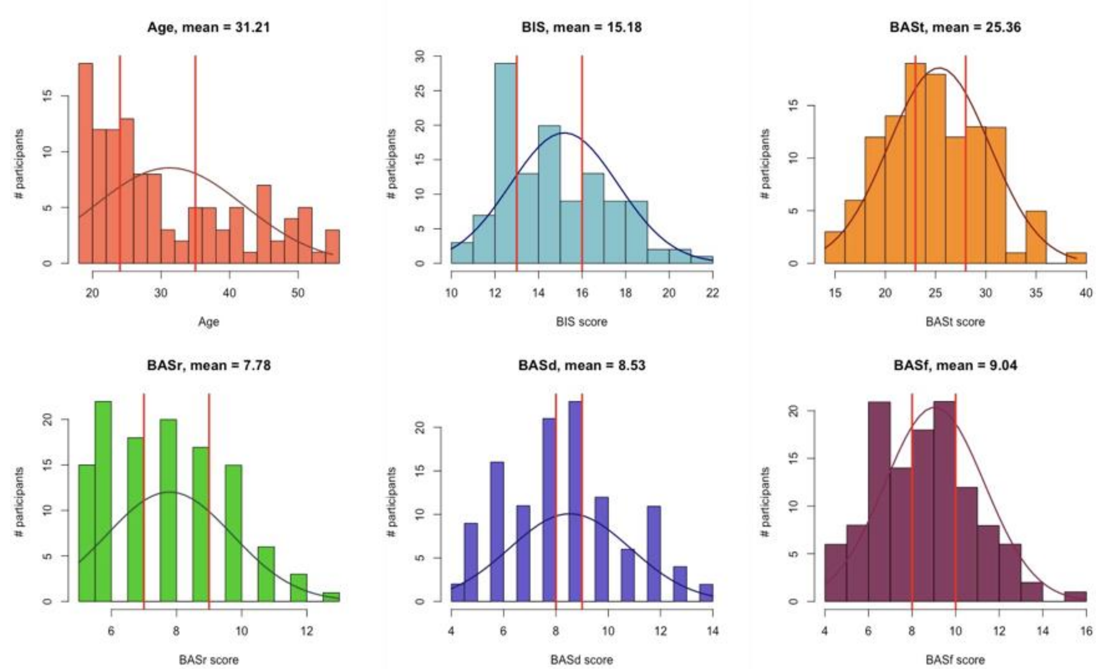

Figure 1 Distribution of scales used in ensuing analyses for Miami (behavioral-only) cohort. Red lines separate terciles.

1.2 Methods & Materials

1.2.1 Participants

Participants were 116 adults (male = 29.1%, mean age = 31.2, range = 18-55, *SD* ± 10.9) recruited from the University of Miami student and community populations. Over the course of 3 visits, participants completed a number of behavioral tasks. Self-reports of various stable (personality), clinical, and sub-clinical traits were obtained from participants during one of three visits. Demographic information can be found in Table 1.

Table 1. Descriptive statistics of demographic and individual difference measures among all participants.

	Mean	*Median*	*Range*	*SD*
N = 116				
Age	31.21	27	18, 55	10.89
Gender (%Male)	29.1%			
BIS	15.18	15	10, 22	2.47
BAS-total	25.36	25	14, 39	5.08
BAS-rr	7.78	8	5, 13	1 .94
BAS-drive	8.54	8	4, 14	2.31
BAS-fs	9.04	9	4, 16	2.29

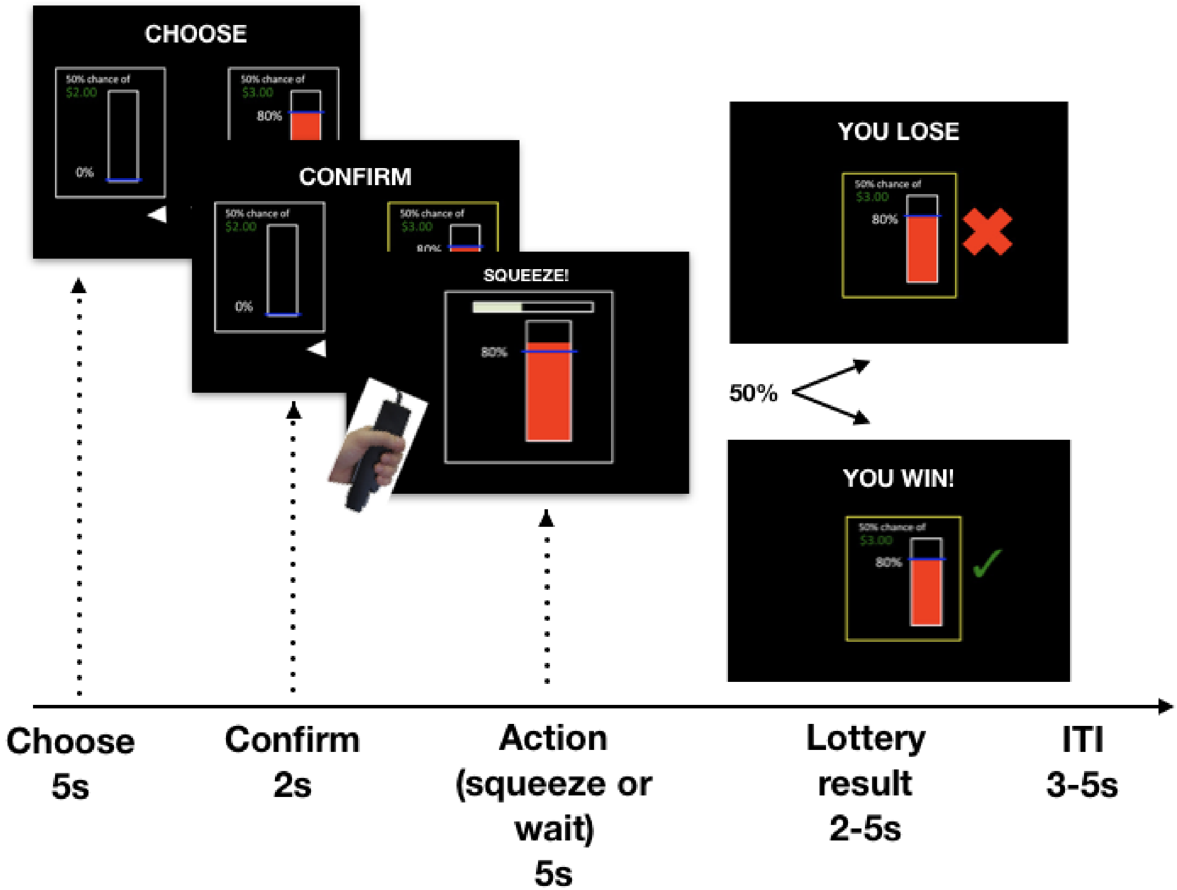

Figure 2 **Effort-based decision-making (eDM) task.** Subjects are presented on each trial with two possible lotteries: an effort option (EF), and a no-effort (NF) option. Subjects choose between the two lotteries, wherein choosing the EF option requires the subject to expend effort equivalent to a determined proportion of their individualized maximal voluntary contraction (%MVC) for up to 3s, while choosing the NF option requires subjects simply to wait for an equivalent amount of time. Lotteries have variable and unequal possible payouts. Subjects select a lottery: a NF lottery choice requires them to wait for 5s, while an EF lottery requires them to squeeze. Subjects receive visual feedback via a bar that moves vertically in real-time according to their current %MVC. Each arm of the bandit will have a stable 50/50 win/loss probability across trials. The subject is given up to 5s to choose a lottery, and must complete 3 of 5 total seconds at target %MVC in order to play the lottery. If a subject fails to complete the required %MVC or to hold that effort level for the required amount of time (3s), the computer registers a loss and moves on to the next trial. NF lottery requires no effort, while EF lottery target %MVCs vary by trial between 10-90 %MVC. Subjects complete 90 trials, and all experimental sessions reference subjects' specific MVC (registered by calibration early in each session) in order to normalize across sessions and individuals.

1.2.2 Procedure

Estimation of Maximal Voluntary Contraction.

After informed consent and before task instructions were given, each participant's maximum voluntary contraction (MVC) was established in order to normalize effort costs across participants. Participants were asked to hold the dynamometer and squeeze as hard as they could with their dominant hand for 3 seconds. They were shown real-time feedback in the form of a moving dotted line that changed slope as a function of increased/decreased squeezing intensity. They were shown a number that corresponded to their MVC, and told that they would repeat this processes. To motivate them to find their true MVC, they were told that they would try again, and if they could bring move the output line above a marker indicating their previously recorded MVC, they would win an additional $1.00 at the end of the experiment. This process was repeated again a third time, giving subjects a total of three chances to find their true MVC. The single maximal force reached between all three attempts was used as the MVC for that participant.

1.2.3 Paradigm

The effort-based decision-making (eDM) task in this study has successfully been shown to elicit broad variability in individualized effort discounting curves. This dual-choice lottery paradigm orthogonalizes rewards and effort costs by parameterizing the required effort costs – operationalized as the amount of force to be delivered to a hand-held dynamometer as a function of each individual's maximum voluntary contraction (MVC) – and reward levels. Subjects complete 90 trials, and all experimental sessions will reference subjects' specific MVC (registered by calibration early in each session) in order to normalize across sessions and individuals. A schematic of the task can be seen in Figure 2.

A major consideration within the broader goals of the study for which these data were collected was to examine the effects of rewarding or punishing outcomes on future behavior in specific clinical and subclinical populations. A common set of behavioral phenomena resulting from externalizing and internalizing symptoms includes changes to future behavior as a function of variation in affective states. Rewarding outcomes, especially when unexpected, can trigger increases in approach behavior, while punishing or negatively outcomes may have the opposite affect, leading to increases in avoidant behavior.

With this in mind, a proportion of trials were coded as deterministic at 100% in order to examine whether highly rewarding outcomes, or repetitive rewarding outcomes – i.e., "streaks," – led to increases in the general allocation of effort. A description of how these regressors were coded into our analyses is described below.

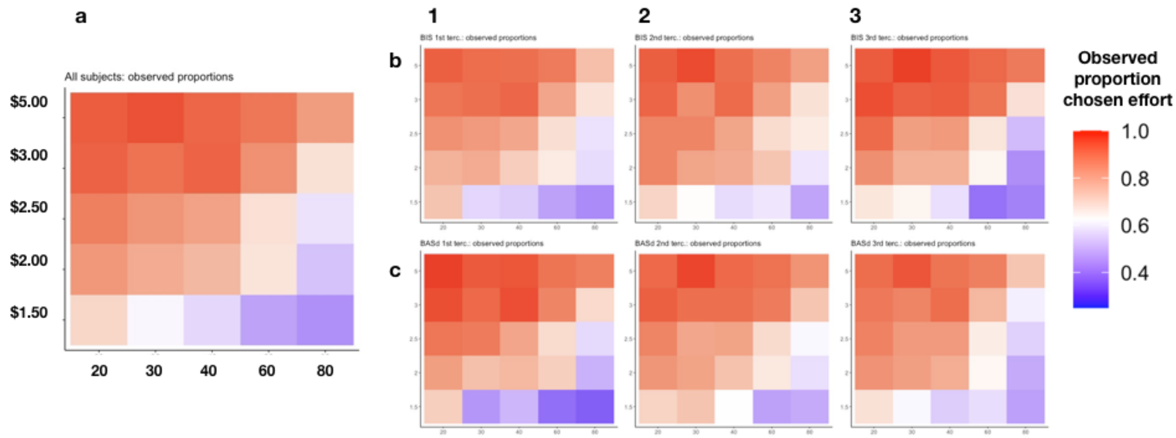

Figure 3: **Proportion of observed choices for high-effort lottery.** a. Proportions of all subjects choices, averaged over all reward-cost combination types. b. Proportions of choices according to BIS splits by first tercile (1), second tercile (2), and third tercile (3). c. The same splits, by BAS-drive. These data illustrate the variety of choices across all participants within two selected subscales.

However, we took extra time to inform participants that reward probabilities would be equivalent (50%) between all decision types, on all trials, in order to preclude the assumption that higher demand trials would lead to more highly rewarding outcomes. Effort discounting itself relies on the intuition that rewards are valued more highly in reference to what a person believes they deserve given the effort they have expended. Therefore, we explicitly informed the participants of the equivalent reward probabilities between block types, seeking to neutralize the prediction factor to get the participants to focus on tradeoffs given current information.

1.2.4 Hardware & Software

Stimuli and feedback generated using PyGame and in-house code were presented to participants using a PC. In order to register responses, participants selected one of two possible options using a keyboard, and a handheld dynamometer (MP150, BIOPAC Systems, USA) to pay the effort costs required to play the lottery. Dynamometer force recordings were measured at 10 Hz.

1.3 Models & Model Selection.

1.3.1 Model-free: GLMs and LMERs

Behavioral task choices were first analyzed with a generalized logistic mixed effects model using the lme4 package (Bates et al., 2015[23]) in R (version 3.6.1; R Core Team, 2019). The dependent variable was chosen effort (1=effortful lottery choice, 0=no effort lottery choice) (see Fig. 3 for examples of individual differences in the % of effortful lottery choices). The null model included a subject-specific intercept and the following regressors of interest: reward magnitude (R(effort) – R(no effort), and effort cost as %MVC.

1.3.2 Model-free parameter selection

To account for an average likelihood above chance (mean = 72.2% proportion effortful choices) of choosing effort, we initially ran linear models per participant to estimate a subject-specific response-bias parameter, *beta*. This parameter can be entered into second level models to account for subject-wise constant effects on the intercept. It should be noted that the increased likelihood (>50%) of choosing effort is not a confound. Inzlicht and colleagues' (2017[24]) review of what they call "the effort paradox" shows that studies of both human and non-human organisms offer evidence of a bias towards effort expenditure, especially after energy has already been spent towards a goal. A "learned industriousness" explanation proposes that, in the real world, more effort often leads to more reward, and therefore organisms may show an increased willingness to engage with effortful tasks out of a desire to maximize rewards (Eisenberger et al., 1989[25]). Another theory suggests that the aversive nature of exerting effort may make subsequent rewards seem, by contrast, more valuable (e.g., sunk costs, per Arkes & Blumer, 1985[26]). While our experimental stimuli were not optimized to disentangle the source of this bias, our analyses make a point of accounting for it.

Monetary value, reward magnitude, or utility

Given the possibility that effort discounting operates in units subjective value and not monetary value, we first investigated the contribution of different types of reward to participant decisions. While regression models are not designed to return subjective value, we attempted to determine the best characterization of the value of rewards by comparing three models using different regressors representing reward levels: reward value R (raw value of the effort lottery reward), reward magnitude M (value of the effort lottery minus value of the no-effort lottery), or utility U (reward magnitude to the 0.8 power ($M^{0.8}$), which provides an average utility estimate for humans, as suggested by Abdellaoui and colleagues, 2007[27]). All three models, differentiated by reward-related regressors, were significant above the null model, as per analysis of variance (ANOVA) comparisons: value R (model: chi^2 = 1429.7, DF = 1, BIC = 9198.4, $P < 0.001$; regressor: F_{10420} = 989.69, $P < 0.001$); magnitude M (model chi^2 = 1762.5, DF = 1, BIC = 8865.6, $P < 0.001$; regressor: F_{10420} = 1197.0714, $P < 0.001$); utility U (model: chi^2 = 1762.5, DF = 1, BIC = 8849.7, $P < 0.001$; regressor: F_{10420} = 1281.08, P<0.001). Following model Bayesian Information Criterion (BIC), we proceeded in all logistic regression analyses using a utility regressor, given as:

$$Utility_t = (Reward_{effort_t} - Reward_{no\ effort_t})^{0.8}$$

Effort cost, or utility cost

In order to determine how the cost of the lottery presented to participants was computed, we compared two logistic regressions using different regressors denoting effort: the raw %MVC (as an integer, e.g., 80% MVC = 80), or utility cost C (given as a scaled %MVC^2, as shown in parabolic models of effort discounting). Both models, differentiated by regressors representing effort, were significant above the null model, as per analysis of variance (ANOVA) comparisons: %MVC (model: $c^2 = 453.2$, DF = 1, BIC = 8849.7, $P < 0.001$; regressor: $F_{10420} = 1281.08$, $P < 0.001$); magnitude M (model: $c^2 = 447.82$, DF = 1, BIC = 8855.1, $P < 0.001$; regressor: $F_{10420} = 1269.58$, $P < 0.001$). Given comparisons of model BICs, we proceeded with %MVC as our effort regressor in all following models. For all mixed-effects model analyses, the optimizer "bobyqa" was used, with a maximum number of 2×10^5 iterations. P-values were determined using Likelihood Ratio Tests.

Our models are specifically built to integrate trait measures by estimating weight-parameters that can differentially affect predictive power. To make this intuitive, we can imagine a person who tends to choose effort more often than their peers. They happen to have a higher BAS-drive score than others, but this does not give us any insight into the mechanisms by which this person makes these specific cost-benefit tradeoffs. We do not know whether their preference is due to a response bias (at the intercept) towards expending behavior, or if they do not discount as steeply (i.e., they don't find effort as costly), or if they simply find rewards more rewarding than most people. Behaviorally, these effects are indistinguishable, and we do not know how or whether the "system" underlying their motivation (here, BAS-drive) contributes to their behavior. However, by estimating parameters that modulate the contributions of a self-report measure to specific behaviors, we gain not only predictive power over simple linear regression, but also insight into possible mechanisms by which these cognitive computations are made.

1.3.3 Behavioral choice models and model selection.

In order to first determine subject-specific response biases β_0, logistic regressions were run on each subject:

$$P(effort) \sim \beta_0 + U_t + MVC_t$$

The null mixed-effects logistic regression, which assumes no contribution of individual differences, was given as:

$$P(effort) \sim \beta_j + U_t + MVC_t + trial_t + U_{t-1} + effort_{t-1} + win_{t-1:3} + (1|SID)$$

Where β_j denotes the intercept weight β_0 of subject j to account for response bias (effort v no-effort bias), U_t denotes the utility (reward value of the effortful lottery minus the reward value of the no-effort lottery, to the 0.8 power) on trial t, MVC_t denotes the effort cost (%MVC) required on trial t, $trial_t$ denotes the number of trials completed so far on trial t, and previous trial effects ($U_{t-1}, effort_{t-1}, win_{t-1:3,n}$) denote the utility of the previous trial's outcome, whether or not the subject expended effort on the previous trial, and a regressor to represent a win streak of 3 or more in a row, respectively. Subject ID number was entered as a random effect.

The enriched regressions were given as:

$$P(effort) \sim id_j \times \beta_j + id_j \times U_t + id_j \times MVC_t + id_j \times trial_t + id_j \times U_{t-1} + id_j \times effort_{t-1} + id_j \times win_{t-1:3} + (1|SID)$$

Where the null model parameters hold, and the value of a specific ID measure of subject j is indexed by id_j. Each iteration of the model defined id_j as a distinct ID to evaluate the weight of that specific measure on various task-specific variables. In this analysis, ID measures corresponded to one of: demographic variables age, race, gender, and the five BISBAS measures inhibition (BIS), activation system total (BAS-t), reward-responsiveness (BAS-r), drive (BAS-d), and fun-seeking (BAS-f). Results from these analyses can be seen in Table 2.

Logistic Regression Results

Across all participants, decisions were guided by main effects of reward magnitudes and costs of possible choices (effort task: Utility, t = 208, %MVC, t = 143; both p<0.001). Within enriched regressions, those including demographic variables Race (model chi^2 = 70.9, DF = 18, BIC = 8951, $P < 0.001$), age (model chi^2 = 29.1, DF = 9, BIC = 8910, $P < 0.001$), and gender (model chi^2 = 23.4, DF = 9, BIC = 8792, $P < 0.001$) were significant. Among BIS/BAS scales, BIS (model chi^2 = 45.1, DF = 9, BIC = 8771, $P < 0.001$), BAS-drive (model chi^2 = 34.5, DF = 9, BIC = 8781, $P < 0.001$), and BAS-total (model chi^2 = 23.7, DF = 9, BIC = 8915, $P < 0.001$) contributed significantly above the null to the probability of choosing effort. While none of these demographic or BIS/BAS measures had a significant effect on the intercept, there were many significant interactions that give some indication of directionality. Full results can be found in Supplementary Tables 1 & 2[Supplementary] Table 1.

Table 2: **Logistic regression ANOVA results.** Results from two-way ANOVAs comparing each enriched model to the nested (null) model. Models including parameters for Race, Age, Gender, BAS-drive, and BIS are significant above the null, but factorizing race diminishes this result.

GLM	*npar*	*AIC*	*BIC*	*logLik*	*deviance*	*Chisq*	*Df*	*Pr(>Chisq)*	
NULL	8	8798	8856	-4391	8782	-	-	-	
Race	26	8763	8951	-4355	8711	70.9	18	3.2E-08	***
BIS	17	8771	8894	-4368	8737	45.1	9	8.80E-07	***
BASd	17	8781	8904	-4374	8747	34.5	9	7.40E-05	***
Age	17	8787	8910	-4376	8753	29.1	9	0.00062	***
BASt	17	8792	8915	-4379	8758	23.7	9	0.0047	**
Gender	17	8792	8916	-4379	8758	23.4	9	0.0054	**
BASr	17	8802	8925	-4384	8768	13.5	9	0.14	

BASf	17	8802	8926	-4384	8768	13.4	9	0.15
Hispanic	17	8806	8930	-4386	8772	9.29	9	0.41

1.3.5 Discounting models

While logistic regression can quantify the contributions of specific task variables to subject choices, linear models do not offer a plausible account of the cognitive computations underlying this type of behavior. Formal characterizations of utility discounting, borrowed from behavioral economics, may provide insight into a putative model of how the brain computes tradeoffs between effort and reward. Here, we examined 1) whether estimated free parameters in discounting models provide an improved explanation of discounting behavior, and 2) whether the contributions of individual differences to such behavior better explain variance in cost-benefit tradeoffs regarding effort.

Parabolic, free power parabolic, sigmoidal.

Many existing studies of effort-related choices utilize parabolic models (Hartmann et al., 2013[11]) to describe the shape of effort discounting. However, a systematic investigation from Klein-Flügge and colleagues (2015[28]) shows that the discounting of rewards by physical effort (instead of, for example, delay, which follows a hyperbolic shape), better follows a sigmoidal curve. In order to verify that our subjects were following a sigmoidal discounting curve, we first ran omnibus models of each shape fitted to all 116 subjects' data. Bayesian Information Criterion, which penalizes additional parameters, confirm that the sigmoidal model best characterizes the effort discounting behavior of our participants (Table 3).

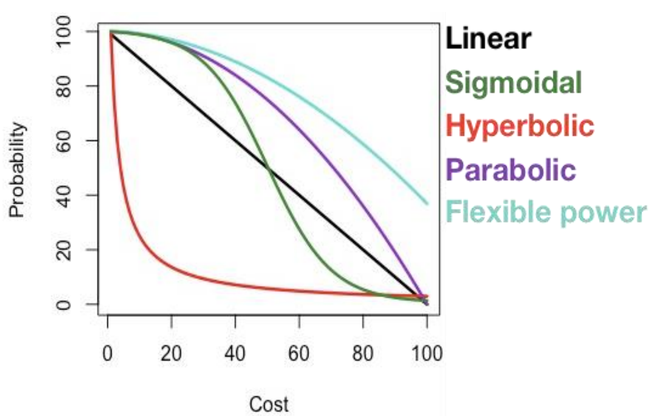

Figure 4 **Illustrations of the shapes given by possible discounting models.** The linear model implies a constant integration of effort independent of reward amount. The hyperbolic model describes how additional costs have large impacts on discounting value when added to small costs but a small impact when added to high costs (e.g., delay discounting). The parabolic model describes a shape wherein rewards are devalued by an increasing amount with growing effort/cost levels. The flexible power function matches the number of parameters of the sigmoidal model but with features more similar to the parabolic model. The sigmoidal model (winning model per table 3 below), is designed to accept initially concave shape to allow small costs to have little effect on value, whereas high effort levels that are unattainable flatten the curve.

Table 3 **Model fits for different possible shapes of discounting**. All models fit in omnibus across all subjects.

k	*beta*	*llik*	*BIC*	*p*	*gamma*	*model*
0.0537	0.9049	2005.28	4035.26	75.3298	NA	sigmoidal
0.0083	0.9900	2040.24	4096.96	NA	NA	hyperbolic
0.0000	0.7745	2023.09	4062.65	NA	NA	parabolic
0.0000	0.7745	2019.89	4056.25	NA	2.4016	Flexible power

A central consideration in the design of the experiment was to understand how individual difference (ID) metrics contribute to effort discounting. These ID measures may be integrated into formal models of discounting to characterize a constant (additive) or modulatory (multiplicative) effect on decisions about effort expenditure.

Computational models.

Each model contained idiosyncratic parameters characterizing the degree k to which rewarding outcomes were discounted, or devalued, by effort, individual difference weight parameters indexed by (b, μ, κ), and the turning point of the sigmoid, p. Two main features were varied to build the model space: first, as described above, we adjusted the model's functional form to modify the shape of the discounting curves for four possible types of discounting: parabolic, power (free parabolic), hyperbolic, and sigmoidal. Second, we adjusted the position of the individual difference parameter i of subject j (i_j) and their weights (b, μ, κ), to best describe the effect of ID measures on participant decisions.

A key innovation of this study is a computational model that reliably estimates, from participant choices, weight parameters that modulate effort-based decision utilities as a function of the subclinical individual difference measures. This is to say, models that were fit to observed decisions per subject estimated specific ID weight parameters. We compared several models to formally characterize the contributions of individual differences to the discounting of rewards by effort costs.

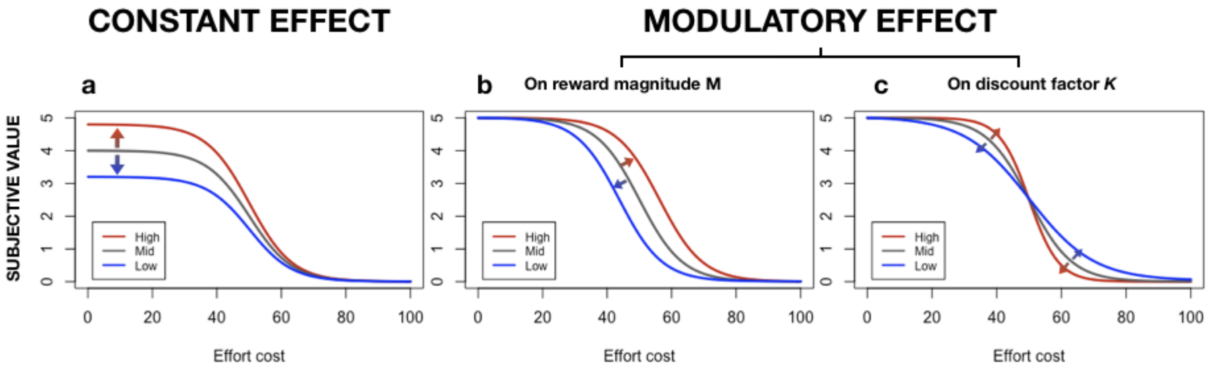

Figure 5. **Illustration of plausible effects of individual difference (ID) measures on the shape of effort discounting** (here, by cost only). Variations in trait approach/avoid measures may integrate into effort discounting functions as weights on ID values. Additive models, as illustrated here in (a), lead to constant (additive) changes in at the intercept of the value function. Multiplicative models, seen here in two possible forms (b and c), modulate the shape of the discounting curve. (b) illustrates changes in an ID weight modulating M, leading to proportional changes magnitude discounting and therefore indifference points. (c) illustrates IDs weighing on discount factor K, where variation in IDs leads to changes in the steepness of reward discounting by effort.

Baseline model (sigmoidal model)

The null model assumes no contribution of individual differences to discounting. We fit the model parameters across all participant decisions using maximum-likelihood estimation. Choices under the baseline model are characterized by an increasing devaluation of reward with growing effort costs, according to the function:

$$U = \beta_h + M\left(1 - \left(\frac{1}{1 + e^{-k(C-p)}} - \frac{1}{1 + e^{kp}}\right)\left(1 + \frac{1}{e^{kp}}\right)\right)$$

where utility U denotes the subjective value, M denotes the reward magnitude, and C the cost (in this case, physical effort). Following previous reports (Klein-Flugge et al., 2015[28], per Wu et al., 2011[29]) that physical effort discounts subjective value as opposed to pure monetary units (here, dollars and cents), we have set the utility parameter α on reward magnitude M to $\alpha = 0.8$ (again following Abdellaoui et al., 2007[27]). Two parameters, estimated by the data, describe the steepness (k) of the utility devaluation, and the turning point of the sigmoid (p). The turning point free parameter allows the sigmoid to become less steep at higher levels of effort, where the highest level of effort demands feel increasingly unattainable. Additionally, the baseline (null) model assumes that individual difference measures do not contribute to participant choices, but estimates constant β_h (intercept) to capture a general bias towards choosing to expend effort.

Additive

The additive-effect only model captures possible constant changes to decision utility, wherein the baseline probability of choosing to work for reward either increases or decreases linearly as a function of ID weights. Each separate individual difference measure of participant j is represented by (i_j), while weight b denotes the effect of said ID measure on the intercept. The additive model allows IDs to directly contribute only to changes at the intercept, thus entailing that they have no effect on the steepness of discounting. Choices under the additive model are governed by the utility function:

$$U_{add} = \beta_h + b(i_j) + M\left(1 - \left(\frac{1}{1+e^{-k(C-p)}} - \frac{1}{1+e^{kp}}\right)\left(1+\frac{1}{e^{kp}}\right)\right)$$

Multiplicative: k only

The multiplicative-k model sees ID weights place on rewards as a linear function of BIS/BAS, modulating the slope of discounting parameter k. The multiplicative model assumes that these specific individual differences do not contribute to changes at the intercept, but rather they govern the magnitude of i_j on the rate of effort discounting indexed by parameter κ. Choices under the multiplicative-k model are governed by the utility function:

$$U_{mult.k} = \beta_h + M\left(1 - \left(\frac{1}{1+e^{(-k+\kappa(i_j))(C-p)}} - \frac{1}{1+e^{(k+\kappa(i_j))p}}\right)\left(1+\frac{1}{e^{(k+\kappa(i_j))p}}\right)\right)$$

Multiplicative: Reward only

The multiplicative-reward model sees the weights participants place on counterpart payoffs as a linear function of IDs modulating the reward magnitude. The multiplicative model assumes that IDs modulate utility weights $\mu(i_j)$ on the reward magnitude M. Choices under the multiplicative-reward model are governed by the utility function:

$$U_{mult.r} = \beta_h + (1+\mu(i_j))M\left(1 - \left(\frac{1}{1+e^{-k(C-p)}} - \frac{1}{1+e^{kp}}\right)\left(1+\frac{1}{e^{kp}}\right)\right)$$

Multiplicative-additive (multadd): Full model

The multiplicative-additive model (multadd) is a combination of the effects of the above models. In this instantiation, additive weight parameter b modulates ID i_j at the intercept, while parameter weight κ modulates a separate ID measure y_j at slope parameter k. The additive effect varies the initial probability of choosing effort as a function of i_j, regardless of slope, while the multiplicative effect varies the steepness of discounting as a function of cost and a distinct ID measure, y_j. This model allows possible interactions between more than one individual difference measure, and choices under this model are governed by the utility function:

$$U_{mult.add} = \beta_h + (1+\mu(i_j))M\left(1 - \left(\frac{1}{1 + e^{(-k+\kappa(y_j))(C-p)}} - \frac{1}{1 + e^{(k+\kappa(y_j))p}}\right)\left(1 + \frac{1}{e^{(k+\kappa(y_j))p}}\right)\right)$$

Connecting utility to choices

Participant decisions, parametrically modulated by discrete units of effort costs and reward, were binary in nature (no effort = 0, effort = 1). Thus, we calibrated each subjective value function to behavior using the softmax function with inverse temperature parameter (lambda). The probability of a participant choosing the effortful option on a specific trial is given by:

$$P(effort)_j = \frac{1}{1 + exp\ (-\lambda(U(l_{t,j}) - U(h_{t,j}))}$$

where temperature *l* defines the steepness of the softmax function, and thus how sensitive the choice probability is to $\left(U(l_j) - U(h_j)\right)$, or the difference in subjective value between the no-effort lottery on trial t given subject j, $l_{t,j}$, and effort lottery on trial t, $h_{t,j}$. Bayesian Parameters were estimated using Maximum Likelihood Estimation (MLE), by maximizing the log-likelihood of all participant decisions, per model. Twenty random starting points were defined within the optimization function to capture the global minimum. Standard errors and p-values for the parameters were obtained from the inverse of the Fisher Information Matrix.

Modeling results.

The null model successfully predicted behavior far above chance ($R^2 = .828$), but the integration of trait-weighting parameter estimates and the BIS/BAS subscales increased prediction significantly. Specifically, the best-fitting model, which sees the BAS-drive variable and its estimated weight rho modulate the reward magnitude, outperformed this model significantly ($R^2 = 0.921$, LRT $\chi^2(2)$ $p < 2.42E\text{-}11$). Results for this specific model found in Table 5; full modeling results can be found in Supplementary Tables 1-7.

Model performance.

Null models (without individual difference parameters) were first fit to each subjects' choices in order to compute individual estimates for decision biases (beta), discount factors (k), and turning points (p). To evaluate which model fit best, we next ran an omnibus analysis over all subject choices. Resulting parameter MLEs are estimates of the type (e.g., additive, multiplicative) and weight placed on individual difference measures in the computation of effort/reward tradeoffs.

Selection.

Three types of model selection criteria were computed: Bayesian Information Criterion (BIC) to penalize the use of additional parameters over the null model, R^2 to evaluate the proportion of variance from observed choices, and likelihood ratio test (LRT), which considers degrees of freedom as number of additional estimated parameters from the null, to determine if the difference in likelihood scores among the null (M_0) and enriched ($M_{1..n}$) models is statistically significant. The results from these can be seen in Table 4 below. The model with an age parameter on the reward magnitude term, $(1+\mu_{age}(age_j))M$, appears at first to be the winning model overall. However, the skew of ages within this particular cohort (as seen in Figure Figure 8: Distribution of scales used in ensuing analyses for Berkeley cohort.) suggests that this effect may not be robust. Given the normal distribution of BIS/BAS parameters, we can see that the model with a BAS-drive parameter within the multiplicative reward term $(1+\mu_{BASd}(BASd_j))M$ is the best model. Here, the weight parameter μ_{BASd} = -0.01($BASd_j$), suggesting that as BASd increases, subjective value decreases.

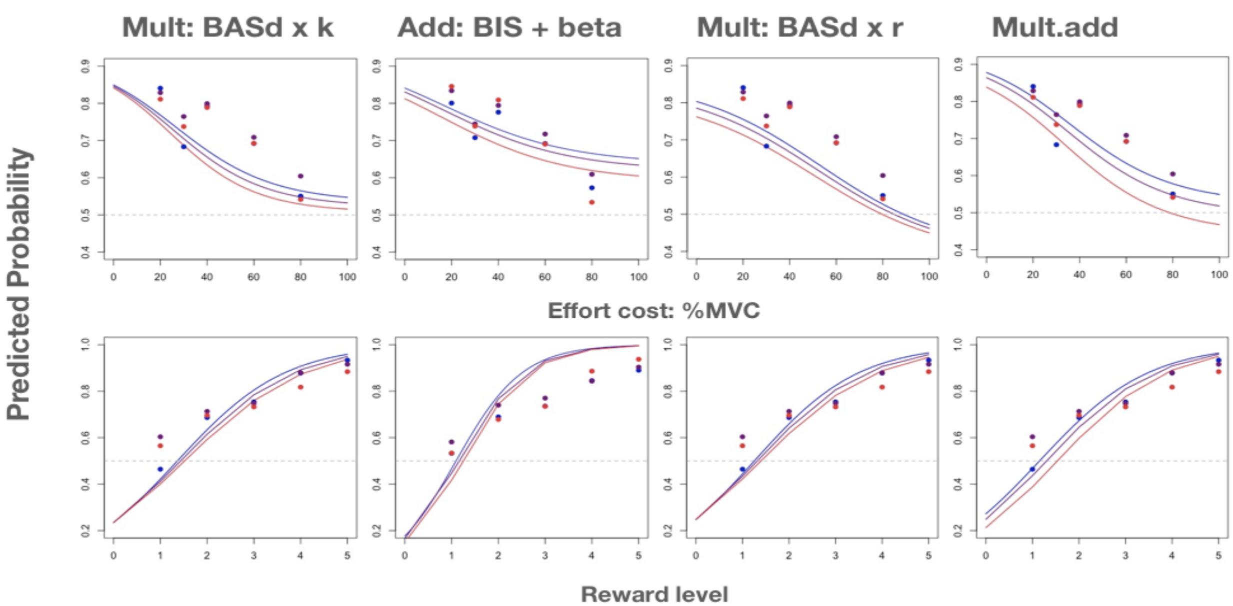

Figure 6. **Model fitted discount curves.** Curves show changes in shape and slope as a function of 1) ID measure and 2) manner of contribution to utility function (e.g., additive, multiplicative, etc). Top row: predicted probabilities of choosing effort as a function of effort costs. Bottom row: predicted probability of choosing effort as a function of reward offered by effort lottery.

Therefore, the winning individual difference model is the *multiplicative reward: BASd*, governed by the utility function:

$$U_{mult.r} = \beta_h + (1+\mu_{BASd}(BASd_j))M\left(1-\left(\frac{1}{1+e^{-k(C-p)}}-\frac{1}{1+e^{kp}}\right)\left(1+\frac{1}{e^{kp}}\right)\right)$$

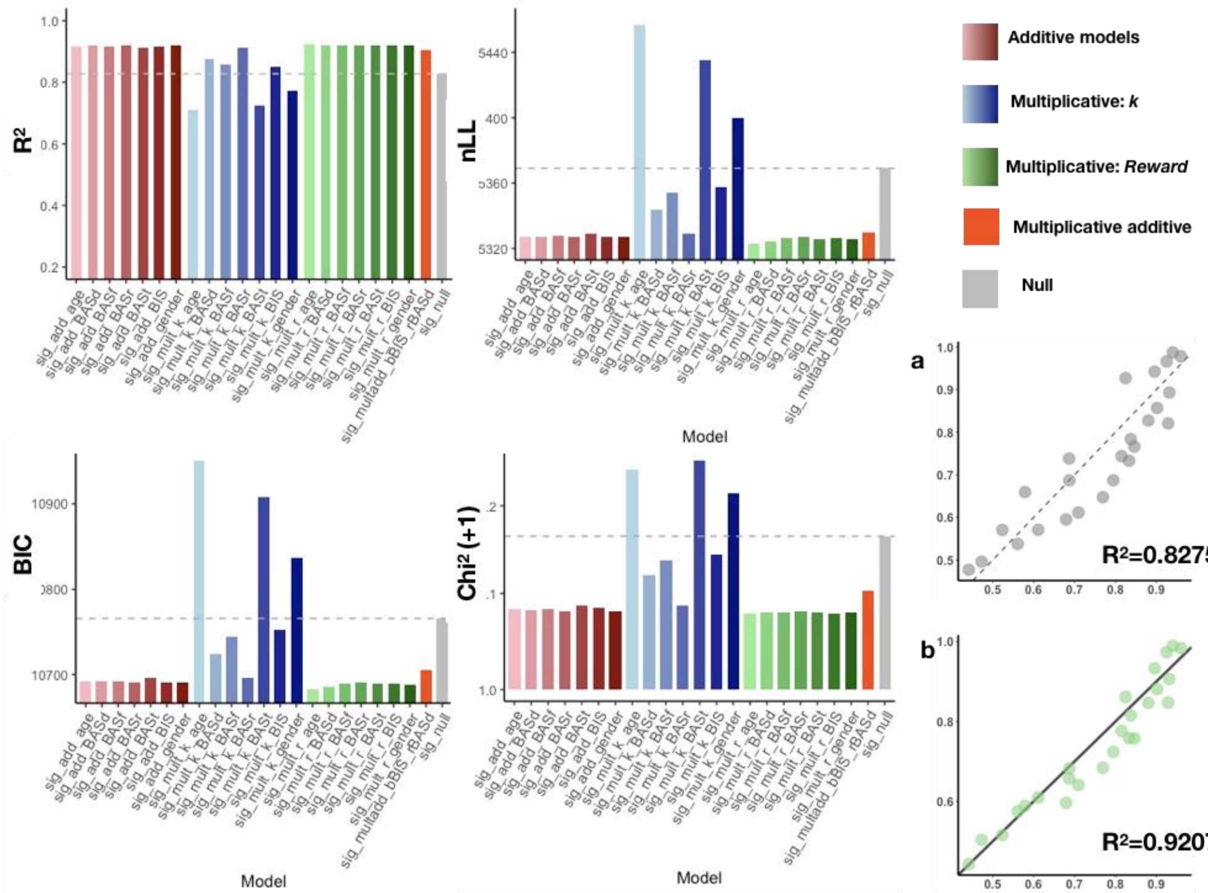

Figure 7. **Model metrics.** Four indicators of model fit show all 22 enriched models' performance against the null (basic) model (null model metrics shown in grey, with grey dashed line for reference). Most additive and reward-modulating multiplicative models perform above the null, while multiplicative models with ID measures modulating discount factor K tend to perform more poorly.

Table 4: **Performance metrics for computational models.** The best performing individual difference model includes a negative modulatory effect ($\mu_{BASd} = -9.73E-03$) of BAS-drive on reward magnitude, suggesting that those scoring lower on BAS-drive tended to value reward magnitude more highly, while those scoring higher on BAS-drive derived less subjective value on higher rewarding outcomes.

Model	*# est'd parameters*	*nLL*	*BIC*	χ^2	R^2	$P(>\chi^2)$
Mult: Age on reward	4	5323.13	10683.29	0.0782	0.9216	1.23E-11
Mult: BASd on reward	4	5324.45	10685.94	0.0793	0.9207	2.42E-11
Mult: gender on reward	4	5325.34	10687.73	0.0799	0.9203	3.82E-11
Mult: BASt on reward	4	5325.94	10688.93	0.0797	0.9205	5.19E-11
Mult: BIS on reward	4	5326.03	10689.11	0.0785	0.9214	5.43E-11
Mult: BASf on reward	4	5326.31	10689.67	0.0793	0.9208	6.27E-11
Mult: BASr on reward	4	5326.83	10690.70	0.0803	0.9200	8.17E-11
Add: BIS on intercept	4	5326.97	10690.99	0.0848	0.9158	8.78E-11

Add: BASr on intercept	4	5327.00	10691.03	0.0813	0.9191	8.88E-11
Add: gender on intercept	4	5327.08	10691.20	0.0810	0.9194	9.29E-11
Add: age on intercept	4	5327.32	10691.69	0.0828	0.9176	1.05E-10
Add: BASd on intercept	4	5327.36	10691.75	0.0819	0.9186	1.07E-10
Add: BASf on intercept	4	5327.83	10692.71	0.0832	0.9173	1.37E-10
Add: BASr on intercept	4	5329.26	10695.56	0.0870	0.9136	2.83E-10
Add: BASt on intercept	4	5329.37	10695.78	0.0870	0.9136	3.00E-10
Mult.add: add BIS, mult BASd on rew.	5	5329.86	10706.03	0.1028	0.9056	3.10E-09
Mult: BASd on K	4	5343.73	10724.49	0.1207	0.8769	4.86E-07
Mult: BASf on K	4	5353.89	10744.82	0.1371	0.8582	9.90E-05
Mult: BIS on K	4	5357.73	10752.51	0.1430	0.8514	7.70E-04
Null	3	5369.05	10765.87	0.1642	0.8275	1.00E+00
Mult: gender on K	4	5399.82	10836.68	0.2146	0.7718	1.00E+00
Mult: BASt on K	4	5435.29	10907.62	0.2545	0.7233	1.00E+00
Mult: age on K	4	5456.52	10950.09	0.2431	0.7092	1.00E+00

Table 5: **Results and parameter descriptions of winning model.** Parameter μ has a negative weighting on reward magnitude, suggesting that higher BAS-d

Parameters	*Tests of model-fit*	*parameter description*	*multiplicative: Reward*
β_h		constant for choosing effortful option	0.21133
μ		effect of ID on magnitude	-9.73E-03
k		discount factor	0.03216
p		turning point in the sigmoid	43.313
λ		inverse temperature parameter for softmax decision function	1.3252
	nLL	negative log-likelihood	5324
	BIC	Bayesian information criterion	10686
	χ^2	test of independence	0.07934
	R^2	variance explained	0.9207
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2) p < 2.42E-11

1.4 Discussion

Examinations of the contribution of individual differences to behavior are increasingly relevant to fields at the confluence of clinical and basic science. Here, we have presented one technique for using computational methods to characterize interactive processes between individual differences and motivated decision-making that may offer additional insight into putative cognitive models of effort discounting. Preliminary analyses show that while properties of the effort-reward tradeoff process were not related to variables or outcomes of previous trials, task choices were largely modulated by Behavioral Inhibition System (BIS) scores, and interactions between reward magnitude and the BAS sub-score related to motivation, Drive. Specifically, trait avoidance and motivation interact with response bias and effort costs respectively to influence the discounting of rewards by physical effort. Increases in the BAS subscore characterizing trait motivation contributes multiplicatively to the value of a decision's reward magnitude. Increases in the Behavioral Inhibition System subscore lead to concurrent linear decreases in response bias towards choosing effort.

Demographic variables like gender and race also perform above the null; however, these demographics cannot give us insight beyond changes in behavior, and into cognitive processes. The contribution of age to willingness to expend effort, however, may be a more complicated story. Indeed, age is shown to be a significant factor in all choice types. Age has demonstrable effects on cognitive and underlying neural processes, and studies examining age-related discrepancies in effort discounting suggest that effort is perceived as more costly for older adults (Westbrook et al., 2013). Given the younger skew of the data (median = 27 years, 77% of participants < 40 years), these data are not optimized for hierarchical analysis of age differences, but still provide some insight into the effects of age on willingness to allocate effort.

These results offer preliminary evidence that the magnitude of reward is negatively modulated by BAS-drive scores. The estimated weight parameter in this model, similar to the coefficients estimated from the logistic regressions, is negative and therefore suggests that as BAS-drive increases, the subjective value of the reward magnitude term decreases. These results may require further analyses in order to parse and explain such initially counterintuitive findings.

With data from 116 subjects, this analysis gives strong indication that these effects are robust, suggesting an extant relationship between model-estimated parameters and traits relating to inhibitory and appetitive motivation. Beyond the immediate analyses, this analytical framework may provide a path towards 1) the integration of diagnostic measures into computational analyses of effort-related behavior, 2) the estimation of subject-specific free-parameters that may have a closer link to underlying cognitive processes than self-report, and 3) a first step in connecting data-driven individual difference estimates with task-elicited neural activity (reported on in detail in the following chapter).

Chapter 1 Supplementary materials.

Supplementary Table 1: Logistic mixed effects regression results for BAS: total, reward-responsiveness, drive.

	BASt			***BASr***			***BASd***		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.94	0.01 – 67.56	0.979	4.65	0.14 – 155.69	0.391	0.47	0.02 – 10.08	0.626
BASt	1.04	0.88 – 1.22	0.658						
Response bias	2.08	0.02 – 256.19	0.766	0.42	0.01 – 27.82	0.688	6.17	0.16 – 245.32	0.333
utility	6.35	4.43 – 9.10	**<0.001**	4.13	3.11 – 5.50	**<0.001**	6.17	4.69 – 8.13	**<0.001**
Effort cost	0.97	0.96 – 0.99	**<0.001**	0.97	0.96 – 0.98	**<0.001**	0.98	0.97 – 0.98	**<0.001**
trial	1	0.99 – 1.00	0.531	1	0.99 – 1.01	0.842	1	0.99 – 1.01	0.734
Previous mag.	0.81	0.67 – 0.98	**0.033**	0.94	0.81 – 1.10	0.452	0.83	0.71 – 0.96	**0.012**
streak	0.63	0.32 – 1.24	0.177	0.59	0.34 – 1.03	0.064	0.7	0.41 – 1.18	0.185
Previous effort	1.32	0.72 – 2.42	0.366	1.27	0.78 – 2.07	0.332	1.01	0.63 – 1.61	0.974
*BASt * Response bias*	1	0.83 – 1.19	0.963						
*BASt * utility*	0.98	0.96 – 0.99	**0.001**						
*BASt * Effort cost*	1	1.00 – 1.00	0.868						
*BASt * trial*	1	1.00 – 1.00	0.07						
*BASt * Previous mag.*	1.01	1.00 – 1.02	**0.039**						
*BASt * streak*	1.02	0.99 – 1.04	0.214						
*BASt * Previous effort*	0.98	0.96 – 1.01	0.192						
BASr				0.92	0.59 – 1.43	0.705			
*BASr * Response bias*				1.21	0.72 – 2.04	0.474			
*BASr * utility*				0.98	0.95 – 1.02	0.268			
*BASr * Effort cost*				1	1.00 – 1.00	0.381			
*BASr * trial*				1	1.00 – 1.00	**0.046**			

*BASr * Previous mag.*				1.01	0.99 – 1.03	0.523			
*BASr * streak*				1.06	0.99 – 1.14	0.078			
*BASr * Previous effort*				0.96	0.90 – 1.01	0.137			
BASd							1.21	0.86 – 1.71	0.277
*BASd * Response bias*							0.87	0.58 – 1.31	0.513
*BASd * utility*							0.94	0.91 – 0.97	**<0.001**
*BASd * Effort cost*							1	1.00 – 1.00	0.956
*BASd * trial*							1	1.00 – 1.00	0.128
*BASd * Previous mag.*							1.02	1.00 – 1.04	**0.013**
*BASd * streak*							1.04	0.98 – 1.10	0.233
*BASd * Previous effort*							0.99	0.94 – 1.04	0.585
Random Effects									
σ^2	3.29			3.29			3.29		
τ_{00}	1.80 SID			1.77 SID			1.79 SID		
ICC	0.35			0.35			0.35		
N	116 SID			116 SID			116 SID		
Observations	10420			10420			10420		
Marginal R^2 / Conditional R^2	0.290 / 0.541			0.289 / 0.538			0.294 / 0.543		

Supplementary Table 2: Logistic mixed effects regression results for BAS: fun-seeking, and BIS.

	BASf			***BIS***		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	1.99	0.04 – 92.48	0.726	0.51	0.00 – 146.58	0.814
BASf	1.02	0.68 – 1.53	0.915			
beta.intercept	1.77	0.02 – 136.07	0.796	14.81	0.02 – 11852.07	0.429
utility	4.9	3.67 – 6.54	**<0.001**	1.17	0.75 – 1.82	0.486

percentageOfMVC	0.98	0.97 – 0.99	**<0.001**	1.01	0.99 – 1.02	0.211
trialNumber	0.99	0.98 – 1.00	**0.006**	0.99	0.97 – 1.00	**0.023**
prev_mag	0.88	0.75 – 1.02	0.097	0.97	0.77 – 1.23	0.814
streak	0.91	0.53 – 1.59	0.75	0.65	0.27 – 1.54	0.326
prev_eff	1.14	0.69 – 1.89	0.613	1.14	0.53 – 2.45	0.733
*BASf * beta.intercept*	1.01	0.64 – 1.58	0.979			
*BASf * utility*	0.97	0.94 – 0.99	**0.022**			
*BASf * percentageOfMVC*	1	1.00 – 1.00	0.794			
*BASf * trialNumber*	1	1.00 – 1.00	0.268			
*BASf * prev_mag*	1.01	1.00 – 1.03	0.116			
*BASf * streak*	1	0.95 – 1.06	0.874			
*BASf * prev_eff*	0.97	0.92 – 1.03	0.323			
BIS				1.11	0.76 – 1.63	0.589
*BIS * beta.intercept*				0.87	0.56 – 1.36	0.544
*BIS * percentageOfMVC*				1	1.00 – 1.00	**<0.001**
*BIS * trialNumber*				1	1.00 – 1.00	0.24
*BIS * prev_mag*				1	0.99 – 1.02	0.885
*BIS * streak*				1.03	0.97 – 1.09	0.367
*BIS * prev_eff*				0.98	0.94 – 1.03	0.512
Random Effects						
σ^2	3.29			3.29		
τ_{00}	1.79 SID			1.79 SID		
ICC	0.35			0.35		
N	116 SID			116 SID		
Observations	10420			10420		
Marginal R^2 / Conditional R^2	0.289 / 0.539			0.298 / 0.546		

Supplementary Table 3: Basic (unenriched) model fits.

Parameters	*Tests of model-fit*	*parameter description*	*basic model*
β_h		constant for choosing effortful option	0.1402
k		discount factor	0.048
p		turning point in the sigmoid	8.551
λ		inverse temperature parameter for softmax decision function	1.085
	nLL	negative log-likelihood	5369.05
	BIC	Bayesian information criterion	10765.87
	χ^2	test of independence	0.1642
	R^2	variance explained	0.8275

Supplementary Table 4: Age mode fits. Age-weight kappa (parameter weight) is negative, suggesting that the higher the age, the less discounting we see.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*multiplicative: Reward*	*multiplicative: K*
β_h		constant for choosing effortful option	0.118	0.0374	0.4024
b		effect of ID on constant	0.0015	-	-
ρ		effect of ID on reward	-	0.0032	-
κ		effect of ID on K	-	-	-9.23E-05
k		discount factor	0.034	0.032	0.097
p		turning point in the sigmoid	0.34	46.3	3.78
λ		inverse temperature parameter for softmax decision function	1.2	1.104	1.095
	nLL	negative log-likelihood	5327.3	5323.1	5456.5
	BIC	Bayesian information criterion	10692	10683	10950
	χ^2	test of independence	0.0827	0.0782	0.2431
	R^2	variance explained	0.9176	0.9216	0.7093
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2)	χ^2 (2)	χ^2 (2)
			$p < 1.1e-10$	$p < 1.2e-11$	p = 1

Supplementary Table 5: BIS model fits.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*multiplicative : Reward*	*multiplicative : K*
β_h		constant for choosing effortful option	0.379	0.0528	0.18914
b		effect of ID on constant	-0.0135	-	-
ρ		effect of ID on reward	-	0.0054	-
κ		effect of ID on K	-	-	-9.76E-05
k		discount factor	0.0352	0.032	0.0517
p		turning point in the sigmoid	30.304	46.76	10.76
λ		inverse temperature parameter for softmax decision function	1.225	1.119	1.181
	nLL	negative log-likelihood	5327	5326	5357.7
	BIC	Bayesian information criterion	10691	10689	10753
	χ^2	test of independence	0.0848	0.0785	0.143
	R^2	variance explained	0.9158	0.9214	0.8513
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2)	χ^2 (2)	χ^2 (2)
			p < 8.8e-11	p < 5.4e-11	p < 7.7e-04

Supplementary Table 6: BAS-total model fits.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*multiplicative : Reward*	*multiplicative : K*
β_h		constant for choosing effortful option	0.1812	0.20394	1.07E-01
b		effect of ID on constant	2.78E-05	-	-
ρ		effect of ID on reward	-	-2.79E-03	-
κ		effect of ID on K	-	-	2.26E-04
k		discount factor	0.0364	0.03233	0.02279
p		turning point in the sigmoid	27.58	42.094	1.171
λ		inverse temperature parameter for softmax decision function	1.228	1.309	0.6139
	nLL	negative log-likelihood	5329	5326	5435
	BIC	Bayesian information criterion	10696	10689	10908
	χ^2	test of independence	0.08704	0.07966	0.25446
	R^2	variance explained	0.9136	0.9205	7.23E-01
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2)	χ^2 (2)	χ^2 (2)
			$p < 2.9e-10$	$p < 5.19E-11$	$p = 1.0e+00$

Supplementary Table 7: BAS-reward responsiveness model fits.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*multiplicative : Reward*	*multiplicative : K*
β_h		constant for choosing effortful option	0.10571	0.13217	0.17707
b		effect of ID on constant	5.99E-03	-	-
ρ		effect of ID on reward	-	1.51E-03	-
κ		effect of ID on K	-	-	-1.37E-04
k		discount factor	0.03313	0.03252	0.03741
p		turning point in the sigmoid	37.3	40.622	27.58
λ		inverse temperature parameter for softmax decision function	1.2194	1.2025	1.2282
	nLL	negative log-likelihood	5327	5327	5329
	BIC	Bayesian information criterion	10691	10691	10696
	χ^2	test of independence	0.08125	0.08025	0.08703
	R^2	variance explained	0.9191	0.92	0.9136
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2)	χ^2 (2)	χ^2 (2)
			p < 8.89E-11	p < 8.17E-11	p = 2.83E-10

Supplementary Table 8 : BAS-drive model fits.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*multiplicative : Reward*	*multiplicative : K*
β_h		constant for choosing effortful option	0.18947	0.21133	0.23466
b		effect of ID on constant	-3.65E-03	-	-
ρ		effect of ID on reward	-	-9.73E-03	-
κ		effect of ID on K	-	-	6.02E-04
k		discount factor	0.03361	0.03216	0.04196
p		turning point in the sigmoid	35.361	43.313	13.971
λ		inverse temperature parameter for softmax decision function	1.2216	1.3252	1.2243
	nLL	negative log-likelihood	5327	5324	5344
	BIC	Bayesian information criterion	10692	10686	10724
	χ^2	test of independence	0.08189	0.07934	0.12074
	R^2	variance explained	0.9186	0.9207	0.8769
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2)	χ^2 (2)	χ^2 (2)
			$p < 1.07E-10$	$p < 2.42E-11$	$p = 4.86E-07$

Supplementary Table 9: BAS- fun seeking model fits.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*multiplicative : Reward*	*multiplicative : K*
β_h		constant for choosing effortful option	0.16991	0.16884	0.20877
b		effect of ID on constant	-6.01E-04	-	-
r		effect of ID on reward	-	-4.16E-03	-
k		effect of ID on K	-	-	2.71E-04
k		discount factor	0.03422	0.03213	0.0468
p		turning point in the sigmoid	33.048	43.939	11.512
λ		inverse temperature parameter for softmax decision function	1.2227	1.2617	1.1901
	nLL	negative log-likelihood	5328	5326	5354
	BIC	Bayesian information criterion	10693	10690	10745
	χ^2	test of independence	0.08322	0.07927	0.13707
	R^2	variance explained	0.9173	0.9208	0.8582
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2)	χ^2 (2)	χ^2 (2)
			$p < 1.37\text{E-}10$	$p < 6.27\text{E-}11$	$p = 9.90\text{E-}05$

Chapter 2

Connecting integrated individual difference measures in effort-based decision-making to functional brain activity.

2.1 Introduction

Motivation is a complex and adaptive process that is critical for survival, and its behavioral instantiation is shown, following the previous chapter, to be modulated by ostensibly latent but integrated individual differences (IDs). These IDs, and the ways in which they weigh on motivation, involve multiple behavioral functions mediated by an array of interacting neural circuits. How humans differ from one another in their valuation of costs and benefits has been linked to differences in facets of personality, psychopathology, and brain structure and function. Investigations into the relationship between individual differences in traits and brain structure (Bonnelle et al., 2016[30]), dopaminergic mechanisms (Treadway et al., 2012[2]), and behavior (Chong et al., 2015[1]) have offered various ways to dissect motivation into its component elements, providing possible clues for more systematic and accurate classifications of clinical and subclinical inter-individual variation.

The existence of a connection between behavioral signatures of motivation and the brain function and structure supporting them have been well studied in model organisms, and more recently in humans as well. Treadway and colleagues describe findings that the existence of individual differences (IDs) in dopamine (DA) receptor availability in the mesoaccumbens DA pathway regulates motivational differences. There is additional evidence that in nucleus accumbens (NAc), DA mediates its effects through D1- and D2-like receptors (Salamone, 2009[8]). The anterior cingulate cortex (ACC) has been proposed to monitor task difficulty, possibly indexed by conflicts in information processing (Botvinick et al., 2004[31]; Ridderinkhof et al., 2004[32]). Previous examinations offer evidence that the ACC, particularly dorsal portions of the ACC, plays major a role in effort-based decision making (Botvinick, 2007[33]; Rushworth et al., 2004[34]; Walton et al., 2003[35]). Furthermore, on an anatomical level, the ACC sends direct connections to the NAc (Croxson et al., 2005[36]; Kunishio & Haber, 1994[37]). These findings suggest that the ACC and NAc may be of primary importance in mediating effort-based decisions in humans. However, there remains a dearth of knowledge regarding the neurocomputational mechanisms that might account for apathy across populations. These functional regions and their relationship to the facilitation of human motivation, and its associated variation between people, are therefore of clearly defined interest to this work. As of the writing of this thesis, we have yet to discover whether and how self-report level and computationally estimated IDs are borne out in real time in the human brain during cost/benefit choice behavior specific to effort. This second chapter aims to investigate a plausible link between neural data and an integrative computational mechanism underlying individual differences in motivated decisions. While investigations into effort-based decision-making and the models of decisions that humans use to guide them are not new, the approach I have aimed to take – to identify specific, individual-level traits and integrate those into a formal mathematical model with both explanatory and predictive power – is novel. In this chapter, I hope to combine computational modeling approaches with functional magnetic resonance imaging (fMRI) as a step towards a method with greater sensitivity and specificity of classifying aberrant motivation.

2.2 Materials & Methods

2. 2.1 Participants

The study was approved by the Committee for Protection of Human Subjects at the University of California, Berkeley. Data from 64 neurologically healthy individuals were included in the reported analyses. Participants were 68.7% female (3 DNR), aged 18-55 (mean = 27.4 ± 7.3). The monetary compensation participants received at the end of the experiment depended on their performance on the task inside the scanner, and varied between $15 and 25. A comparison of the Berkeley versus Miami cohorts can be seen in Table 7 below.

Table 6: Descriptive statistics of demographic and individual difference measures.

	Mean	*Median*	*Range*	*SD*
N = 67				
Age	27.38	26	18, 55	7.35
Gender (%Male)	31.3%			
BIS	14.4	14	7, 19	2.22
BAS-total	25.94	25	13, 40	6.23
BAS-rr	8.39	8	5, 16	2.36
BAS-drive	9.07	9	4, 15	2.68
BAS-fs	8.48	8	4, 15	2.67

Table 7: Comparison of group variances using Levene's test. Groups differ significantly outside of race, likely due to diversity of ethnic and racial pool of Miami, FL, versus Berkeley, CA. Standard deviation of

	F-val	*Pr(>F)*	*Signif.*
Age	2.24	0.14	-
Gender	2.11	0.15	-
Race	21.9	4.4e-06	***
BIS	1.2	0.27	-
BAS-total	7.79	0.0056	**
BAS-rr	2.97	0.086	-
BAS-drive	1.42	0.23	-
BAS-fs	2.37	0.12	-

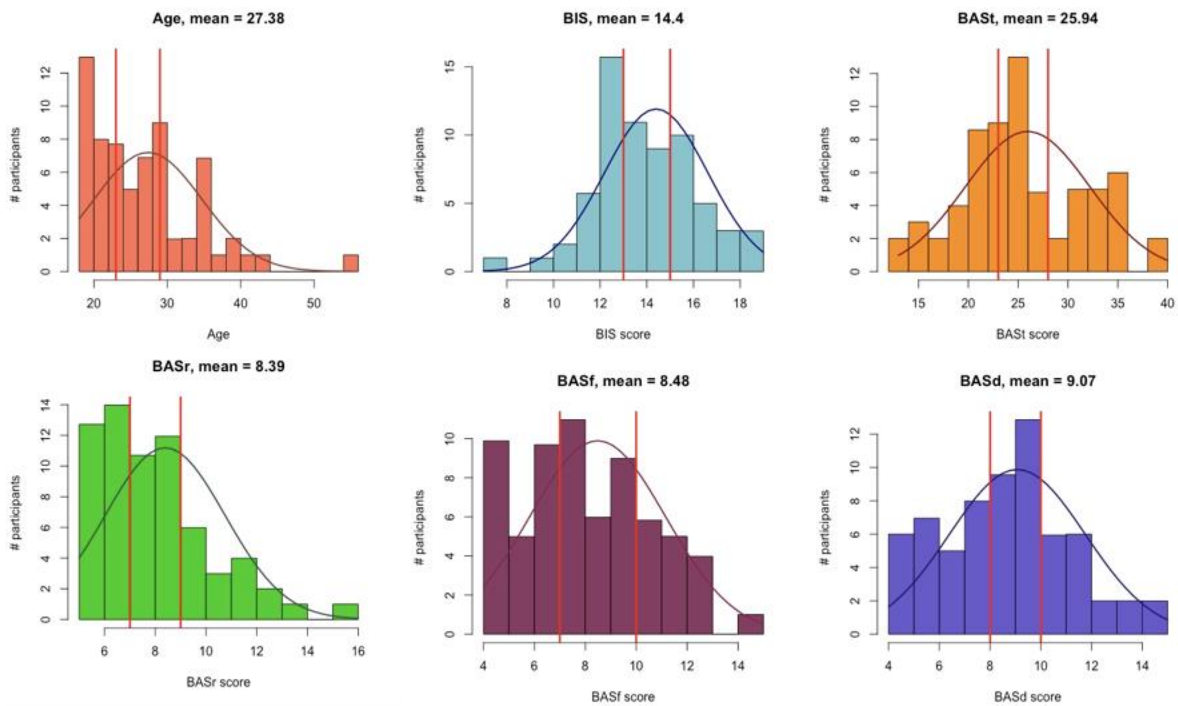

Figure 8: Distribution of scales used in ensuing analyses for Berkeley cohort.

2.2.2 Behavioral Task

Participants completed the same experimental paradigm as described in Chapter 1, but the behavioral data described here were collected during functional magnetic resonance imaging (fMRI), using a different, MR-compatible grip force recording device. As in the behavioral-only version, participants were given 2 breaks throughout (after trials 30, 60) to prevent fatigue. For this reason, data were initially preprocessed as three separate runs, and analyzed in the same GLM (described in detail below). Unlike the behavioral-only task, participants were asked to squeeze with their non-dominant hand, in order to respond using a button box held in their dominant hand.

2.2.3 Hardware

Stimulus presentation was programmed in-house on a PC using Python library PyGame (Shinners et al., 2011[38]; http://pygame.org/). Stimuli and feedback generated using PyGame and in-house code were presented to participants via MR-compatible back-projection system (Simulink), viewed through a mirror placed above the receiver coil. Participant choices were recorded with a button box, and effort output was recorded using an MR-compatible dynamometer (Current Designs, Inc., Philadelphia, PA), which was connected to the experimental computer using a fiberoptic cable. Grip force was recorded using the handheld dynamometer. The PC screen provided subjects with real-time visual feedback on the force being exerted.

2.2.4 Imaging Procedure

Brain imaging data were collected at the UC-Berkeley Brain Imaging Center on a Siemens 3T TIM Trio system, using a 32-channel phase-arrayed head coil. High-resolution T1-weighted magnetization-prepared rapid gradient-echo (MP-RAGE) anatomical images were acquired for spatial normalization. Function T2*-weighted echo planar images were acquired during task scans (three runs of 30 trials each, plus two 30s breaks). Images were collected using the following scanning parameters: TR=2000ms, TE=26ms, flip-angle=60°, 32 transverse slices, voxels=3mm^3.

2.3 Imaging Analyses

2.3.1 Preprocessing

All neuroimaging data analyses were completed in SPM12 (Wellcome Department of Imaging Neuroscience, London, UK). Image preprocessing, denoising, and quality assurance steps applied to structural and functional fMRI data include: slice time and distortion field correction, functional scan realignment, coregistration of functional data to anatomical images, normalization to MNI standard space, band-pass filtering, and spatial smoothing (FWHM at 6mm Gaussian kernel). Hypotheses were tested using the general linear model (GLM) by regressing subject-specific effort-discounting and individual difference parameters against the hemodynamic responses to task-related events of interest (effort versus no-effort decisions, win versus loss lottery outcomes).

2.3.2 First-level analyses

Mixed-effects logistic regression analyses

Following the same preliminary analytical methods as Chapter 1, behavioral task choices from scanner participants were first analyzed with a generalized logistic mixed effects model using the lme4 package in R. The dependent variable was chosen effort (1=effortful lottery choice, 0=no effort lottery choice). The null model included a subject-specific response-bias parameter, beta, and the following fixed effects: reward magnitude (*R(effort) – R(no effort)*), and effort cost as %MVC. Results from these analyses can be found in Supplementary Table 1.

Computational modeling of behavior

Following the same non-linear modeling analyses as Chapter 1, basic sigmoidal models, without subject-specific individual difference parameters, were run on each subject in order to compute individual maximum likelihood estimates per subject. Discount factor *k,* turning point *p,* and constant *beta*, along with BIS/BAS measures and age, would all be entered as covariates in second-level analyses on imaging data to examine whether any of these individualized parameters correlated with % signal change at various events in the fMRI experiment. Subject specific measures and additional covariates can be found in Supplementary Table 3. First level computational modeling results are described in further detail in Supplementary Table 4.

General Linear Models of fMRI data

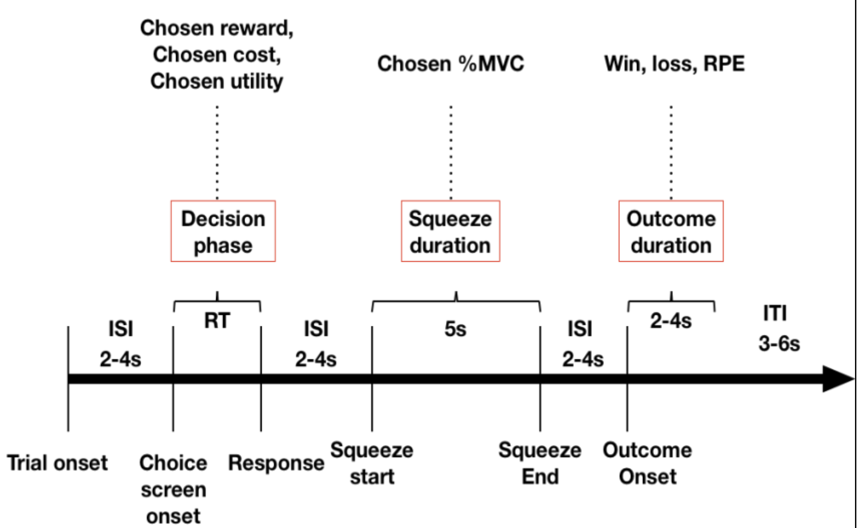

Figure 9. **Representation of how events were modeled in imaging GLMs.** Modeled events highlighted in red boxes; above them, dotted lines point to (a non-exhaustive list of) examples of parametric modulators on events. Subjects were given a total of 5 seconds to make their decisions, and a total of 5 seconds to squeeze (or wait for no-effort lottery choices). The decision event was modeled as a function of reaction time (subject-driven). The squeeze events were modeled as squeeze duration minus 2s to account for action preparation. All outcome durations, ISIs, and ITIs were jittered and not dependent upon subject behavior.

FMRI data were analyzed using voxel-wise time series analysis within the framework of the General Linear Model (GLM). To this end, a design matrix was generated with a simulated hemodynamic response function (HRF; gamma convolution, phase = 0 s, SD = 3s, mean lag = 6s) and its first temporal derivative. Several event types were distinguished.

We designed an initial model (GLM0) that included regressors to model the squeeze event, which showed that motor demands were indeed recruiting the expected sensorimotor cortical responses, seen in Figure 6 below. This GLM used only decision, squeeze duration, and outcome as regressors. Additionally, given the length of the action/no-action event in the GLM, and the large signal resulting from the squeeze, a number of instantiations were run to best estimate the true timing of motor activity: 1) using variable event durations from subject data, 2) using a fixed even duration (5s), equal to the amount of time subjects were given to complete the action phase (either squeeze or no-squeeze), and 3) an amount equivalent to subject-specific timing as in (1), but with ±2s at each end of the squeeze duration in order to account for mental preparation beforehand, and signal decay after. The GLMs that best modeled squeeze timing followed (1), using simply the subject-specific squeeze duration. All GLMs henceforth relied on this timing structure.

Next, we included a number of separate regressors modeling the main three events: incentive display (choice screen), effort exertion with a boxcar function (7s), and outcome display. These three distinct categorical regressors were, respectively, modulated by the following parameters in different GLMs: expected reward magnitude, expected utility, expected effort costs, reward prediction error (RPE) on the previous trial, whether/how much was won on the previous trial, and whether/how much effort was expended on the previous trial (incentive display); level of effort demand and length of effort expenditure (effort exertion); level of monetary earnings, whether/level of effort expended, and current RPE (for outcome display). To correct for motion artifacts, participant-specific realignment parameters were included as covariates of no interest in all the GLMs employed to analyze functional activations.

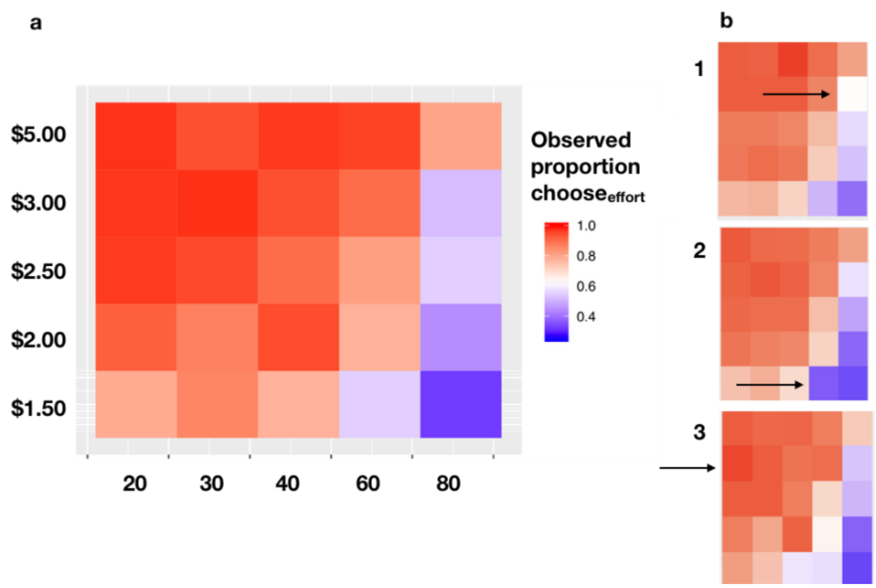

Figure 10: **Observed choices.** a) Observed probabilities across all subjects. b) Observed probabilities by BISBAS terciles (1 = first, 2 = second, 3 = third). Variation in behavior can be seen by changes in heatmap colors; arrows point to a selection of specific effort/reward combinations showing changes from overall. This representation is shows that there are behavioral disparities between those scoring higher and lower on BIS/BAS subscores. The relationship between these measure neural responses to effort/reward tradeoffs and are not yet known.

Logistic regression analyses of behavior.

The effect of this cohort's trait scores on behavior differed greatly from the Miami cohort. Interestingly, the best most significant trait contribution was reported sex. Indeed, while both males and females strongly preferred to expend effort (mean = 76.93% effort choices), males expended effort nearly 12% more often (mean = 84.99% effort choices, N=21), while females expended effort at a rate closer to 73% (mean = 73.26%, N=46).

Behavior from inside the scanner was put through the same set of model-free and model-based behavioral assays as the Miami data. This is to say, behavioral task choices were first analyzed with mixed-effects logistic regressions, including the same main regressors of interest: %MVC as cost, utility of reward magnitude as reward, trial number, and previous trial information. Predicted curves are shown below, and give insight into the manner in which these ID measures may contribute to effort discounting.

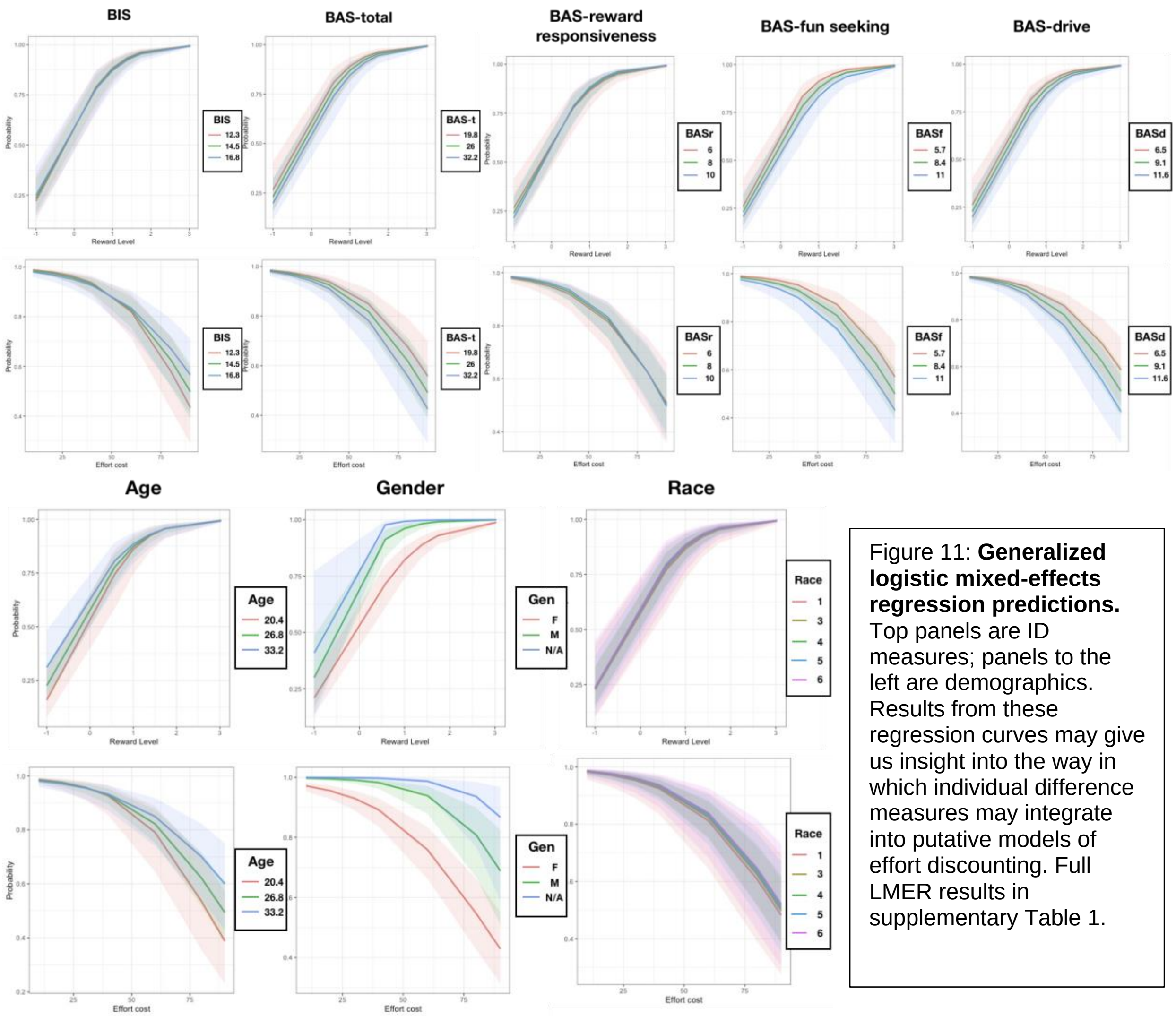

Figure 11: **Generalized logistic mixed-effects regression predictions.** Top panels are ID measures; panels to the left are demographics. Results from these regression curves may give us insight into the way in which individual difference measures may integrate into putative models of effort discounting. Full LMER results in supplementary Table 1.

2.4 Results

2.4.1 Modeling results.

The Berkeley and Miami cohort show some important similarities, but also some notable differences. As in the Miami group, BAS-drive seems to be an important predictor of choosing to effort: as BAS-d increases, so does the propensity for choosing effort at higher reward magnitudes, but not at higher costs. These results show that gender contributes most strongly to the computation of effort-reward tradeoffs. Similarly to the Miami cohort, BAS-drive, a proxy for motivation, also adds predictive power as modulator of reward magnitude.

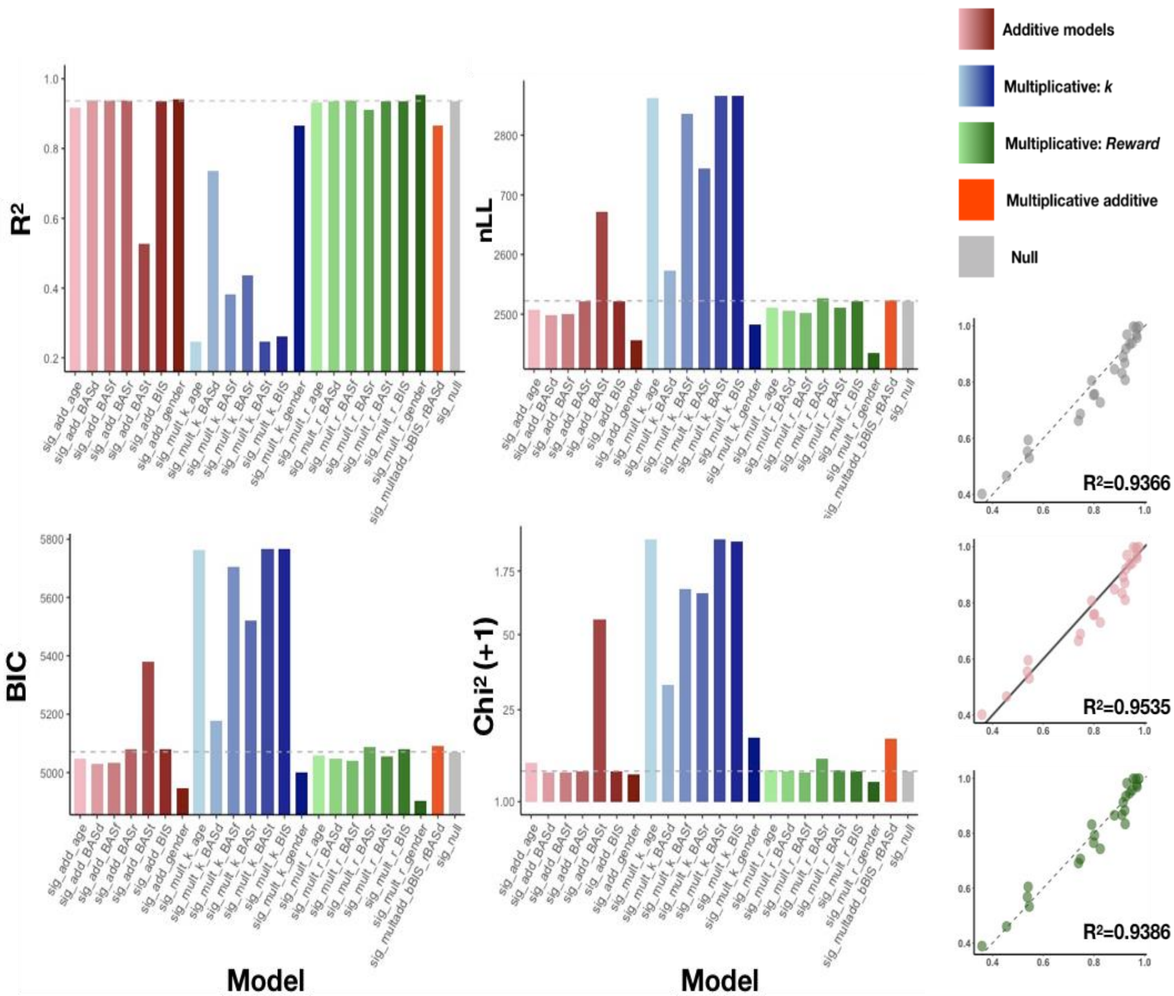

Figure 12 **Discounting model metrics.** Four indicators of model fit show all 22 enriched models' performance against the null (basic) model (null model metrics shown in grey, with grey dashed line for reference). Most additive and reward-modulating multiplicative models perform above the null, while multiplicative models with ID measures modulating discount factor K tend to perform more poorly.

2.4.2 Imaging Results

As a first check, we examined visuomotor systems during the squeeze event at both the individual and group-levels. Individuals with no discernable motor activity over baseline during the squeeze events were excluded (N=64 after excluding 3 participants). All ensuing models included squeeze duration and squeeze cost regressors in order to control for the large signal resulting from effort expenditure.

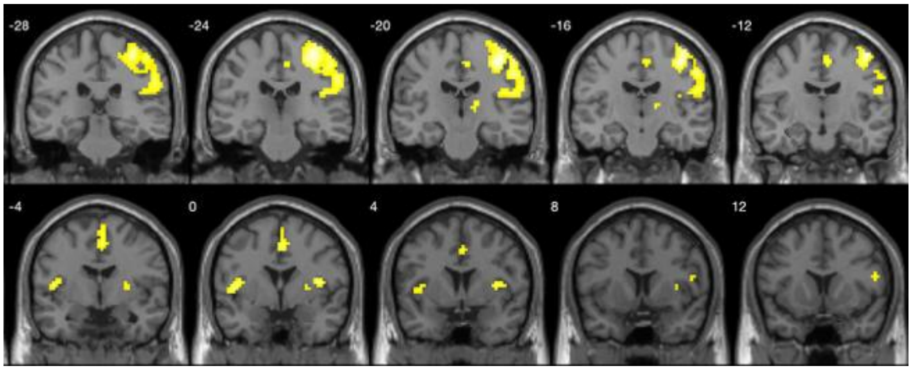

Figure 13. **Statistical parametric map of squeeze event**. BOLD signal associated with the cost of effort during squeeze event. Regions showing significant increase in activation with increases in effort costs. Top panels show activity in SMA and somatosensory cortices while bottom show activity in dorsal ACC and insula. Peak voxel statistics shown in Table 2.

Table 8. Local maxima for group-level action-phase events (p-values adjusted for search-volume).

	MNI coordinates (peak voxel)			*z-stats*	P_{uncorr}	P_{fwe}	* *survives correction*
	x	*y*	*z*				
Squeeze							
R Postcentral Gyrus	39	-19	50	7.32	< 0.001	< 0.001	*
L Insula	-45	-1	5	6.65	< 0.001	< 0.001	*
L Dorsal PFC	-36	47	32	3.98	< 0.001	0.158	
L Postcentral Gyrus	-9	-25	47	3.38	< 0.001	0.662	
Effort Cost							
R Postcentral Gyrus (SMA)	30	-25	53	6.68	< 0.001	< 0.001	*
R Insula	42	2	14	4.26	< 0.001	0.081	

Next, we examined various regressors during the choice phase. These included the evaluation of expected (chosen) effort costs modulated by the level of effort required by the chosen lottery, expected reward value modulated by chosen reward magnitude (difference of effort lottery value minus no-effort lottery value), subjective value of chosen lottery (derived from utility function of winning computational model) and four previous trial regressors: 1) outcome (whether won or lost), 2) reward magnitude of won trials, 3) RPE of the previous trial, and 4) RPE of previous trial given effort expenditure (Figure 10).

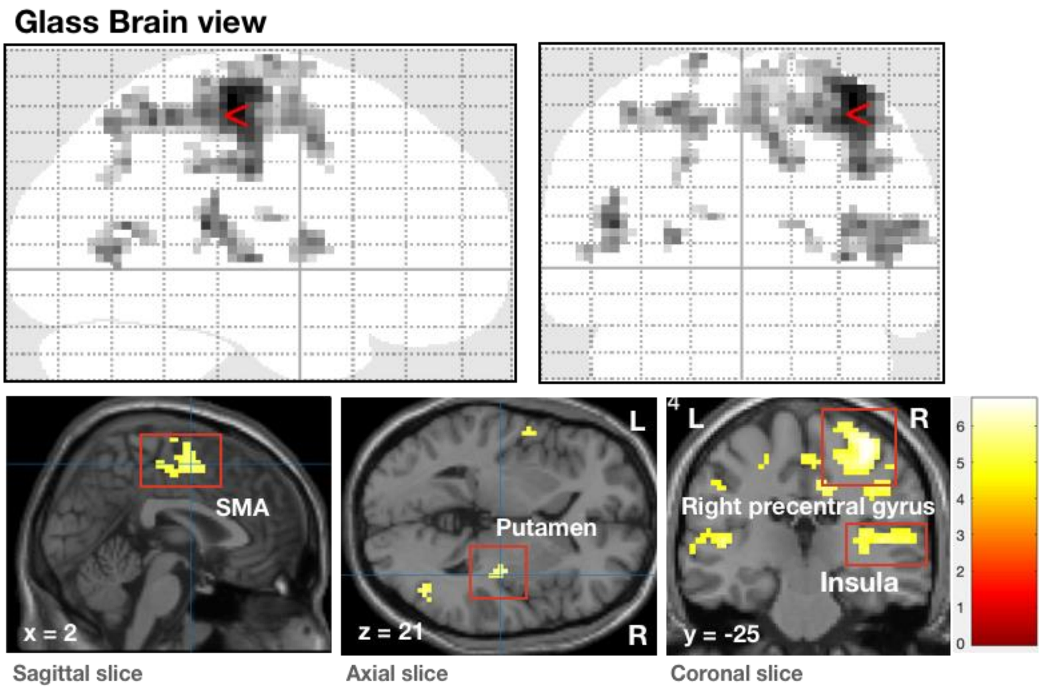

Figure 14. **SPM of neural response to expected effort costs**. BOLD signal associated with effort evaluation. Lefthand panels show BOLD signal change in supplementary motor area and dACC; righthand panels show activity in right somatosensory cortex, both during the choice phase in response to variation in effort lottery costs. Legend heatmap colors correspond to z-statistics.

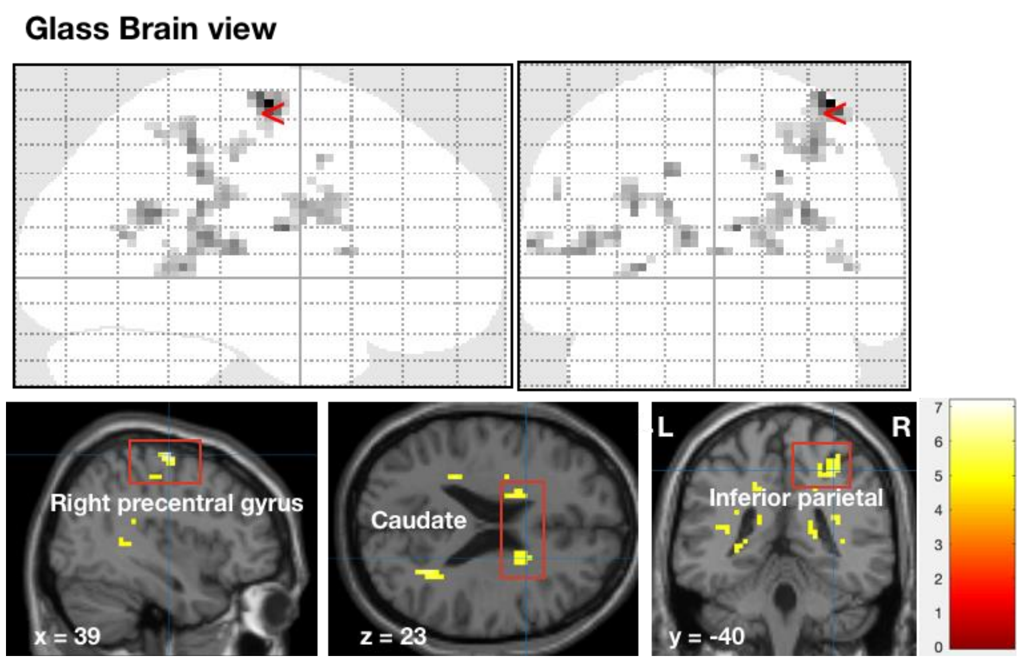

Figure 15. **SPM of neural response to expected reward magnitude**. BOLD signal change associated with expected reward during decision-phase. Slice views show activity in right precentral gyrus, bilateral caudate, and inferior parietal gyrus. Legend heatmap colors correspond to z-statistics.

Table 9. Local maxima for group-level decision-phase events (p-values adjusted for search-volume).

	MNI coordinates (peak voxel)			z-stats	P uncorr	P fwe	* survives correction
	x	y	z				
Expected effort cost							
R Precentral Gyrus	36	-29	59	5.83	< 0.001	< 0.001	*
Posterior superior temporal Gyrus	-45	-34	14	5.50	< 0.001	< 0.001	*
L Precentral Gyrus	-15	-13	70	5.17	< 0.001	0.002	*
L Putamen	-24	-2	8	5.14	< 0.001	0.002	*
Expected reward							
R Precentral Gyrus	39	-16	65	6.11	< 0.001	< 0.001	*
L Superior parietal gyrus	-18	-37	38	5.07	< 0.001	0.003	*
R Caudate	12	-1	20	4.79	< 0.001	0.012	*
Previous Trial Outcome							
L Superior Frontal Gyrus	-12	2	65	3.44	< 0.001	0.943	
R Superior Parietal Gyrus	21	-37	50	3.37	< 0.001	0.97	
L Superior Parietal Gyrus	-21	-43	38	3.29	< 0.001	0.985	
Previous Trial Win: Reward Mag							
Posterior Thalamic Radiation	-36	-52	2	4.29	< 0.001	0.095	
Previous Trial Win: Reward Mag							
Posterior Thalamic Radiation	-36	-52	2	4.29	< 0.001	0.095	
L Thalamus	-9	-7	17	3.66	< 0.001	0.576	
R Thalamus	6	-7	16	3.34	< 0.001	0.883	
Previous Trial Loss: RPE							
L Superior Frontal Gyrus	-12	2	65	3.53	< 0.001	0.904	
Previous Trial RPE given effort							
R Hippocampus	33	-31	7	3.54	< 0.001	0.882	
L Hippocampus	-24	-28	-4	3.14	0.001	0.998	

Last, we examined two specific outcome-related regressors, win (parametric modulators reward magnitude and reward magnitude given effort expenditure) and loss (with parametric modulators RPE and RPE given effort). We included these modulating terms in order to parse whether changes in BOLD signal were reflective of allocating effort. That is to say, did rewards feel more rewarding after working hard?

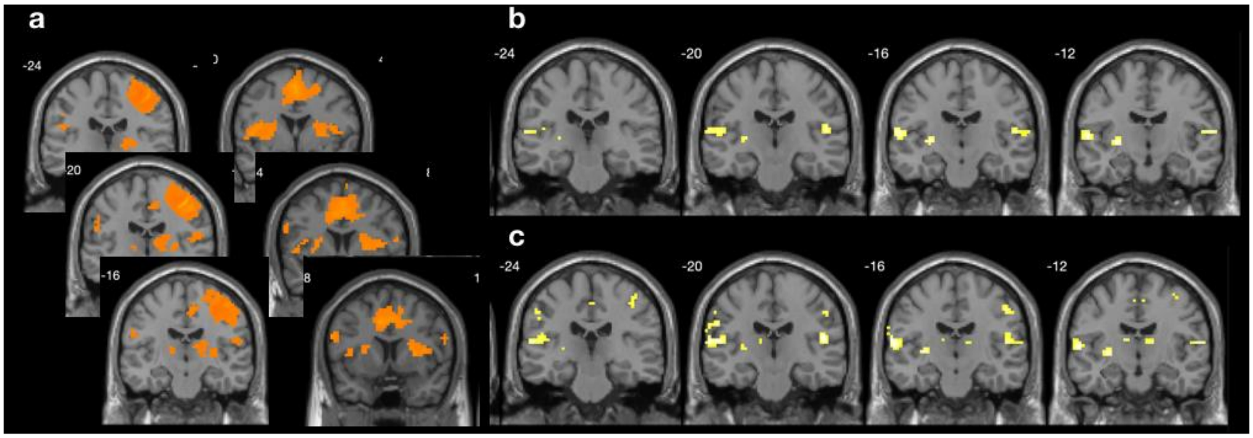

Figure 16: a, Win given effort expenditure. RPE (bottom row = given effort). **Reward magnitude after effort.** Win modulated by reward level after choosing high effort. Activity in primary motor cortex (PMC), supplementary motor area (SMA), bilateral putamen, bilateral thalamus, and dorsal anterior cingulate (dACC). RIGHT: **Reward prediction error.** Top panels: RPE, bottom panel, RPE given effort. RPEs of all trials seen most in insula. Losses after effort expenditure see additional activity in bilateral putamen, insula, thalamus, and bilateral somatosensory cortex. [add figure for RPE (both wins, losses)]

Table 10. Local maxima for group-level outcome-phase events (p-values adjusted for search-volume).

	MNI coordinates (peak voxel)			*z-stats*	P_{uncorr}	P_{fwe}	** survives correction*
	x	*y*	*z*				
Win Magnitude							
R Postcentral Gyrus	3	-25	56	3.79	< 0.001	0.533	
L Precentral Gyrus	15	2	20	2.67	0.003	1	
Win Mag. given effort							
R Postcentral Gyrus	39	-16	50	6.06	< 0.001	< 0.001	*
L Precentral Gyrus	-45	-1	11	4.92	< 0.001	0.007	*
R Thalamus	15	-19	5	4.88	< 0.001	0.008	*
L Supramarginal Gyrus	-42	-34	41	4.2	< 0.001	0.113	
RPE							
L Putamen	-30	-13	2	4.21	< 0.001	0.131	
Superior Temporal Gyrus	-57	-13	8	4.16	< 0.001	0.158	
L Precuneus	-3	-43	38	4.03	< 0.001	0.237	
R Lingual Gyrus	0	-67	5	4	< 0.001	0.258	
R Supramarginal Gyrus	51	-19	14	3.8	< 0.001	0.44	
RPE given effort							
Superior Temporal Gyrus	-60	-16	11	4.19	< 0.001	0.144	

L Putamen	-30	-13	3	4.09	< 0.001	0.198

2.4.3 Second Level: Analyses of individual difference measures

Covariates of interest including individual difference measures and maximum likelihood estimates computed from our final utility model were entered into second-level multiple regression analyses with event contrasts showing significant percent signal changes over the mean. We tested for regions showing both positive and negative correlations with covariates in order to assess the magnitude of correlation between task-specific activations and these subject-specific measures. These regressors and results are shown in Table 5 below.

Table 11: Subject-specific covariates with significant correlations with task activations. Discount factor K correlates significantly with computations of effort/reward tradeoffs. Significant percent signal changes in response to chosen cost regressors are positively correlated with discount factor K, while those associated with chosen reward magnitude are associated with BAS-drive levels.

	MNI coordinates (peak voxel)			*z-stats*	*P uncorr*	*Pfwe*	** survives correction*
	x	*y*	*z*				
K: Chosen Cost							
R Inferior frontal gyrus	60	26	14	5.49	< 0.001	< 0.001	*
R (post.) superior frontal gyrus	12	14	50	4.41	< 0.001	0.046	*
K: Chosen Mag							
R Middle frontal gyrus	45	26	35	6.3	< 0.001	< 0.001	*
R superior frontal gyrus PFC	6	42	50	4.47	< 0.001	0.045	*
BASd: Chosen utility							
R Insula	42	2	8	4.07	< 0.001	0.155	
R Thalamus	25	-25	2	3.71	< 0.001	0.420	
L Insula	-42	-4	8	3.53	< 0.001	0.519	

Subject-specific discount factors k, derived from behavioral data using Maximum Likelihood Estimation, were significantly associated with increased recruitment of neural resources at the decision event-specific activations. Specifically, variation in k was positively correlated with activity in right inferior frontal gyrus (IFG, $P_{FWE} < 0.001$, $Z_E = 5.49$) and the posterior portion of the right superior frontal gyrus, which houses components of the supplementary motor area (SMA, $P_{FWE} = 0.046$, $Z_E = 4.41$) in response to the chosen effort costs. The IFG, a region typically associated with semantic processing, was not active in action execution but was nevertheless modulated by the effort implied during the decision event.

Notably, k was negatively correlated with activity in response to chosen reward magnitude in right middle frontal gyrus ($P_{FWE} < 0.001$, $Z_E = 6.3$) and the right superior frontal gyrus in the prefrontal cortex (SFG, $P_{FWE} = 0.045$, $Z_E = 4.47$).

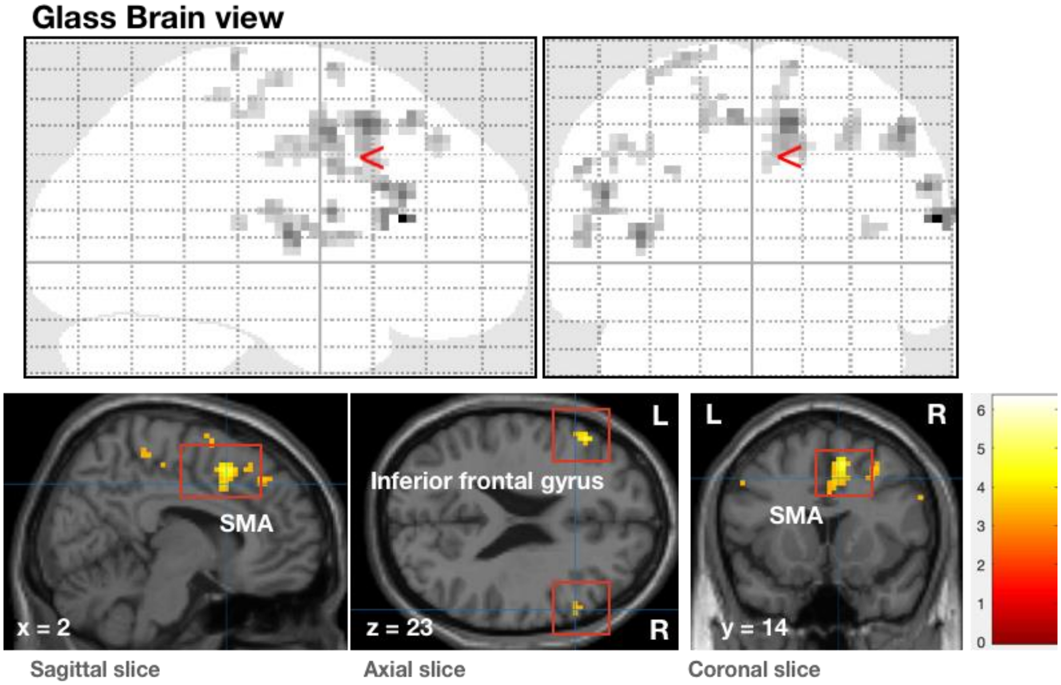

Figure 17: **Association between expected effort cost at decision-phase activativity, and subject-specific discount factor K ($P_{uncorr} < 0.001$). .** Parameter K is estimated from decision data and represents the weight that individuals place on cost, and effects the steepness of discounting. Those who are more sensitive to effort (i.e., discount more), have higher *k* values, while the opposite holds true of those less sensitive to effort costs. These associations are shown in supplementary motor area and inferior frontal gyrus.

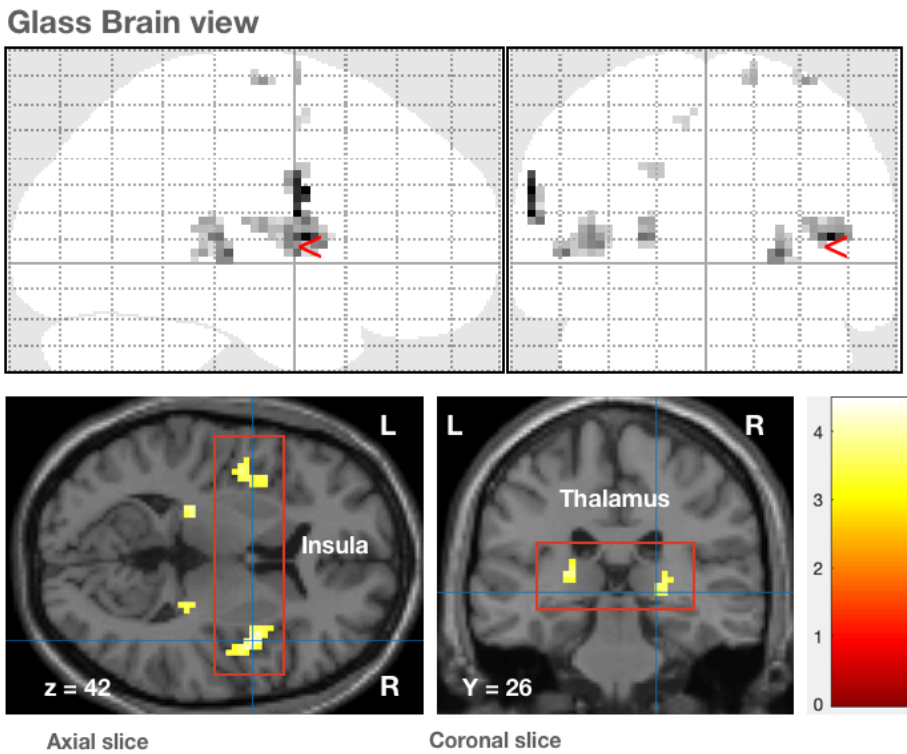

Figure 18: **Association between expected effort utility at decision-phase activations and BAS subscale Drive (P_{uncorr} < 0.001).** Increased recruitment of neural resources at the decision phase modulated by utility of the chosen lottery was observed in those higher on BAS-drive. Drive is defined as the persistent pursuit of desired goals, and is the BAS subscale most directly related to appetitive motivation. This effect was found in bilateral insula and thalamus, known to be involved in anticipation of effort production and action preparation (Kurniawan et al., 2013).

Of the BIS/BAS subscales, only BAS-drive (scale relating to the persistent pursuit of desired goals, the BAS subscale most directly related to appetitive motivation) showed a significant correlation magnitude with activity in relation to task elicited activity. Specifically, variation in BAS-drive was associated with activity in bilateral insula ($P_{uncorr} < 0.001$, R: $Z_E = 4.01$, L: 3.53) and right thalamus ($P_{uncorr} < 0.001$, $Z_E = 3.71$) in response to the utility of the chosen lottery option.

2.5 DISCUSSION

Overlapping task-related activity in regions of interest.

This analyses revealed an extensive but specific network of regions that showed activation correlating with chosen utility, chosen cost, experienced cost, and reward magnitude. The regions most clearly shown to be involved in computations underlying effort-reward tradeoffs were also recruited during the experience of effort exertion, and in some cases, the evaluation of outcomes (e.g., supplementary motor area, primary motor cortex, and insula). Broadly, this analysis yielded results with varying possible interpretations. In the first instance, it has been shown that the computations of effort evaluation and ensuing action execution recruit many of the same regions, including supplementary motor area (SMA, Bonnelle et al., 2016[30]), thalamus (Zénon et al., 2016[39]), dorsal anterior cingulate cortex (dACC, Hadland et al., 2003[40]; Assadi et al., 2009[41]), premotor regions (Massar et al., 2015[42]), and insula (Fauth-Bühler et al., 2010[43]). Additionally, the presence of reward signals in these motor-related regions is consistent with findings that reward signatures are not limited to regions classically associated with the mesocortical circuit, but in fact can be found throughout the brain (Vickery et al., 2011[44]).

Effort and reward monitoring.

The current findings suggest that a number of these overlapping regions show increased activity during the period of evaluating effort-reward tradeoffs. While causality remains outside of the scope of this particular analysis, many examinations of functional and structural connectivity of motor and reward systems have evidenced strong relationships between them (Hampton et al., 2017[45]). An analysis of the resting state data collected alongside this task may provide further support for the contribution of individual differences at the neural level to variation in behavioral apathy and subject-specific parameter estimates.

Notably, BOLD signal related to rewarding outcomes following effortful trials was sensitive to the amount of effort invested, while prediction error signals were not. Given the assumption that high effort is felt as more costly to the subject, the effect of this sensitivity is akin to creating a form of net valuation signal, such that relative to worse outcomes, the incremental value of better outcomes after high effort is attenuated and leads to smaller prediction errors compared with those after low effort.

Relationship to current individual difference measures.

The reported second level analyses provided only some evidence for a relationship between BIS/BAS and neural signals related to effort/reward tradeoffs: BAS-drive was associated with percent signal change in right somatosensory cortex to utility evaluation ($p_{\text{uncorr}}<0.001$, $p_{\text{FWE}}=0.378$, $Z_E=3.82$). However, second-level multiple regression analyses with individualized, subject-level discounting estimates k as a covariate revealed significant signal change both to expected cost and reward magnitude (see Table 5 for specific regions and statistics). Given that factor k offers an estimate for the weight that individuals place on the cost of effort within the utility function, a correlation with activity in SMA and IFG may suggest that these regions contribute to the weight individuals place on costs when evaluating effort/reward tradeoffs. This result further evidences the contribution of computational modeling and its ability to provide individualized parameter estimates from behavior that may be clearly connected to the neural activity underlying said behavior.

Where to go from here.

An initial primary focus of this examination was to determine the association between individual differences in stable traits to activity in reward-related mesocortical regions (e.g., major portions of the basal ganglia (BG), including striatum (Str), nucleus accumbens (NAcc), caudate, and putamen) during decisions involving physical effort. While some signal change in putamen and caudate were found in response to expected magnitude and reward prediction errors, there remains a noted lack of activity from deeper structures such as Str and NAcc. This may be attributable to any one of a number of factors: 1) these populations may have been more focused on costs and interoceptive signaling (i.e., effort and discomfort respectively) than rewarding outcomes, or 2) achieving high signal to noise ratio in subcortical regions such as striatum is notoriously difficult (Colizoli et al., 2020[46]), and while we specifically used 32-channel 3-T phased-array coil for this reason, we cannot rule out the possibility that signal dropout in subcortical regions contributed to the weak/null effects in these regions.

A final note regarding variation in data: an important difficulty found in analyzing these data was the tactile difference between the MR-compatible dynamometer, and the behavioral-only device. A small percentage of participants noted discomfort using the former device at higher levels of effort, and the somatosensory and insular activity found weighing on most event regressors may overwhelm much of the signal in various phases of the task.

The analysis described here contains a subset of the entire Berkeley cohort. As such, the methodologies described here can be extended to the larger group. However, given the size of the entire Berkeley cohort (nearing 150 subjects), it will be important to run group analyses not only among the entire cohort, but also among smaller subsections (e.g., different levels of various clinical and subclinical measures). Such measures can be integrated into both computational models and second-level GLM analyses in a similar manner. While, at this point in analysis, BIS/BAS did not seem to yield strong or easily interpretable results, the results from the more robust Miami cohort suggest that integrating other, more clinically specified measures may provide further, subject-specific indications of motivated approach/avoidance mechanisms at the biological level.

Chapter 3

Choosing when to help whom: a formal account of socially motivated effort expenditure

3.1 Introduction

Models of social decision-making historically rely on monetary costs to the self in computations of utility, or subjective value. However, many prosocial acts, in which an agent might incur costs to themselves in order to reward another, regularly entail a primary biological cost: energy. Prosocial acts are an evolutionarily adaptive and crucial component of human social coexistence, and in real world scenarios often require the allocation of effort. The way in which such metabolic costs affect the valuation of a decision that is other- versus self-rewarding is understudied. Additionally, it is unknown how an agent integrates information about the identity of another into such self-cost/other-benefit valuation processes. This chapter uses a novel effort-based social decision-making task and computational modeling to examine how person perception modulates willingness to engage in an effortful task. By parametrically modulating reward, effort, and the identity of the recipient, we show that the contributions of specific social information are interacting but computationally separable. Using economic models of utility discounting, we find that preferences for the allocation of effort for other-rewarding outcomes vary as a function of the recipient's perceived warmth and competence. Specifically, the warmth of a recipient has a constant effect on the probability of willingness to expend effort, while the competence of a recipient may modulate the steepness of the utility discounting curve.

3.1.1 Social perception and valuation

Socially interactive behaviors make up a critical set of tools within the mammalian behavioral arsenal, central to the survival of a species at both the individual and group levels. Within this set of behaviors, anywhere from the level of dyad to intragroup dynamics, prosocial acts are a key component of cohesion, wherein the success of others may ultimately abet one's own well-being (Fehr & Rockenbach, 2004[47]). Humans have developed many unique social behaviors in response to unique social challenges. Among these are template-like cognitive representations that allow for quick categorization and valuation between individuals and groups (Hutchison & Martin, 2015[48]; Freeman & Johnson, 2016[49]). In evaluating actions to take towards conspecifics, an agent must consider the costs and benefits to themselves, while also weighing the costs and benefits to another, that will lead to a desired outcome. Models of distributional preferences (Charness & Rabin, 2002[50]) have shown that agents integrate the consideration of another's needs into their own utility function when making decisions that may affect another. Such models are constrained, however, by two assumptions: (1) that social valuation occurs uniformly across all agents and the subjects of their evaluation, and (2) that all cost types in such utility functions can be represented equally by monetary costs. Here, we employ a computational approach that integrates dimensions of social perception and a primary biological cost, effort, to capture variation in the discounting of vicarious rewards by physical effort.

There has until recently been little focus on the ways in which individual components of the identity of another person may contribute to models that represent distributional preferences. Following proposals from social psychology that stereotypes can be organized along specific dimensions (Asch, 1945[51]), recent work suggests that the categorical information contained within stereotypes modulates social valuation, and thereby behavior, in social economic games (Jenkins et al., 2018[52]). Specifically, Jenkins and colleagues employ features of prominent framework of stereotype content that organizes person perception onto two orthogonal dimensions (Fiske et al., 2007[53]): (1) warmth, or the degree to which a person's perceived intentions are good, and (2) competence, or the degree to which a person has the ability to act on said intentions. The authors show that these core dimensions of social perception can be integrated into models of social valuation in order to meaningfully predict variation in unequal treatment. Typical of various studies modeling behavior in social economic games (Rangel, Camerer, & Montague, 2008[54]), the cost-benefit tradeoffs considered in this work are restricted to monetary costs to the self. Whether and how such findings translate to decisions that weigh effort costs against vicarious rewards is not known. The current work examines how these dimensions of social perception may integrate into decisions about one's motivation to benefit others.

3.1.2 Effort allocation and reward discounting

Motivation can be defined as a set of cost/benefit valuation processes that allow an organism to overcome an obstacle in order to reach a goal. First, an organism must compute the costs associated with the action that will lead to procuring that goal, and the value of the goal itself (Chong et al., 2016[1]). Examinations of the neural, cognitive, and behavioral signatures of motivated decision-making have increased in recent years, offering further, robust computational characterizations of how organisms approach, understand, and make effort-based decisions (Chang et al., 2019[55]; Lockwood et al., 2017[56]; Inzlicht et al., 2017[24]; Hamid et al., 2016[57]).

Systematic investigations into nonhuman animals and humans alike have put forth various instantiations of task designs that force decision-makers to undertake energetically costly, goal-directed behavior (see Chong and colleagues, 2016, for review[1]). Findings show, across studies and organisms, that decision-makers follow a robust pattern of behaviors resulting from specific cost-benefit computations, distinct from similar reward devaluation tendencies such as temporal discounting (Klein-

Flügge et al., 2015[28]; Hartmann et al., 2013[11]; Floresco et al., 2008[58]). These series of computations follow a concave curve, wherein a high-rewarding goal achievable through low-cost action is conferred a higher subjective value compared to a low-rewarding goal achievable only through high-cost action. This change in subjective utility as a function of reward and effort costs is known as effort discounting and follows a curvilinear utility function. In humans, when both effort costs and monetary rewards are for the self, this curve follows a sigmoidal shape, as shown by Klein-Flügge and colleagues. However, few studies have focused on effort allocation for vicarious rewards, and to our knowledge, none have investigated social identity as a potential factor in such decisions. Additionally, a direct comparison between the shape of social versus individual effort discounting curves has not yet been put forth.

3.1.3 Socially motivated effort expenditure

Our desire and ability to act in prosocial ways has played a crucial role in our evolutionary history, wherein fruitful interpersonal interaction forms a keystone to functioning society. It is believed that group social cohesion has an adaptive value for many social species, and the survival of such species relies at least in some part on prosociality (de Wall, 2008[59]). For that reason, acting in a prosocial manner is a natural reinforcer: to survive among conspecifics, social play between two agents should offer benefits to at least one party (Trezza et al., 2010[60]). Recent work suggests that humans are willing to engage in prosocial acts that benefit others, even when such acts come at a personal cost (Jenkins et al., 2018[52]; Hsu et al., 2008[61], Fehr & Schmidt, 1999[62]). Models described by these studies attempt to parse the interaction between an individual's concern for their own outcomes, and their concern for the outcomes of others (Blanco et al., 2011[63]; Fehr & Schmidt, 1999[62]). The majority of work examining prosocial economic behavior focus on monetary costs to self, while in everyday life, many prosocial interactions come at an energetic cost: effort. While laboratory experiments examining social decision-making account for distributional preferences over an anonymous other (Engelmann & Strobel, 2004[64]; Charness & Rabin, 2002[50]), in real-world scenarios in which one works hard for the benefit of

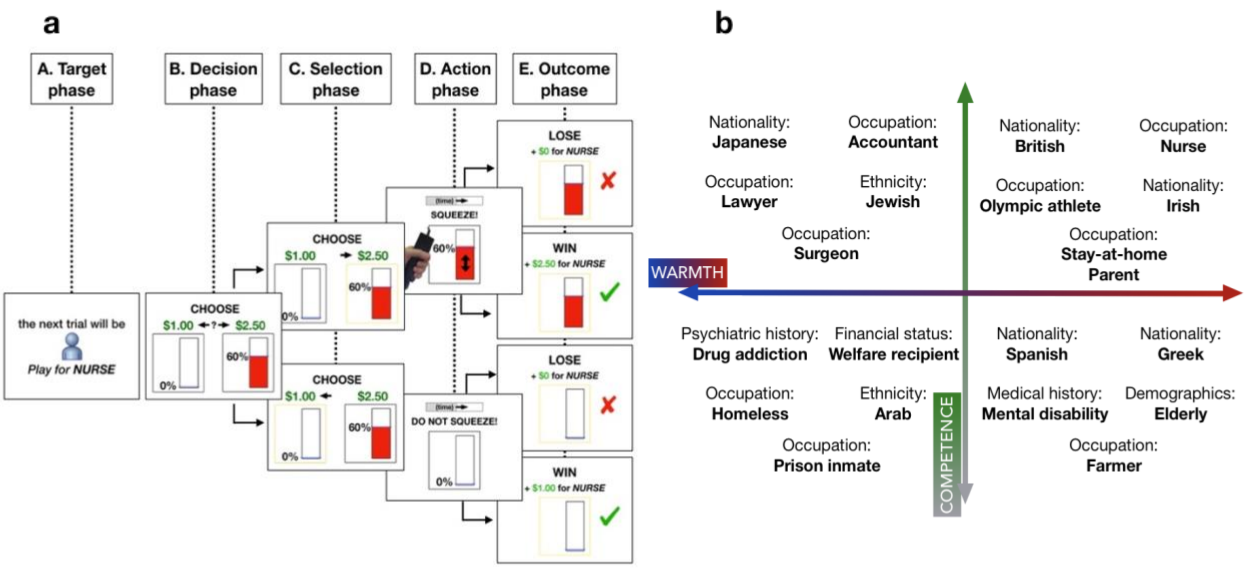

Figure 19 a, **Effort-based social decision-making lottery (seDM) task.** A. Trials begin in the target phase, wherein participants view recipient information (e.g., “On this trial, you will play for Nurse”). B. The decision phase displays the amount of money offered by either lottery option, and effort required (as % MVC) to play the higher stakes lottery. The low-rewarding lottery always requires 0% effort. Participants have up to 5s to select a lottery. C. The selection phase highlights the chosen lottery and prepares participant for the action phase. D. Participants either squeeze (following an effort lottery choice) or wait (following a no-effort lottery choice). Effort exertion was performed by squeezing a force dynamometer with dominant hand, which translated on the screen as a red bar gradually filling a rectangle. Either action phase lasts a total of 5 seconds. E. Participants are shown the outcome of the lottery, set at 50% probability of winning. Outcomes for both lotteries are probabilistic at 50%. b, heatmap of observed high effort decision proportions given cost and reward levels, averaged over all participant trials. Top left panel (high reward, low cost) shows highest average proportion of effort lotteries chosen (91%), while bottom right panel (low reward, high cost) shows the lowest average proportion of effort lotteries chosen (21%). b, twenty social groups selected to span the warmth and competence space.

another person, the identity of the beneficiary is often known. In the following experiment, we build on the idea that identity information in the form of stereotype content contributes to computations about resource allocation. Utilizing an effort-based laboratory game, we show that variation in the perception of different groups’ warmth and competence offers novel insight into how individuals are willing to accept direct, effort costs in order to benefit those groups.

3.2 Methods and Materials

Participants: The experiment was performed at the University of California at Berkeley. The research was approved by the Committee for Protection of Human Subjects at UC Berkeley. Participants provided written informed consent prior to participation. A total of fifty-two (52) participants were recruited from the university’s undergraduate and graduate student populations. Eligible study participants were between the ages of 18 and 28. Of the 52 experimental participants, seven were excluded due either to hardware challenges, failing attentional checks, or switching grip-force device between hands at some point in the experiment. The following analyses include a total of 45 participants (female = 26 (1 did not report), mean age = 22.9).

3.2.1 Design

Participants completed 3 blocks of 42 trials (126 total) comprised of 6 decisions for themselves, and 120 decisions for 20 targets (6 decisions per target). Each trial presented a choice between two possible lotteries (Fig. 1): one that required zero eff ort for a low reward ($1.00), and another that required a variable amount of effort (20, 30, 40, 60, 80%) for a variable reward ($1.50, 2, 3, 3.50, 5). Three breaks were included to prevent participant fatigue. All reward outcomes for both lotteries were probabilistic at 50%, to allow use to examine whether previous rewards had an effect on current decisions. For all subjects, regardless of choice, the interval between decision onset and outcome onset was fixed (5s) in order to avoid confounding effort discounting with temporal discounting.

3.2.2 Procedure

The primary function of this study was to examine whether humans compute effort-related self-cost/other-rewarding tradeoffs as a function of social perception. To examine this, we designed an effort-based social decision-making task based on rodent and human research (Hartmann et al., 2015[11]; Prévost et al., 2010[65]). In this paradigm, participants made decisions about whether or not to exert a specified level of effort (a personal cost) for a specified level of monetary reward (a reward for someone else) and executed the chosen option (effort or no effort). To precisely characterize effort output, we used five different levels of effort scaled as a percentage of each individual participant's maximal voluntary contraction (MVC). Participants decided between two possible lotteries: one that required 0% effort to play for a 50% chance of winning the specified reward ($1.00), and another that required a variable level of effort costs to play a lottery for a higher reward. Upon choosing the no-effort lottery, participants simply waited, and then played the lottery. Upon choosing the effort lottery, participants were required to expend the specified degree of effort by squeezing the handheld device before being able to play the lottery. Following an effort lottery choice, if participants did not squeeze hard enough or for enough time, they failed and moved on to the next trial. Otherwise, assuming a successful completion,

3.2.3 Targets

The crucial manipulation in this task relates to varying the recipient of the monetary reward. Participants made decisions to earn payments either for themselves (self-rewarding, 5% of trials), or for a designated recipient (other-rewarding, 95% of trials), wherein the recipient is defined as 1 of 20 possible targets. All targets followed

Jenkins and colleagues (2017), which describes these targets as spanning the dimensions of low to high warmth and competence. Twenty counterparts and tradeoffs between 5x5 cost/reward were chosen pseudo-randomly such that the possible target-effort-reward combinations were distributed equally across participants.

3.2.4 Hardware

In order to register responses, participants selected one of two possible options using a PC keyboard, and a handheld dynamometer (MP150, BIOPAC Systems, USA) to pay the effort costs required to play the lottery. Dynamometer force recordings were measured at 10Hz/ms.

3.2.5 Statistical analyses of behavioral data

All analyses of behavioral data were performed in R. Running various linear regressions, we examined the probability of choosing effort given cost, reward magnitude, warmth, competence, age, gender, and trial number to account for possible fatigue. To assess how previous trials might affect current decisions, we also included regressors for previous wins (binary scale), previous reward magnitudes (discrete scale), and reward prediction errors (RPE). Finally, we also ran regression models to examine interactions between all non-correlated, aforementioned regressors. Results from these regressions can be found in Table 1. All modeling analyses were performed on other-rewarding trials only. For all the analyses, the alpha level was set to $P < 0.05$ two-tailed.

3.3 Models & Model Selection.

3.3.1 Computational modeling

A central consideration in the design of the experiment was to understand whether social identity information contributed to effort-discounting in a constant (additive) or modulatory (multiplicative) manner (shown in FIG. 2). To preliminarily characterize these possible effects, hierarchical linear models were fit prior to computational models.

First, in order to best understand the general shape of the discounting curve that these types of decisions – tradeoffs between effort costs to self and other-rewarding outcomes – follow, we fit omnibus models of four discounting model types.

Table 12

k	*beta*	*llik*	*BIC*	*p*	*gamma*	*model*
4.881	0.423	2433.54	4891.90	79.390	NA	sigmoidal

0.023	0.921	2419.18	4854.91	NA	NA		hyperbolic
4.07E-04	0.514	2393.75	4804.05	NA	NA		parabolic
3.83E-05	0.494	2389.79	4796.13	NA		2.547	power

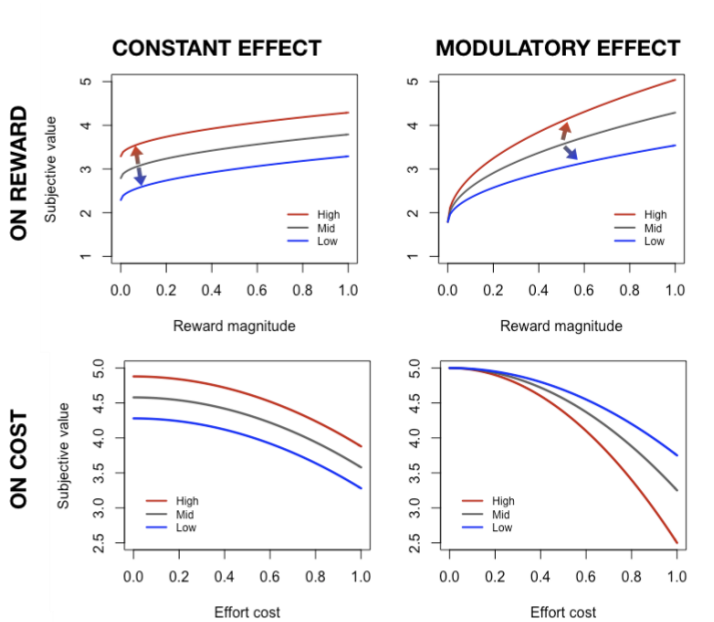

Figure 20: Illustration of changes to subjective value by the integration of warmth and competence information. Grey curves represent a baseline model of vicarious reward discounting by effort. Red (blue) curves indicate that the inclusion of a positive (negative) weight parameter at the additive term (constant effect) leads to an additive bonus (penalty) at the intercept (left panels). The inclusion of a positive (negative) weight parameter on the multiplicative term leads to a modulatory change in the slope of discounting (right panels).

Each model contained idiosyncratic parameters characterizing the degree k to which rewarding outcomes were discounted, or devalued, by effort, warmth and competence weight parameters (b, μ, κ), and the steepness of the power decrease γ. Two main features were varied to build the model space: first, we adjusted the model's functional form to modify the shape of the discounting curves for four possible types of discounting: parabolic, power (free parabolic), hyperbolic, and sigmoidal. Second, we adjusted the position of the social perception variables (w_j, c_j) and their weights (b, μ, κ), to best describe the effect of warmth and competence on participant decisions.

A key innovation of this study is a computational model that reliably estimates, from participant choices, weight parameters that modulate effort-based decision utilities as a function of the content of stereotype information. This is to say, observed decisions per target within an omnibus model (using all participants' choices) generated specific warmth and competence weight parameters. We accounted for potential overfitting by reporting k-fold cross-validated log-likelihoods, with the same testing-training splits for model comparison.

These weights were entered into simulation models and in a k-fold LOO cross validation (iterations = 1500), the winning model performed at R=.89 ($p < 0.001$). Out-of-sample predictions were made by holding out some proportion of the targets, fitting the model on the remaining targets, and then using this fitted model to predict the response rates of the held-out targets.

We compared several models to formally characterize the contributions of warmth and competence to the discounting of vicarious rewards by effort costs.

Baseline model (flexible power function)

The baseline model assumes no contribution of stereotype content to discounting. We fit the model parameters per participant and per target using maximum-likelihood estimation. Choices under the baseline model are characterized by an increasing devaluation of reward with growing effort costs, according to a power function:

$$U = \beta_h + M^{\alpha} - kC^{\gamma}$$

where utility U denotes the subjective value, M denotes the reward magnitude, and C the cost (in this case, physical effort). Following previous reports (CITE Klein-Flugge, following Wu et al.) that physical effort discounts subjective value as opposed to pure monetary units (here, dollars and cents), we have set the utility parameter α on reward magnitude M to $\alpha = 0.8$. This value represents an estimate of the average utility parameter of healthy humans (Abdellaoui et al., 2007[27]). Three parameters, estimated by the data, describe the steepness (k) and the curvature (power parameter, g) of the utility devaluation. Additionally, baseline (null) model assumes that social identity information (warmth, competence) does not contribute to participant choices, but estimates constant β_h (intercept) to capture a general preference for choosing to expend effort.

Additive

The additive-effect only model captures possible constant changes to decision utility, wherein the baseline probability of choosing to work for a target j's reward either increases or decreases linearly as a function of (warmth, competence). Warmth and competence parameters (w_j, c_j) and weight b denote the given warmth and competence of target j (quantitative ratings based on Jenkins et al. 2017[52]), and the weight placed on that dimension, respectively. The additive model assumes that warmth and competence directly contribute only to changes at the intercept, and have no effect on the steepness of discounting. Choices under the additive model are governed by the utility function:

$$U_{add} = \beta_h + b(w_j, c_j) + M^{\alpha} - kC^{\gamma}$$

Multiplicative: k only

The multiplicative-k model sees the weights participants place on counterpart rewards as a linear function of warmth and competence modulating the slope of discounting parameter k. The multiplicative model assumes that the perception of w,c does not contribute to changes at the intercept, but rather they govern the magnitude of (w_j, c_j) on the rate of effort discounting via parameter κ. Choices under the multiplicative-k model are governed by the utility function:

$$U_{mult.k} = \beta_h + M^{\alpha} - (k+\kappa(w_j, c_j))C^{\gamma}$$

Multiplicative: reward only

The multiplicative-reward model sees the weights participants place on counterpart payoffs as a linear function of warmth and competence modulating the reward magnitude. The multiplicative model assumes that warmth and competence modulate utility weights $\mu(w_j, c_j))$ on the reward magnitude M. Choices under the multiplicative-reward model are governed by the utility function:

$$U_{mult.r} = \beta_h + (1+\mu(w_j, c_j))M^{\alpha} - kC^{\gamma}$$

Multiplicative-additive (multadd): Full model

The multiplicative-additive model (multadd) is a combination of the effects of the above models. In this instantiation, additive weight parameter b modulates (w_j, c_j) at the intercept, while parameter weight κ modulates variables (w_j, c_j) at slope parameter k. The additive effect varies the initial probability of choosing effort as a function of (w_j, c_j), regardless of slope, while the multiplicative effect varies the steepness of discounting as a function of cost, and warmth and competence.

$$U_{mult.add} = \beta_h + b(w_j, c_j) + M^{\alpha} - (k+\kappa(w_j, c_j))C^{\gamma}$$

Connecting utility to choices

Participant decisions, parametrically modulated by discrete units of effort costs and reward, were binary in nature (no effort = 0, effort = 1). Thus, we calibrated each subjective value function to behavior using the softmax function with inverse temperature parameter (lambda). The probability of a participant choosing the effortful option on a specific trial is given by:

$$P(effort)_{t,j} = \frac{1}{1 + exp\ (-\lambda(U(l_{t,j}) - U(h_{t,j}))}$$

where temperature *l* defines the steepness of the softmax function, and thus how sensitive the choice probability is to $U(l_{t,j}) - U(h_{t,j})$, or the difference in subjective value between the no-effort lottery on trial t given target j, $l_{t,j}$, and effort lottery on trial t, $h_{t,j}$. Bayesian Parameters were estimated using Maximum Likelihood Estimation (MLE), by maximizing the log-likelihood of all participant decisions, per model. Twenty random starting points were defined within the optimization function to capture the global minimum. Standard errors and p-values for the parameters were obtained from the inverse of the Fisher Information Matrix.

3.4 Results

3.4.1 Discounting social rewards by physical effort.

Following previous examinations of individual and social reward discounting by physical effort (Lockwood et al., 2017[56]; Bonnelle et al., 2014[30]), we first hypothesized that increases in effort costs would lead to decreases in the utility of rewarding outcomes.

Initial model-free analyses confirm that our task elicits the expected behavior, wherein individuals discount rewarding outcomes by effort costs.. While on average most participants elected to choose the effortful lottery more often (59%) when playing a self-rewarding trial, other-rewarding trials still followed the same pattern of decreasing probability of effort expenditure as a function of increasing costs, and increasing probability of effort tracking increasing rewards. With regard to the proportion of effortful lotteries chosen, we first found using analyses of variance (ANOVAs) that willingness to expend effort to another person's benefit differs as a function of effort ($F_{4,695} = 404.48$, $P<0.001$) and cost ($F_{4,695} = 135.48$, $P<0.001$). Thus, as expected, when a lottery increased in required effort, participants were less willing to select it; concordantly, as a lottery increased in offered reward, participants' willingness to choose that lottery increased.

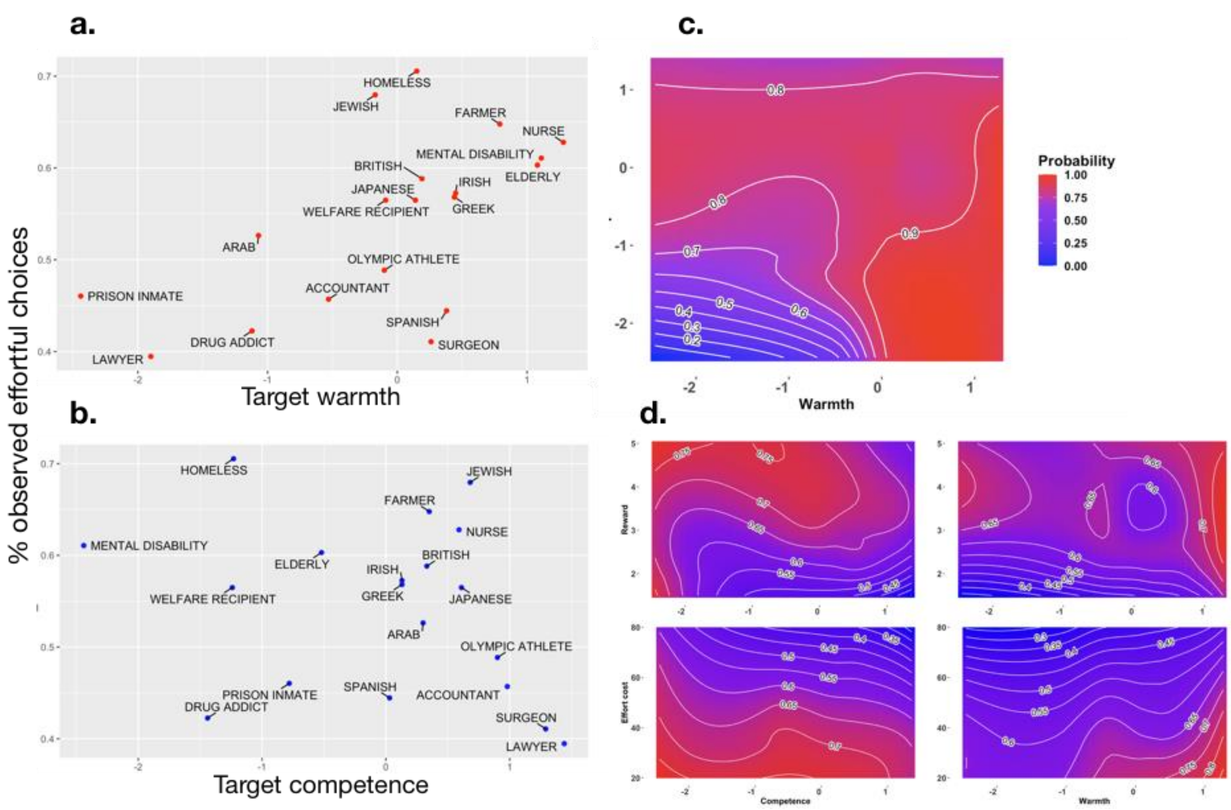

Figure 21. **Effect of target warmth and competence on probability of choosing effort for vicarious rewards.** (a) and (b) show relationship between specific targets' warmth and competence and proportion of effortful choices made (warmth r = 0.58, competence r = -0.12). (c) shows the probability of effortful choices as a function of warmth (x-axis) and competence (y-axis). (d) shows the same probabilities but includes both effort (bottom panels) and reward values (top panels).

Second, we hypothesized that the probability of choosing effort for another's reward would vary as a function of that target's perceived warmth and competence. More specifically, we expected that increases in a target's perceived warmth, or positive intentions towards others, would lead to monotonic increases in effortful choices, while increases in a target's perceived competence, or ability to act on those intentions, would see a reduction in effortful choices. Using mixed logistic regressions that integrated social perception dimensions, we found a main effect of warmth ($F_{4,695} = 39.73$, $P<0.001$), but not competence ($F_{4,695} = 6.79$, $P=0.86$). However, there was a significant interaction of Competence × Cost ($F_{4,695} = 10.55$, $P<0.002$). This initial model-free analysis suggests that while warmth confers a constant effect on willingness to expend effort, the perception of cost may be specifically modulated by the competence of a counterpart.

Finally, we investigated three ways in which a target's group membership could integrate into the computation of decision outcomes: (1) additively, wherein stereotype perception contributes to the baseline probability of choosing effort independently of effort discounting, i.e., regardless of reward or cost, (2) multiplicatively on reward (mult-rew), wherein the contribution of social perception directly modulates the magnitude of a target's reward, and (3) multiplicatively on k (mult-k), wherein the contribution of social perception directly modulates the steepness of reward discounting. Under an additive model, a target's warmth and competence have a constant, linear effect on the utility of a choice. Under a mult-rew model, warmth and competence modulate the utility weight participants place on the reward magnitude of the chosen lottery. Under the mult-k model, warmth and competence modulate the utility weight on discount factor k, or the degree to which a reward is devalued by effort costs.

We found that stereotyping contributes to social valuation during effort discounting in specific but dissociable ways. While the null model successfully predicts other-rewarding effort-based decisions ($R^2 = 0.59$), the best-fit model robustly predicted behavior ($R^2 = 0.7$, $p < 1.3e\text{-}07$) towards targets. Under this model, a multiplicative-additive instantiation with a warmth term on the intercept and a competence term modulating k, target warmth has a constant effect on the intercept, corresponding to preliminary results from logistic regressions. Additionally, target competence modulates discount factor k, suggesting that increases in perceived competence lead to an increasingly steeper devaluation of vicarious rewards by effort.

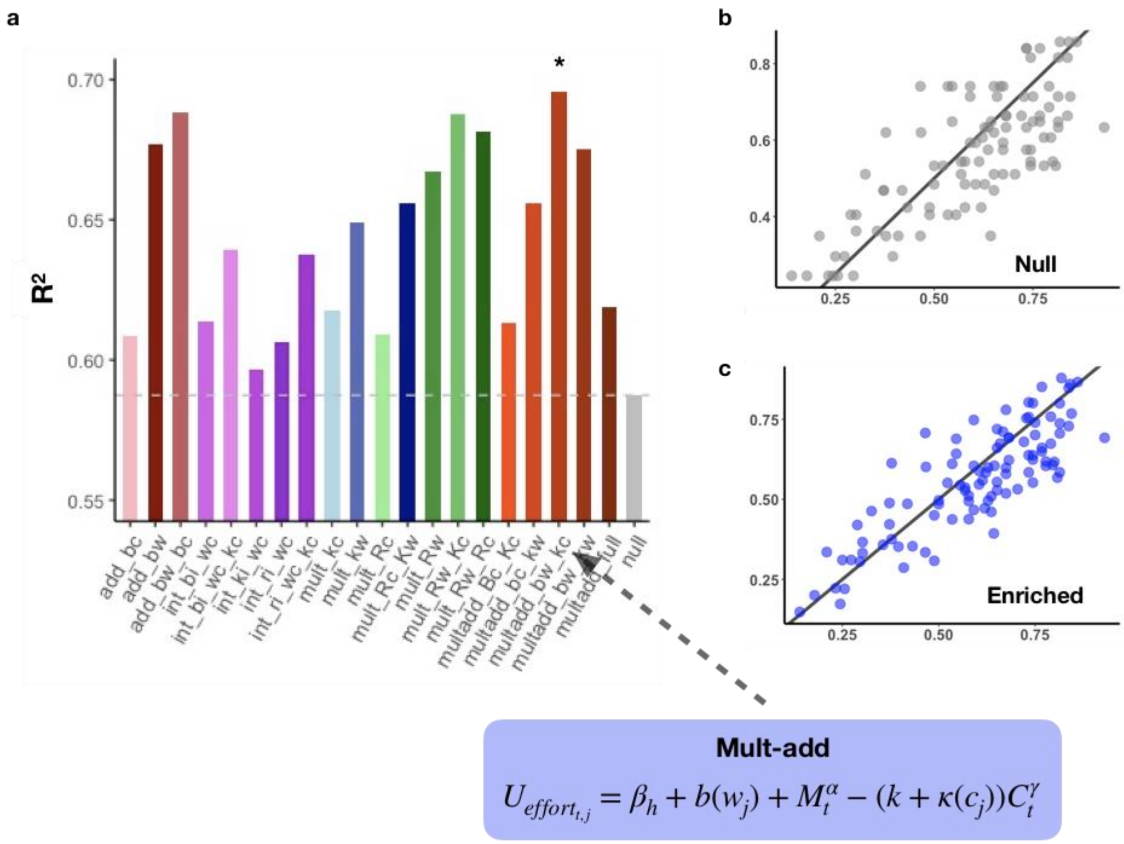

***Figure 22*. Model fits**. a, Model comparison results, with model number shown on x-axis and R^2 values from each model on y-axis. Models were additive (1-3), interactive (4-8), multiplicative (9-15), or multiplicative-additive (16-20); model 21 is the null model. b, scatterplots showing correlations between observed and predicted effortful choice probabilities: top panel, null model R = .75; bottom panel, enriched model R = .84. c, functional form of the winning utility function: where the utility of an effortful choice on trial t given target j is equal to the effort bias parameter β plus the warmth of target j modulated by weight parameter b, reward utility parameter M^{α} minus discount factor k plus competence of target j modulated by weight parameter κ, together weighing on cost C and curvature parameter γ.

3. 5 Discussion

The present study examined how stereotype content may influence people's motivations to benefit others when expending their own metabolic resources. When faced with the choice of whether to accept a cost to the self in order for another person's monetary gain, subjects reliably integrated social information into the decision computation. The perceived valence of the counterpart's intentions, operationalized here as the warmth dimension, contributes an additive bonus (penalty) to the utility of a decision involving effort. This is to say, the more positive (negative) a receiver's stereotypic intentions are perceived to be, the more (less) likely a person is to expend effort for the receiver's benefit. On the other hand, a counterpart's competence interacted with effort costs in such a way that the weight on cost was modulated by an amount proportional to competence in the computation of the subjective value of an effort-based choice. This is to say, as the more (less) a receiver's is perceived as able to act on their stereotypic intentions for their own benefit, the less (more) likely a person is to expend effort for the receiver's benefit. Taken together, these results confirm that while prosocial motivation may be an integral component of human social life, societally crystalized notions about the traits and intentions of different social groups have an observable and quantifiable effect on the degree to which humans are motivated to accept metabolic costs in order to help one another.

We showed that perceived warmth and competence influence effort-based behavior by amplifying or diminishing 1) one's baseline willingness to accept costs on another's behalf and 2) one's sensitivity to the cost of effort as those costs rise. It is possible that an increased willingness to benefit those whom we consider warmer returns a higher value of vicarious reward than those whom we consider less warm. Additionally, the observation that competence indexes a counterpart's ability to help themselves may explain subjects' increased sensitivity to effort at higher costs when counterparts are seen as more competent. While competence may not affect the reward one feels by helping a counterpart, it may weigh more heavily on the perception of cost in a manner consistent with the notion that "they can do it themselves."

This result contributes to our understanding of the ways in which humans engage with cost-benefit tradeoffs in social contexts. Lockwood and colleagues found that individuals are broadly less likely to accept effort costs for others' rewards, or to avoid others' losses, than they are to accept costs to gain their own rewards (avoid losses). Additionally, we can consider metabolic energy to be a type of resource that varies between people and can be distributed. is consistent with previous findings that warmth and competence contribute to behavior in social economic games in interacting but separable ways by modulating concerns for fiscal equity (Jenkins), wherein individuals are averse to receiving more than warm counterparts and to receiving less than competent counterparts. Future work may offer insight into whether and how stereotype information contributes to the utility of working to avoid losses for a specific counterpart in the same manner as working to endow gains for a counterpart.

These findings may be consistent with studies of social interaction in model organisms, which show that phasic dopamine release in the nucleus accumbens (NAcc) increases in response to pro-social behavior (Wu et al., 2011[29]). Additional work suggests that elevated dopamine release in NAcc has been associated with increased motivational vigor in animal models of effort-related reward seeking behavior (Ko & Wanat, 2016[66]). Whether increased motivation lead to increases in pro-social effort-related behavior, or vice versa, remains unknown. Future work may add to this understanding by investigating a plausible mechanistic overlap between regions involved in these socially motivated effort-based decisions.

Chapter 3 Supplementary materials.

Supplementary Table 10: Logistic mixed effects regression results for warmth and competence data.

		Null model			**Competence -only**			**Warmth-only**			**Full model**	
Predictors	*Log-odds*	*Conf. Int (95%)*	*P-Value*	*Log-odds*	*Conf. Int (95%)*	*P-Value*	*Log-odds*	*Conf. Int (95%)*	*P-Value*	*Log-odds*	*Conf. Int (95%)*	*P-Value*
Intercept	0.70 **	0.22 – 1.17	**0.004**	0.72 **	0.24 – 1.21	**0.003**	0.73 **	0.25 – 1.20	**0.003**	0.73 **	0.25 – 1.22	**0.003**
Effort Cost	-0.00 ***	-0.00 – -0.00	**<0.001**	-0.00 ***	-0.00 – -0.00	**<0.001**	-0.00 ***	-0.00 – -0.00	**<0.001**	-0.00 ***	-0.00 – -0.00	**<0.001**
Reward Magnitude	0.52 ***	0.43 – 0.61	**<0.001**	0.53 ***	0.44 – 0.62	**<0.001**	0.51 ***	0.42 – 0.60	**<0.001**	0.53 ***	0.44 – 0.62	**<0.001**
Trial Number	0.00	-0.00 – 0.00	0.125	0.00	-0.00 – 0.00	0.098	0.00	-0.00 – 0.00	0.116	0.00	-0.00 – 0.00	0.092
Warmth				0.43 ***	0.21 – 0.64	**<0.001**				0.50 ***	0.29 – 0.72	**<0.001**
Warmth:Effort cost				0.00	-0.00 – 0.00	0.458				0.00	-0.00 – 0.00	0.062
Warmth: Reward mag				0.03	-0.06 – 0.12	0.555				0.00	-0.09 – 0.09	0.985
Warmth:Trial				0.00	-0.00 – 0.00	0.107				0.00	-0.00 – 0.00	0.107
Competence							0.02	-0.19 – 0.23	0.860	0.01	-0.20 – 0.22	0.933
Competence: Effort cost							-0.00 **	-0.00 – -0.00	**0.002**	-0.00 *	-0.00 – -0.00	**0.011**
Competence: Reward mag							-0.01	-0.11 – 0.08	0.776	-0.02	-0.11 – 0.07	0.658
Competence: Trial							0.00	-0.00 – 0.00	0.996	0.00	-0.00 – 0.00	0.814
Random Effects												
σ^2	3.29	3.29	3.29	3.29								
τ_{00}	2.13 SID	2.24 SID	2.16 SID	2.26 SID								
ICC	0.39	0.40	0.40	0.41								
N	45 SID	45 SID	45 SID	45 SID								
Observations	4695	4695	4695	4695								
Marginal R^2 / Conditional R^2	0.135 / 0.475	0.153 / 0.496	0.141 / 0.482	0.157 / 0.500								
* $p<0.05$ ** $p<0.01$ ***												

Computational modeling results.

Supplementary Table 11. Model fits, basic sigmoidal model.

Parameters	*Tests of model-fit*	*parameter description*	*basic model*
β_h		constant for choosing effortful option	0
k		discount factor	0.0005
γ		free exponential on cost in power function	1.8779
λ		inverse temperature parameter for softmax decision function	0.7165
	nLL	negative log-likelihood	3269.779
	BIC	Bayesian information criterion	6590.877
	χ^2	test of independence	2.9139
	R^2	variance explained	0.5886

WARMTH-ONLY MODELS

Supplementary Table 12: Model fits, warmth-only models.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*Mult K*	*multadd*
β_h		constant for choosing effortful option	0	0.2423	0.1576
β_w		effect of warmth on beta	0.3369	-	0.3925
k		discount factor	0.0003	0.0049	0.0025
k_w		effect of warmth on k	-	-0.001	0.0002
γ		free exponential on cost in power function	1.9894	1.3922	1.5364
λ		inverse temperature parameter for softmax decision function	0.7074	0.7088	0.7162
	nLL	negative log-likelihood	3239.56	3256.26	3241.35
	BIC	Bayesian information criterion	6504.78	6538.17	6508.35
	χ^2	Test of independence	2.3936	2.6859	2.4176
	R^2	variance explained	0.6757	0.6411	0.6752
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (1) p < 3.9e-08	χ^2 (1) p < 2.4e-04	χ^2 (2) p < 6.7e-07
	LRT v add	likelihood ratio test of model vs additive			χ^2 (1) p < 2e-05
	LRT v mult	likelihood ratio test of model vs multiplicative			χ^2 (1) p < 1.0e-04

COMPETENCE ONLY MODELS

Supplementary Table 13: Model fits: competence-only models.

Parameters	*Tests of model-fit*	*parameter description*	*additive*	*multiplicative*	*multadd*
β_h		constant for choosing effortful option	0	0.1611	0.2051
β_c		effect of competence on beta	-0.1887	-	-0.1757
k		discount factor	0.0005	0.0031	0.0039
k_c		effect of competence on k	-	0.0004	0
γ		free exponential on cost in power function	1.8985	1.4928	1.4406
λ		inverse temperature parameter for softmax decision function	0.7203	0.7208	0.7217
	nLL	negative log-likelihood	3259.53	3263.78	3261.4
	BIC	Bayesian information criterion	6544.71	6553.23	6548.47
	χ^2	Degrees of freedom	2.6929	2.683	2.7012
	R^2	variance explained	0.609	0.6178	0.6135
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (1) p < 1.4e-03	χ^2 (1) p < 1.4e-02	χ^2 (2) p < 1.5e-02
	LRT v add	likelihood ratio test of model vs additive			χ^2 (1) p = 1
	LRT v mult	likelihood ratio test of model vs multiplicative			χ^2 (1) p = 0.12

WARMTH + COMPETENCE MODELS: ON THE SAME TERM

Supplementary Table 14: Model fits: warmth and competence on the same term.

Parameters	*Tests of model-fit*	*parameter description*	*add. interaction*	*R mult. interaction*	*K mult. interaction*	*full (mult-add interaction)*
β_h		constant for choosing effortful option	0	0.0049	0.3656	0.3013
β_w		effect of warmth on beta	0.3093	-	-	0.4138
β_c		effect of competence on beta	-0.1561	-	-	-0.1747
R_w		effect of warmth on reward	-	0.1367	-0.0018	0.0006
R_c		effect of competence on reward	-	-0.06	0.0009	0
k		discount factor	0.0004	0.0005	0.0109	0.0064
k_w		effect of warmth on k	-	-	-0.0018	0.0006
k_c		effect of competence on k	-	-	0.0009	0
γ		free exponential on cost in power function	1.9652	1.9003	1.2231	1.3341
λ		inverse temperature parameter for softmax decision function	0.7164	0.7285	0.7116	0.7162
	-LL	negative log-likelihood	3232.95	3236.58	3253.04	3235.91
	BIC	Bayesian information criterion	6491.55	6498.81	6531.73	6497.5
	χ^2	test of independence	2.2606	5.567	2.4751	3.0976
	R^2	variance explained	0.6873	0.1979	0.6611	0.5706
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2) $p < 1.0e\text{-}08$	χ^2 (2) $p < 2.3e\text{-}04$	χ^2 (2) $p < 2.3e\text{-}04$	χ^2 (4) $p < 8.0e\text{-}07$
	LRT v add int.	likelihood ratio test of model vs additive int.				χ^2 (2) $p = 1$
	LRT v K mult int	likelihood ratio test of model vs K multiplicative int.				χ^2 (2) $p < 2.0e\text{-}04$

WARMTH + COMPETENCE MODELS: ON SEPARATE TERMS

Supplementary Table 15: Model fits: warmth and competence on separate terms.

Parameters	*Tests of model-fit*	*parameter description*	*add. warmth, mult. competence*	*mult. warmth, add. competence*	*full*
β_h		constant for choosing effortful option	0.2084	0.2091	0.1358
β_w		effect of warmth on beta	0.3094	-	0.2367
β_c		effect of competence on beta	-	-0.1592	-
R_w		effect of warmth on reward	-	-	0.0341
k		discount factor	0.0035	0.0039	0.002
k_w		effect of warmth on k	-	-0.0007	-
k_c		effect of competence on k	0.0003	-	0.0002
γ		free exponential on cost in power function	1.4641	1.4464	1.5859
λ		inverse temperature parameter for softmax decision function	0.7133	0.7152	0.718
	nLL	negative log-likelihood	3238.6	3249.23	3237.58
	BIC	Bayesian information criterion	6502.86	6524.12	6500.82
	χ^2	test of independence	2.2383	2.5	2.242
	R^2	variance explained	0.6953	0.6537	0.6937
	LRT v basic	likelihood ratio test of model vs basic	χ^2 (2) p < 1.7e-07	χ^2 (2) p < 2.4e-04	χ^2 (4) p < 4.7e-07
	LRT v add w, mult c	likelihood ratio test of model vs additive			χ^2 (1) p < 0.31
	LRT v mult w, add c	likelihood ratio test of model vs multiplicative			χ^2 (1) p < 6.4e-04

Supplementary Table 16: Model fit metrics of all enriched and basic models.

Model	npar	llik	BIC	chisq	rsq	pv.chisq
Additive: competence	9	3260	6545	2.693	0.609	1.365E-03
Additive: warmth	9	3240	6505	2.394	0.6757	3.859E-08
Additive: w, c	10	3233	6492	2.261	0.6873	1.003E-08
Multiplicative k: comp.	9	3264	6553	2.683	0.6178	1.434E-02
Multiplicative k: warmth	9	3256	6538	2.686	0.6411	2.355E-04
Multiplicative k: w, c	10	3253	6532	2.475	0.6611	2.313E-04
Multiplicative R, K: c, c	10	3252	6529	2.507	0.6559	1.318E-04
Multiplicative R: comp.	9	3262	6549	5.921	0.1441	4.029E-03
Multiplicative R, K: w, c	10	3240	6505	2.264	0.6879	3.062E-07
Multiplicative R: warmth	9	3242	6509	5.613	0.1903	1.268E-07
Multiplicative R: w, c	10	3237	6499	5.567	0.1979	6.179E-08
Mult-add: Bc, Kc	10	3261	6548	2.701	0.6135	1.517E-02
Mult-add: Bc, Kw	10	3249	6524	2.5	0.6537	3.45E-05
Mult-add: Bw, Kc	10	3239	6503	2.238	0.6953	1.697E-07
Mult-add: Bw, Kw	10	3241	6508	2.418	0.6752	6.691E-07
Mult-add: Bw, Rw, Kc	11	3238	6501	2.242	0.6938	4.749E-07
Mult-add: Bw, Rw	10	3240	6505	2.445	0.6718	2.754E-07
Mult-add: Bw, Bc, Kc	11	3235	6495	2.309	0.6864	1.178E-07
Multadd: Bwc, Kwc	12	3236	6497	3.098	0.5706	7.959E-07
null	8	3270	6591	2.914	0.5887	1E+00

Concluding Remarks

Motivation and its instantiation in choice behavior varies profoundly across individuals and contexts. The decision to allocate energy towards a specific goal will depend upon the computation of energetic costs (effort), the raw value of the goal (reward), the interaction of the two (utility), and the integration of many possible, often invisible, variables. Such variables may include whether the cost is mental versus physical, whether the goal is primary (e.g., food) or secondary (e.g., monetary), whom it rewards (self- versus other-rewarding), and a host of state (mood) or trait (personality, clinical diagnoses) measures that may weigh on various terms within the utility function. In this dissertation, I have discussed a broad approach towards integrating some of these variables into effort-based utility discounting functions.

The first chapter of this thesis discusses the challenges posed by current methods of relating individual differences to variation in motivated decision behavior. In an attempt to formalize this relationship, we offer a computational method of integrating stable, trait-level individual differences into a putative cognitive model of effort-based decision-making. The large cohort of participants offers a rich dataset from which we are able to infer robust findings related to the contribution of individual differences to effort-based choices.

The second chapter examines whether and how this approach to behavioral choice modeling may relate to task-specific neural activity during effort-related choices. Increasing interest in the neural mechanisms supporting effort/reward computations – including both cognitive and physical effort, and some studies specifically comparing the two – has added considerably to our understanding of these mechanisms. However, the precise involvement of both clinical and sub-clinical individual differences in these computations are still not well understood. The methods outlined here, while not definitive with respect to clinical application, offer a framework for comparing the possible contributions of diagnostic measures with free-parameter model estimates derived solely from behavior. This manner of computational phenotyping will remain a useful tool at the confluence of clinical and basic research.

Finally, the third chapter explores the ways in which humans compute tradeoffs between effort costs to self in order to make monetary gains for others, and how stereotypes regarding the designated other may contribute to variation in choices. While self-rewarding choices seem to follow a specific pattern of discounting, the functional form of effort-related utility equations are distinct. Additionally, these computations integrate information about the identity of the person for whom one is accepting

This approach to studying only a small set of the many aspects of human motivation provides a basis for numerous directions for future research. An integration of clinical diagnostic measures into computational models of effort discounting may offer novel, phenotypic parameter estimates both for behavioral prediction, but also for the aberrant cognitive and neural mechanisms underlying disordered motivation. Additionally, does participant social group membership and how its markers interact with other-rewarding effort-based decisions may help to mitigate the effects of social group information on social valuation? Following these lines of questioning may further our understanding of motivation, its interactions with interindividual differences, and implications in everyday personal and social life.

References

1 Chong, T. T. J., Bonnelle, V., & Husain, M. (2016). Quantifying motivation with effort-based decision-making paradigms in health and disease. In Progress in Brain Research (Vol. 229, pp. 71–100). https://doi.org/10.1016/bs.pbr.2016.05.002
2 Treadway, M. T., Buckholtz, J. W., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., Baldwin, R. M., Schwartzman, A. N., Kessler, R. M., & Zald, D. H. (2012). Dopaminergic Mechanisms of Individual Differences in Human Effort-Based Decision-Making. Journal of Neuroscience, 32(18), 6170–6176. https://doi.org/10.1523/JNEUROSCI.6459-11.2012
https://doi.org/10.1016/bs.pbr.2016.05.002
3 Holt, D. D., Green, L., & Myerson, J. (2003). Is discounting impulsive?: Evidence from temporal and probability discounting in gambling and non-gambling college students. Behavioural processes, 64(3), 355-367
4 Salamone, J. D., & Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. Neuron, 76(3), 470-485.
5 Chong, T. T. J., Apps, M., Giehl, K., Sillence, A., Grima, L. L., & Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. PLoS biology, 15(2), e1002598.
6 Patzelt, E. H., Hartley, C. A., & Gershman, S. J. (2018). Computational Phenotyping: Using Models to Understand Individual Differences in Personality, Development, and Mental Illness. Personality Neuroscience, 1. https://doi.org/10.1017/pen.2018.14
7 Cardinal, R. N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. Neural Networks, 19(8), 1277-1301.
8 Salamone, J. D. (2009). Dopamine, Behavioral Economics, and Effort. Frontiers in Behavioral Neuroscience, 3(September), 1–12. https://doi.org/10.3389/neuro.08.013.2009
9 Hull, C. L. (1943). Principles of behavior (Vol. 422). New York: Appleton-century-crofts.
10 Shafiei, N., Gray, M., Viau, V., & Floresco, S. B. (2012). Acute stress induces selective alterations in cost/benefit decision-making. Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology, 37(10), 2194–2209. https://doi.org/10.1038/npp.2012.69
11 Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. Behavioural Processes, 100, 192–196. https://doi.org/10.1016/j.beproc.2013.09.014
12 Westbrook, A., & Braver, T. (2015). Cognitive effort: A neuroeconomic approach. Cognitive, Affective, & Behavioral Neuroscience, 15, 395–415. https://doi.org/10.3758/s13415-015-0334-y
13 Ang, Y. S., Lockwood, P., Apps, M. A., Muhammed, K., & Husain, M. (2017). Distinct subtypes of apathy revealed by the apathy motivation index. PLoS One, 12(1), e0169938.
14 Radakovic, R., McGrory, S., Chandran, S., Swingler, R., Pal, S., Stephenson, L., ... & Abrahams, S. (2020). The brief Dimensional Apathy Scale: A short clinical assessment of apathy. The Clinical Neuropsychologist, 34(2), 423-435.
15 Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. Journal of personality and social psychology, 67(2), 319.
16 Gray J.A. (1987). The psychology of fear and stress (Vol. 5). CUP Archive.
17 Johnson S.L. , Turner R.J., Iwata N. (2003). BIS/BAS levels and psychiatric disorder: An epidemiological study. Journal of Psychopathology and Behavioral Assessment, 25, 25–36.
18 Angelides, N. H., Gupta, J., & Vickery, T. J. (2017). Associating resting-state connectivity with trait impulsivity. Social Cognitive and Affective Neuroscience, 12(6). https://doi.org/10.1093/scan/nsx031
19 Kahnt T. , Heinzle J., Park S.Q., Haynes J.D. (2010). The neural code of reward anticipation in human orbitofrontal cortex. Proceedings of the National Academy of Sciences, 107, 6010–5.]

20 Scheres A. , Sanfey A.G. (2006). Individual differences in decision making: Drive and reward responsiveness affect strategic bargaining in economic games. Behavioral and Brain Functions, 2, 35.]
21 Laricchiuta D. , Petrosini L., Piras F., et al. (2014). Linking novelty seeking and harm avoidance personality traits to basal ganglia: volumetry and mean diffusivity. Brain Structure and Function, 219, 793–803.
22 Jarbo K. , Verstynen T.D. (2015). Converging structural and functional connectivity of orbitofrontal, dorsolateral prefrontal, and posterior parietal cortex in the human striatum. The Journal of Neuroscience, 35, 3865–78.]
23 Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi: 10.18637/jss.v067.i01.
24 Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. Trends in cognitive sciences, 22(4), 337-349.
25 Eisenberger, R., Weier, F., Masterson, F. A., & Theis, L. Y. (1989). Fixed-ratio schedules increase generalized self-control: Preference for large rewards despite high effort or punishment. Journal of Experimental Psychology: Animal Behavior Processes, 15(4), 383.
26 Arkes, H., & Blummer, C. (1985). The Psychology of Sunk Cost Organizational Behaviour and Human Decision Processes, Vol. 35.
27 Abdellaoui, M., Barrios, C., & Wakker, P. P. (2007). Reconciling introspective utility with revealed preference: Experimental arguments based on prospect theory. Journal of Econometrics, 138(1), 356-378.
28 Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D., & Bestmann, S. (2015). Behavioral Modeling of Human Choices Reveals Dissociable Effects of Physical Effort and Temporal Delay on Reward Devaluation. PLoS Computational Biology. https://doi.org/10.1371/journal.pcbi.1004116
29 Wu, S. W., Delgado, M. R., & Maloney, L. T. (2011). The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. Journal of Neuroscience, 31(24), 8822-8831.
30 Bonnelle, V., Manohar, S., Behrens, T., & Husain, M. (2016). Individual Differences in Premotor Brain Systems Underlie Behavioral Apathy. Cerebral Cortex, 26(2), 807–819. https://doi.org/10.1093/cercor/bhv247
31 Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. Trends in Cogni- tive Sciences, 8, 539-546.
32 Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. Science, 306, 443-447.
33 Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. Cognitive, Affective, & Behavioral Neuroscience, 7, 356-366.
34 Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Banner- man, D. M. (2004). Action sets and decisions in the medial frontal cortex. Trends in Cognitive Sciences, 8, 410-417.
35 Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. Journal of Neuroscience, 23(16), 6475-6479.
36 Croxson, P. L., Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Pinsk, M. A., Gross, C. G., et al. (2005). Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. Journal of Neuroscience, 25, 8854-8866.
37 Kunishio, K., & Haber, S. N. (1994). Primate cingulostriatal projec- tion: Limbic striatal versus sensorimotor striatal input. Journal of Comparative Neurology, 350, 337-356.
38 Shinners, P. PyGame, (2011). Software available from pygame. org.
39 Zénon A, Sidibé M, Olivier E. (2015.) Disrupting the supplementary motor area makes physical effort appear less effortful. J Neurosci. 35:8737–8744.
40 Hadland, K. A., Rushworth, M. F. S., Gaffan, D., & Passingham, R. E. (2003). The anterior cingulate and reward-guided selection of actions. *Journal of neurophysiology*, *89*(2), 1161-1164.

41 Assadi, S. M., Yücel, M., & Pantelis, C. (2009). Dopamine modulates neural networks involved in effort-based decision-making. In Neuroscience and Biobehavioral Reviews (Vol. 33, Issue 3, pp. 383–393). https://doi.org/10.1016/j.neubiorev.2008.10.010
42 Massar, S. A. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. NeuroImage, 120, 104–113. https://doi.org/10.1016/j.neuroimage.2015.06.080
43 Fauth-Bühler, M., Zois, E., Vollstädt-Klein, S., Lemenager, T., Beutel, M., & Mann, K. (2014). Insula and striatum activity in effort-related monetary reward processing in gambling disorder: The role of depressive symptomatology. *NeuroImage: Clinical*, *6*, 243-251.
44 Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and Specificity of Reinforcement Signals throughout the Human Brain. Neuron, 72(1), 166–177. https://doi.org/10.1016/j.neuron.2011.08.011
45 Hampton, W. H., Alm, K. H., Venkatraman, V., Nugiel, T., & Olson, I. R. (2017). Dissociable frontostriatal white matter connectivity underlies reward and motor impulsivity. NeuroImage, 150, 336-343.
46 Colizoli, O., De Gee, J. W., Van Der Zwaag, W., & Donner, T. H. (2020). Comparing fMRI responses measured at 3 versus 7 Tesla across human cortex, striatum, and brainstem. *BioRxiv*.
47 Fehr, E., & Rockenbach, B. (2004). Human altruism: economic, neural, and evolutionary perspectives. Current opinion in neurobiology, 14(6), 784-790.
48 Hutchison, J., & Martin, D. (2015). The evolution of stereotypes. In Evolutionary Perspectives on Social Psychology (pp. 291-301). Springer, Cham.
49 Freeman, J. B., & Johnson, K. L. (2016). More than meets the eye: Split-second social perception. Trends in cognitive sciences, 20(5), 362-374.
50 Charness, G., & Rabin, M. (2002). Understanding Social Preferences with Simple Tests Author. The Quarterly Journal of Economics, 117(3), 817–869. https://doi.org/10.1162/003355302760193904
51 Asch, S. E. (1946). Forming impressions of personality. The Journal of Abnormal and Social Psychology, 41(3), 258.
52 Jenkins, A. C., Karashchuk, P., Zhu, L., & Hsu, M. (2018). Predicting human behavior toward members of different social groups. Proceedings of the National Academy of Sciences, 115(39), 9696 LP – 9701. https://doi.org/10.1073/pnas.1719452115
53 Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. Journal of personality and social psychology, 92(4), 631.
54 Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. Nature reviews neuroscience, 9(7), 545-556.
55 Chang, W. C., Chu, A. O. K., Treadway, M. T., Strauss, G. P., Chan, S. K. W., Lee, E. H. M., ... & Chen, E. Y. H. (2019). Effort-based decision-making impairment in patients with clinically-stabilized first-episode psychosis and its relationship with amotivation and psychosocial functioning. European Neuropsychopharmacology, 29(5), 629-642.
56 Lockwood, P. L., Hamonet, M., Zhang, S. H., Ratnavel, A., Salmony, F. U., & Husain, M. (2017). Prosocial apathy for helping others when effort is required. 0131(June), 1–10. https://doi.org/10.1038/s41562-017-0131
57 Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2016). Mesolimbic dopamine signals the value of work. Nature Neuroscience, 19(1), 117–126. https://doi.org/10.1038/nn.4173
58 Floresco, S. B., Tse, M. T. L., & Ghods-Sharifi, S. (2008). Dopaminergic and glutamatergic regulation of effort- and delay-based decision making. Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology, 33(8), 1966–1979. https://doi.org/10.1038/sj.npp.1301565
59 DeWall, C. N., Baumeister, R. F., Gailliot, M. T., & Maner, J. K. (2008). Depletion makes the heart grow less helpful: Helping as a function of self-regulatory energy and genetic relatedness. Personality and Social Psychology Bulletin, 34(12), 1653-1662.
60 Trezza, V., Baarendse, P. J., & Vanderschuren, L. J. (2010). The pleasures of play: pharmacological insights into social reward mechanisms. Trends in pharmacological sciences, 31(10), 463-469.

61 Hsu, M., Anen, C., & Quartz, S. R. (2008). The right and the good: distributive justice and neural encoding of equity and efficiency. science, 320(5879), 1092-1095.
62 Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. The quarterly journal of economics, 114(3), 817-868.
63 Blanco, M., Engelmann, D., & Normann, H. T. (2011). A within-subject analysis of other-regarding preferences. Games and Economic Behavior, 72(2), 321-338.
64 Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. American economic review, 94(4), 857-869.
65 Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M. L., & Dreher, J. C. (2010). Separate valuation subsystems for delay and effort decision costs. Journal of Neuroscience, 30(42), 14080-14090.
66 Ko, D., & Wanat, M. J. (2016). Phasic dopamine transmission reflects initiation vigor and exerted effort in an action-and region-specific manner. Journal of Neuroscience, 36(7), 2202-2211.