Why Don't You Try Harder? A Model of Effort Choice in Depression

A thesis presented

by

Alisher Batmanov

to

The Department of Economics and Business

in partial fulfillment of the requirements

for the degree of

Master of Arts

in the subject of

Economics

Central European University

Vienna, Austria and Budapest, Hungary

June 2021

Thesis Advisor:

Professor Marc Kaufmann

Author:

Alisher Batmanov

Abstract

I develop a simple model of effort choice for a person in a state of depression, who struggles to mobilize herself to carry out even the most elemental tasks under the impact of depressive anhedonia. This framework incorporates two key features exhibited by depressed individuals: perceiving the task at stake to be overly challenging to perform and diminishing the value of its reward characterized by the lack of motivational incentive. Even though taking up an activity such as exercising at the gym, reading a book or painting a picture in fact helps mitigate her perception of the task's excessive difficulty, a projection-biased person in depression fails to fully anticipate this and correctly evaluate the soothing effect that engaging in these tasks can bring while clouded by her current depressive state. As a consequence, she suboptimally ends up exerting too little effort or withdrawing from the task at all. To test these theoretical predictions, I ran an online experiment in which I elicited participants' willingness to work on real-effort tasks as well as the presence and severity of depressive symptoms using a widely used self-reported questionnaire. Subjects with more pronounced depressive symptoms chose to do 0.1-0.3 fewer tasks and exhibited a lower projection bias equivalent to 0.5-1.0 tasks, or 7.5-15%.[1]

[1]The research supporting this Master's thesis was fully sponsored by Central European University Foundation of Budapest (CEUBPF). The theses explained herein represent the ideas of the author, and do not necessarily reflect the views of CEUBPF.

Acknowledgements

I would like to thank my advisor Marc Kaufmann for his constant support with this thesis, for having offered me research assistance opportunity and for having tolerated me throughout a roller-coaster ride that grad school applications cycle has been for me. I am also grateful to Nomin Margad-Erdene, Ainura Karabayeva, Viktoriya Tsoy, Nurzhan Onashabay and Milana Mukiyeva for numerous ramen parties and soju nights we have had - without these I would have struggled with my mental health much more often.

I want to dedicate this thesis to two people. First, to my mom who has sacrificed a lot to invest in my education and never make me feel deprived of anything. Her strength and love help me in every endeavor I undertake. Love you mom! And second, to my dear friend Ablay Idris who has shown me what a true friendship really is about. I cannot wait until you win over this obstacle and we finally meet again.

1. Introduction

More than 264 million people globally are diagnosed with some form of depressive disorders (James et al., 2018). As an extreme consequence of a major depressive disorder (hereinafter depression), suicide was the cause of death for nearly 800,000 people in 2015 with a number of those who attempted but didn't "succeed" being way much higher (WHO, 2017). In the same year, suicide was the second leading cause of death for young adults aged 15 to 29 (WHO, 2017). At these staggering rates, the salient presence of depression nowadays is more significant than ever before, and despite a widespread disbelief about its severity, one cannot escape from depression by simply ignoring it and not taking actions.

When it comes to coverage of depression in economics, I join De Quidt and Haushofer (2016) in their view that "given this high prevalence, the significant economic cost, and economists' interest in other psychiatric conditions such as substance abuse, it is perhaps surprising that depression has not received a greater attention in the economics literature". Stemming from the work by Hamermesh and Soss (1974), the literature on suicide is relatively bigger, yet the underlying mechanisms behind a person's decision to end her life, which is frequently due to some severe form of depression, to the best of my knowledge have not been thoroughly investigated by economists.

One of the main characterizing features of depression is the state of anhedonia whereby a person is unable to derive pleasure from engaging in activities that once used to be rewarding. In this paper, I contribute to narrowing the gap in economics literature on depression by presenting a simple model of effort provision by a person in a state of depression. Within the framework in which a person decides how much effort to exert in a specific task, I draw some key channels through which anhedonia paralyzes a person's ability to experience pleasure. In particular, I show that a simple projection bias framework can explain how an agent in a state of depression fails to correctly predict her future utility, which leads to effort under-provision or, in some cases, withdrawal from doing the task at stake.

By no means am I the first to apply the mechanism of projection bias to understand mental health outcomes. O'Donoghue and Rabin (2001) formalize the situation in which a person in the current state of depression fails to correctly assess her future utility when she is healthy and, thus, incorrectly makes a choice to commit suicide. Formally, with a 2-period model where the person is depressed (d) in period 1 and healthy (h) in period 2, she receives the utilities $u(\text{life}, d) = -2$ and $u(\text{life}, h) = 3$ in case she decides not to commit suicide at the start of period 1, with $u(\text{suicide}) = 0$. Although it is obviously optimal to live ($-2 + 3 = 1 > 0$ under no discounting), with simple projection bias α she perceives her future payoff from living to be $\tilde{u}(\text{life}, h|d) = (1-\alpha)u(\text{life}, h) + \alpha u(\text{life}, d) = 3 - 5\alpha$. Under a high enough projection bias ($\alpha > 1/5$), she therefore suboptimally decides to commit suicide. For a more general case when a person lives for T periods, Hong and Lee (2015) present a forward looking model in which a person exhibiting projection bias fails to correctly evaluate the stream of her future utility when clouded by a current depressed state. The latter has a lasting impact when a person estimates the present value of her lifetime utility, which in turn prompts her to commit suicide.

Even though this simple and straightforward mechanism provides an intuitive "justification" for a person's choice to commit suicide, it does not tackle any of the more complex aspects describing humans' behavior when in a state of depression, except for its very extreme end result - suicide. The goal of this paper is to address one, arguably most characteristic and prevalent, feature that Beck (1967) in his seminal book on depression in psychiatry titled as a *paralysis of the will*: "they (the depressed) have a major problem in mobilizing themselves to perform even the most elemental and vital tasks ... although they can define for themselves what they should do, they do not experience any internal stimulus to do it".

I focus on how individuals' choices in effort provision framework are affected by the state of depressive anhedonia. According to the Diagnostic and Statistical Manual of mental disorders (DSM, DSM, 2013), a major depressive episode is marked by a period during which a person exhibits depressive mood as well as diminished pleasure in everyday activities. Evidently, this adverse state impacts individual's effort production in various domains of day-to-day life such as accomplishing assigned tasks at work, studying for an upcoming exam, visiting gym, or even going to the movies with friends. In this paper, I model how anhedonic depressive symptoms change individual's decisions of effort provision in costly effort tasks. In particular, I account for two key aspects characterizing an effort choice by a person in such state. First is a perception that all tasks are increasingly difficult, such that even the simplest appear to incur a large cost. Second has to do with a distortion of motivational incentive to engage in the given task as the goal at stake seems to be less appealing. As a result of these two forces, the agent in depression ends up exerting less effort in costly effort tasks when compared to a healthy person.

Additionally, I show that a mechanism of simple projection bias can be utilized to further portray under-provision of effort in a state of depression. As noted by Loewenstein (2005): "Depression provides an excellent example both of projection bias and of its consequences for medical decision making. ... When a person is depressed, they have a problem imagining ever getting better. ... Projection bias will cause people in a state of acute pain, misery and depression to both exaggerate how long they have felt bad and how long they are likely to to continue to feel bad". In the model proposed in this paper, I illustrate that a dysphoric[2] person overstates her perception of the task being overly challenging by underestimating the extent to which engaging in this task can actually dampen her feeling of the task's inflated disutility. For example, a depressed individual contemplating going to the gym will not only exhibit a higher disutility of doing weight training and cardio, but she will also underesti-

[2]Very unhappy, uneasy, or dissatisfied: marked or characterized by dysphoria.

mate the fact that over the course of exercising her perception that the training is excessively difficult will gradually fade away [3]. She errs in doing so as she projects her current depressed state onto the future, when her depressive mood will be milder precisely due to the soothing effect that exercising at the gym brings.

I conducted an online experiment with a total of 115 participants recruited through the Prolific platform with a goal to experimentally test the main predictions of the theoretical model. The design of the study is based on eliciting subjects' willingness to do costly effort tasks using a relatively straightforward Becker-Degroot-Marshak mechanism (Becker et al. (1964)) in a close resemblance to the experiment by Bushong and Gagnon-Bartsch (2020) on interpersonal projection bias. There were three main stages in the experiment: (i) subjects completed a Personal Health Questionnaire-9 (PHQ-9) on depressive symptoms; (ii) then they did 3 matrix counting tasks to get familiarized with the format and made incentivized predictions about doing additional tasks for additional pay at the end of experiment; (iii) lastly, the participants completed 12 more counting tasks, after which they made actual choices about their willingness to do additional work at that moment. Upon completion of the entire experiment, participants received either £4 or £4.50 depending on their incentivized predictions sent to their Prolific account in less than 24 hours.

There are three main findings in this experiment: first, when in a fresh state subjects' predictions about their future willingness to do additional work for £5 were on average 2.65 tasks higher compared to their actual choices that they made having completed all fifteen mandatory tasks, when they are relatively more tired. This level of projection of one's current state is reasonably close to the estimate found by Bushong and Gagnon-Bartsch (2020). Second, when analyzing a relationship between participants' actual/predicted willingness to work and their depressive symptoms, I found that a 10-points increase in the PHQ-9 score

[3] Note that this notion of inflated disutility effect should not be confused with the disutility of effort itself. While the former decreases with the time spent exercising, the latter will nevertheless increase.

(whose range is 0-27 points, where 27 is the most severe depression level) is associated with a 0.1-0.3 tasks decrease in the level of predicted or actual willingness to do more work for £3, £5 or £7 when controlling for a vector of individual covariates. Last, I found that there is a consistent difference in the level of projection by healthy and depressed participants, which I claim might stem from a projection of inflated disutility exhibited by the latter in line with implications of a theoretical model discussion in the previous sections. A 10-points increase in the PHQ-9 score has been found to narrow the gap between the actual and predicted willingness to work for £5 by around 0.5-1.0 tasks.

The rest of the paper is structured as follows: in Section 2, I present a simple model of binary effort choice for a person in a state of depression. Section 3 will generalize this framework to the case of continuous effort choice. The mechanism of projection bias will be added to the existing model in Section 4. Experimental design will be discussed in Section 5, while the results of the experiment will be presented in Section 6. Section 7 concludes.

2. The Model

I am going to consider a very simple model to illustrate a mechanism in which a person with a major depressive disorder decides on her effort choice. Assume that a person is confronted with a task which demands effort e and provides a return $w(e)$ upon successful completion of the task. Exerting effort is costly, so the person incurs a disutility of $d(e)$. For simplicity, assume that the effort choice is binary, $e = \{0, 1\}$, so that $w(e = 1) = \bar{w} > 0$ and $d(e = 1) = \bar{d} > 0$ are constants and $w(e = 0) = d(e = 0) = 0$. Hence, a rational utility maximizing agent chooses to do the tasks if and only if the return $\bar{w}$ exceeds or equals the cost $\bar{d}$. The next step is to depict in what ways this simple utility specification, $u = \bar{w} - \bar{d}$, gets distorted for a person in depressive state.

When faced with a choice of exerting effort in a task that yields reward, any agent evaluates two aspects which comprise her decision: perceived difficulty of the task and importance of the goal's value. For people with depression, anhedonia potentially affects both of these aspects. One of the most widespread and characterizing features among depressed people is their pronounced perception that all tasks are increasingly difficult, such that even the simplest appear to incur a great effort. Think about a university professor who needs to write a reference letter for her student: in the presence of depressive symptoms she would find this task to be exceptionally effort-demanding and even if she does finish writing a letter, the disutility she incurs is much higher than if she were comparatively healthy. In the model, I introduce a fixed (for now) parameter γ, which represents the extent to which a dysphoric person's perception of task difficulty is skewed to loom larger, so that her disutility of effort is given by $(1+\gamma)\bar{d}$, which increases in γ. Note that the benchmark disutility is that of a nondysphoric person, for whom $\gamma = 0$. In this binary choice framework, the more depressed a person gets (and the higher her γ becomes), the more likely she is to withdraw from supplying an effort.

In the next section, once I allow for a continuous effort choice in so-called *unfixed tasks*, the optimal effort e^* that a person chooses to exert will essentially mean how much time she considers optimal to be spent engaging in this task. In this respect, effort in the present model does not portray the intensity of one's labor as, for example, is in Brinkmann and Gendolla (2008), who in their lab experiment found the evidence that compared to healthy controls people with depressive symptoms exert more intensive effort in relatively easy tasks precisely because these tasks loom challenging for them, and less intensive effort in more difficult tasks since they struck them as considerably less feasible. In the current simple model with a binary effort choice a sufficiently high γ can make dysphoric people withdraw from undertaking the task. Importantly, however, whenever $(1+\gamma)\bar{d}$ exceeds the reward from completing the task and the person decides not to exert effort, she does so optimally since the increasing pain from her depressed state augments disutility she derives from engaging in such activity.

Additionally, the agent experiencing anhedonia could also distort her motivational incentive to engage in the given task since for her the goal would appear to be less appealing. Cléry-Melin et al. (2011) in their experiment involving patients with depression found that potential rewards in the form of money did not stimulate their effort production, which came in a stark contrast to healthy controls. They call this phenomenon a specific "incentive motivation deficit" whereby a potential reward fails to energize a production of force. An illustrative example of incentive motivation in Cléry-Melin et al. (2011) is cheering up a depressed friend by inviting to go to a movie just to receive a negative response. In the present setting I model such motivational characteristic of behavior through a modification in which a person distorts her perception of return to effort ($\bar{w}$) associated with completing the task so that she diminishes it by some fraction. Hence, her perceived return is $\tilde{w} = (1-\theta)\bar{w} + \theta \cdot 0 = (1-\theta)\bar{w}$, i.e. it lies somewhere between the true value $\bar{w}$, the reward perceived by a person without depressive symptoms, and the lower bound of potential benefit, which I arbitrarily set equal to zero. Her depressed state triggers a distorted perception of reward's value, which appears to be diminished for her.

Hence, the person makes a decision to exert effort in a specific incentive task iff:

$$u(e=1) = (1-\theta) \cdot \bar{w} - (1+\gamma)\bar{d} \ \geq \ 0 \tag{1}$$

$$\text{where} \quad \gamma, \theta \in [0,1]$$

Rearranging, the optimality condition for the binary effort choice framework whereby a person chooses to exert an effort ($e = 1$) is given by:

$$\bar{w} \geq \frac{(1+\gamma)\bar{d}}{1-\theta} \tag{2}$$

The interpretation of γ in the model seems to be rather straightforward: an individual exhibiting anhedonic depressive symptoms shall perceive the tasks to be more challenging in the sense of causing more disutility of effort in contrast to nondysphoric people. Parameter θ, on the other hand, serves to capture the fact that for a person with major depression the goal's value seems to be diminished, which discourages her from supplying more effort as the goal at stake is less desirable. Even though both of these effects (task difficulty through γ and motivational deficit through θ) unanimously impact the choice of an agent about her effort provision, I shall argue that the interpretations of such choices to exert the effort or not are not quite equivalent in the sense of optimality.

As outlined above, whenever a rise in the perception of task's difficulty ($\gamma \uparrow$) leads a depressed person not to do the task she will not exert the effort, which is optimal. This is not the case for motivational deficit effect. As a brief digression, in the model of projection bias in effort choices by Kaufmann (2021), in an all-or-nothing task a person projects her current rested state onto her future self and thus suboptimally starts to work on this task, which is not worth doing in the first place, to later realize that she has grown tired and is no longer willing to exert the effort. In doing so, she errs in her effort choice as her distorted perception of future disutility enables her to waste effort without receiving any benefit. In a similar vein, a dysphoric individual makes a mistake of withdrawing from doing the task in the current framework because of her distorted perception of the reward.

Owing to these two driving factors influencing effort exertion, individuals with depression withdraw from engaging in many activities that they once found worth spending time on.

A person that used to go to the gym regularly stops exercising, as not only does she find it more effort demanding, but also her evaluation of benefits are clouded, so that she underappreciates the value of a better shape, relieved stress and potential distraction from the daily chores that going to the gym would bring her. For a college student in depression completing a course assignment appears to be an overly burdensome task requiring devilish effort. On top of that, this student starts to question whether it is worth doing it in the first place. She fails to correctly assess the value of her completed assignment, that upon submitting it she would feel the sense of accomplishment, rejoice over a good grade or get closer to graduation.

This simple mechanism can equally be applied to even the most basic things in our everyday life, such as preparing a dinner or going to the movies with friends. Dysphoric person experiences a perception of excessive effort that she needs to incur in order to choose a movie, get dressed and take metro to get to the movie theater. If the burden of accomplishing these seemingly simple tasks is too high, she decides to stay home instead, which is optimal given her current depressive state. Yet if she makes such decision because she additionally underestimates the benefit of spending some quality time with her friends by having a nice walk and chat in the park after the movie, she then errs in doing so, failing to correctly evaluate the healing effect this time with her friends could potentially have. To illustrate the implications of this simple model, let me portray a straightforward example.

Example 1

Suppose there are four people with identical preferences yet different mental health states, who face a choice of whether to go to the gym and exercise or not. Getting to the gym and spending time exercising is costly and the disutility of effort in this particular task is fixed at $d(e) = \bar{d} = 3$ units of utility. The benefit or reward of visiting gym is also fixed and equals $w(e) = \bar{w} = 5$ units of utility: this can represent a boost in energy, feeling of accomplishment or merely sense of managing one's health. For all four, the alternative is to stay at home which yields 0 units of utility. Four people in focus are: Musa ($\gamma = 0, \theta = 0$) who is mentally healthy and exhibits no signs of depression; Dana ($\gamma = 1/3, \theta = 0$) is depressed and exhibits increased perception of difficulty associated with going to the gym; Alan ($\gamma = 5/6, \theta = 0$) is in a similar state to Dana's yet with a more severe feeling of excessive difficulty; Saya ($\gamma = 1/3, \theta = 2/5$) is suffering from depression, bears the same perception of the task difficulty as Dana and, additionally, deflates the reward value that going to the gym brings.

Using the model specified in this section, each of the four people makes a decision about going to the gym as follows:

- For Musa ($\gamma = 0, \theta = 0$): $u(e = 1) = (1 - \theta)\bar{w} - (1 + \gamma)\bar{d} = 5 - 3 = 2 > 0$

So he decides to **visit the gym** and enjoys a utility of 2 units.

- For Dana ($\gamma = 1/3, \theta = 0$): $u(e = 1) = (1 - \theta)\bar{w} - (1 + \gamma)\bar{d} = 5 - (1 + 1/3) \cdot 3 = 1 > 0$

So, she likewise decides to **visit the gym** yet enjoys a smaller utility of 1 unit compared to Musa as her depressed state inflates her perception of effort cost.

- For Alan ($\gamma = 5/6, \theta = 0$): $u(e = 1) = (1 - \theta)\bar{w} - (1 + \gamma)\bar{d} = 5 - (1 + 5/6) \cdot 3 = -0.5 < 0$

So he chooses **not to go to the gym**. Given the severe degree of depression, it is an optimal choice for Alan to stay at home.

- And, finally, for Saya ($\gamma = 1/3, \theta = 2/5$): $u(e = 1) = (1 - \theta)\bar{w} - (1 + \gamma)\bar{d} =$

$$= (1 - 2/5) \cdot 5 - (1 + 1/3) \cdot 3 = 3 - 4 = -1 < 0$$

Albeit the same extent of perceived difficulty as Dana's ($\gamma = 1/3$), she decides **not to go to the gym** as she mistakenly perceives that the value of the goal at stake - benefit from exercising in the gym - is diminished and it appears as not worth exerting an effort. In the absence of this erroneous perception of motivational deficit, it would otherwise be optimal for her to engage in exercising at the gym.

3. Continuous effort choice

In this section I will abstract away from the binary choice and present a more general model accounting for the fact that an agent can choose how much time to spend exerting effort in a task. For simplicity, effort e is chosen on the interval [0,1]. In addition, I am going to assume that the task at stake is characterized by increasing and concave benefit function $w(e)$ and increasing and convex disutility function $d(e)$. As before, a decision-maker in a state of depression discounts the reward by a constant θ, which is independent of the choice of e. In contrast, the parameter measuring inflated disutility effect γ in the framework of continuous effort choice is endogenized and depends on the amount of effort chosen: before exerting any work, a dysphoric agent forms her initial belief about how hard the task is, yet she also anticipates that this perception of an overly cumbersome task will gradually fade over time[4].

The way I present this feature in the model is by imposing that $\gamma(e) \in [0,1]$ measures the magnitude of instantaneous perception that the task is excessively challenging after having worked for e hours. Following the argument above, I claim that $\gamma(e)$ monotonically decreases in e. For simplicity, I assume that it is a linear function of the form $\gamma(e) = 1 - e$. Then, the overall effect $\Gamma(e)$ is just the integral of $\gamma(e)$ over e^* hours that the person has chosen to spend engaging in the task.

Summing up, a depressed person supplies effort e^*, determined by the following maximization problem:

$$e^* = \underset{0 \leq e \leq 1}{\arg\max}\ u(e) = (1-\theta)w(e) - \Big(1 + \Gamma(e)\Big)d(e) = (1-\theta)w(e) - \Big(1 + \int_0^e \gamma(\varepsilon)d\varepsilon\Big)d(e) =$$

[4] I hereby assume that the task in focus will necessarily bear this feature, i.e. despite a steady rise in the total disutility of effort exerted as time passes, a person with depressive anhedonia shall experience a gradual fall in her γ. In fact, many types of activities in real life can be classified as such. For instance, reading a good novel, exercising at the gym, painting or going for a morning jog are all likely to decrease the perceived excessive difficulty through their positive, therapeutic effect on the person's mental state.

$$= (1-\theta)w(e) - \Big(1 + \int_0^e (1-\varepsilon)d\varepsilon\Big)d(e) = (1-\theta)w(e) - \Big(1 + e - \frac{e^2}{2}\Big)d(e) \tag{3}$$

with the following FOC:

$$\underbrace{(1-\theta)\frac{\partial w(e)}{\partial e}}_{\text{Marginal benefit}} = \underbrace{\Big(1 + e - \frac{e^2}{2}\Big)\frac{\partial d(e)}{\partial e} + (1-e)d(e)}_{\text{Marginal cost}} \tag{4}$$

For a mentally healthy person, $\gamma(e) = 0 \implies \Gamma(e) = 0$ and $\theta = 0$, so that the necessary condition for optimality is a usual one: $\partial w(e)/\partial e = \partial d(e)/\partial e$. In this model, a decision-maker in a state of depression, apart from the two distortive effects coming through γ and θ, is otherwise rational, utility-maximizing agent.

4. Depression and Projection Bias

The focus of this section will be to further enrich the current model of effort provision in a state of depression by introducing the mechanism of projection bias. Following the definition by Loewenstein et al. (2003), a person exhibits *a simple projection bias* if there exists $\alpha \in [0,1]$ such that $\forall\ c, s,$ and s':

$$\tilde{u}(c, s|s') = (1-\alpha)u(c,s) + \alpha u(c, s') \tag{5}$$

whereby a person currently in state s' attempts to predict her future instantaneous utility from consuming c when in state s. I will apply this notion to model a particular behavior by a dysphoric person, characterized by *projecting inflated disutility*: having exerted effort for e' hours, a person attempts to predict her future perception of the task being overly difficult when she has been working for e hours:

$$\tilde{\gamma}(e|e') = (1-\alpha)\gamma(e) + \alpha\gamma(e') \tag{6}$$

The more projection-biased a person is (higher α), the more she mispredicts her future $\gamma(e)$ while clouded by her current adverse state after having worked for e' hours. Hence, she projects her current state's γ onto the future, failing to fully evaluate the soothing effect of being engrossed in the task. First, I will portray how projecting inflated disutility affects effort provision in an all-or-nothing tasks. Next, I will extend the continuous effort choice framework to account for projection bias.

4.1. Binary Effort Choice Framework

Let me abstract from the motivational deficit effect in depressed people and for the moment assume that there is no difference in how dysphoric and nondysphoric individuals value the goal at stake ($\theta = 0$). Nonetheless, the effect of inflated disutility of effort among the depressed shall remain there. In line with the previous sections, the effect of changes in γ on the effort provision choice would still result in an optimal decision of whether to undertake the task or not. At the same time, I will extend the simple static model to a discrete two-period model of effort tasks: imagine that a particular task, such as completing a school project, requires an effort exerted in two consecutive periods $t = 1, 2$ and a person decides whether to take up the task at the beginning of the first period. For simplicity assume that $\bar{w}_1 = 0$, $\bar{w}_2 = \bar{w}$ and $\bar{d}_1 = \bar{d}_2 = \bar{d}$ so that a person makes a choice to undertake this task if and only if[5]:

$$u(e_1 = e_2 = 1) = -(1 + \gamma_1)\bar{d} + \bar{w} - (1 + \gamma_2)\bar{d} \geq 0 \implies \bar{w} \geq (2 + \gamma_1 + \gamma_2)\bar{d} \tag{7}$$

where γ_1,$\gamma_2 \in [0, 1]$ are the degrees to which she perceives the task to be excessively difficult to undertake at $t = 1$ and $t = 2$, respectively.

I claim that in a scenario when the extent of depression is less severe in period 2 compared

[5]Agent receives her reward $\bar{w}$ at the end of the second period conditional on effort exerted in both periods.

to period 1, i.e. when $\gamma_2 < \gamma_1$, the projection biased person in the first period shall be clouded by her current adverse state and thus misperceive her future utility when she in fact feels much better[6]. In other words, at $t = 1$ and given her current value of γ_1 she predicts that her future γ_2 will be:

$$\tilde{\gamma}_{2|\gamma_1} = (1 - \alpha)\gamma_2 + \alpha\gamma_1 \tag{8}$$

And thus, her effort exertion condition transforms to:

$$\bar{w} \geq (2 + \gamma_1 + (1 - \alpha)\gamma_2 + \alpha\gamma_1)\bar{d} \quad \Longrightarrow \quad e_1 = e_2 = 1 \iff \alpha \leq \frac{\bar{w} - (2 + \gamma_1 + \gamma_2)\bar{d}}{(\gamma_1 - \gamma_2)\bar{d}} \tag{9}$$

Thus, for a sufficiently projection-biased person the right decision would appear to be not to take up this task at all, although it could in fact be optimal to exert the effort since the marginal return in the second period over-weighs the cost she incurs in the first.

Example 2

Imagine that Alan ($\gamma_1 = 5/6$) from the previous example is a projection-biased student ($\alpha = 1/2$) confronted with the task in his arts class to paint a picture which shall require the work to be exerted for two consecutive days. The process of painting causes him a disutility of effort such that $\bar{d}_1$=$\bar{d}_2$=3, and he enjoys the return to this task only once he completes the picture at the end of period 2 with $\bar{w}_2 = 10$. At the start of period 1, Alan has to decide whether to engage in the task or not. He also acknowledges that he has a major depression with a high parameter of excessive disutility of effort $\gamma_1 = 5/6$, so he takes this into account when making a choice. Another important aspect of this task is that the process of painting itself helps to relieve the stress and mitigate the diminished ability to derive pleasure, so that after spending the first day painting the picture, Alan's γ parameter will decrease from $\gamma_1 = 5/6$ to $\gamma_2 = 1/3$.

[6]Engaging in activity, such as writing a literary essay for a college class, helps a person in depression get distracted from her pain, relieve stress and mitigate perception of excessively cumbersome task.

If Alan were not projections-biased ($\alpha = 0$), he would make a choice based on the optimality condition outlined in (5):

$$10 \geq (2 + 5/6 + 1/3) \cdot 3 \quad \Longrightarrow \quad 10 \geq 9.5 \quad \top$$

so he would prefer to take up the painting task. However, while in the state of severe depression he fails to perceive the benefit he would get from engaging in this process. In other words, his current adverse state does not allow him to acknowledge the fact that the pain in the second period will be mitigated, as he projects his current state onto his future. Thus, his effort provision condition will be:

$$10 \geq \Big(2 + 5/6 + (1/2)(1/3) + (1/2)(5/6)\Big) \cdot 3 \quad \Longrightarrow \quad 10 \geq 10.25 \quad \bot$$

which does not hold, so he decides not to take up this task, although it in fact would bring him positive utility from engaging in it.

4.2. Continuous Effort Choice Framework

Recalling the maximization problem for a dysphoric person in the case of an unfixed task, and setting $\alpha = 0$ as in the case for a binary effort choice:

$$e^* = \arg\max_{0 \leq e \leq 1} \; u(e) = w(e) - \Big(1 + \Gamma(e)\Big)d(e) = w(e) - \Big(1 + \int_0^e \gamma(\varepsilon)d\varepsilon\Big)d(e) =$$

$$= w(e) - \Big(1 + \int_0^e (1 - \varepsilon)d\varepsilon\Big)d(e) = w(e) - \Big(1 + e - \frac{e^2}{2}\Big)d(e) \tag{10}$$

Now, let me modify the objective taking into account the projection bias exhibited by a person, who mispredicts her future γ after working for e hours and at the moment of making decision, i.e. before exerting any effort, has the following perception:

$$\tilde{\gamma}(e|e') = (1-\alpha)\gamma(e) + \alpha\gamma(e') \quad \Longrightarrow \quad \tilde{\gamma}(e|0) = (1-\alpha)(1-e) + \alpha \cdot 1 = 1 - (1-\alpha)e \tag{11}$$

Then, the objective transforms to:

$$\max_{0 \le e \le 1} u(e) = w(e) - \Big(1 + \tilde{\Gamma}(e|0)\Big)d(e) = w(e) - \Big(1 + \int_0^e \tilde{\gamma}(\varepsilon|0)d\varepsilon\Big)d(e) =$$

$$= w(e) - \Big(1 + \int_0^e (1-(1-\alpha)\varepsilon)d\varepsilon\Big)d(e) = w(e) - \Big(1 + e - \frac{(1-\alpha)e^2}{2}\Big)d(e) \tag{12}$$

with the following FOC:

$$\underbrace{\frac{\partial w(e)}{\partial e}}_{\text{Marginal benefit}} = \overbrace{\underbrace{\Big(1 + e - \frac{(1-\alpha)e^2}{2}\Big)}_{> \text{ than when } \alpha = 0}\frac{\partial d(e)}{\partial e} + \underbrace{\Big(1-(1-\alpha)e\Big)}_{> \text{ than when } \alpha = 0}d(e)}^{\text{Marginal cost}} \tag{13}$$

It is clearly seen that for a projection-biased person ($\alpha > 0$) the marginal cost of exerting another hour of her effort in the task is strictly higher than in the absence of projecting inflated disutility, for any given effort level e. Hence, given the concavity of the total benefit function $w(e)$ and convexity of the total disutility of effort $d(e)$, a projection-biased person in a state of depression supplies less effort.

To conclude this theoretical part of the paper and move to the experiment, let me concisely reiterate the chief takeaways from the model in the form of testable hypotheses to be revisited in an experimental setting. First, the model predicts that in a state of depression an agent perceives reward's value to be smaller yet incurs a higher effort cost compared to a healthy control, which leads her to exert less effort. Second, to attempt explain this behavioral pattern I introduce a projection of inflated disutility exhibited by a dysphoric person, in which she fails to correctly predict by how much she will overvalue the disutility of effort having worked on the task for some time. The remaining sections will elaborate on the design of the experiment that allows me to approach and assess these aspects as well as experimental findings and how they relate to model's predictions.

5. Experimental Design

In order to test the implications of the theoretical model in this paper, I ran an experiment involving making effort choices. I recruited a total of 115 participants using *Prolific*, the United Kingdom based online participant recruitment platform.[7] Participants followed a link which took them to a Qualtrics survey, where they completed all parts of the current study. At the beginning of the survey, they had to fill in their unique Prolific ID, which I then used to match the responses to participants' demographic information. At the end of the experiment I provided subjects with a completion code which they had to enter on the Prolific website to secure payment. The experiment had three distinct stages: in the first block participants were given detailed instructions for the entire experiment and then asked to fill in the PHQ-9 depression symptoms questionnaire. In the second and third blocks, they first had to do some effort demanding tasks and upon completing them make a prediction about their future willingness to do additional tasks for additional pay (second block) and state their actual willingness to do more tasks for an extra pay (third block).[8]

5.1 Measuring Depression

Depression, or a presence of depressive symptoms, is a key component comprising the effort choice model in this paper. One should, however, acknowledge that depression cannot be randomly assigned in any possible setting nor can it be painstakingly identified even by a medical professional. Moreover, I view depression, or any other mental health disorder, not

[7] Subjects for the experiment were selected to meet the following criteria: (i) between 18 and 60 years old; (ii) resident of the United Kingdom; (iii) at least 50 past studies they have taken part on Prolific with a 95% approval rate; and (iv) English as a first language.

[8] The experiment instructions in full can be found in the Appendix.

as a binary variable but rather as a shock to a rich set of experiences a person embraces on a daily basis. Apart from the clinical screening procedures, the first best approach in the mental health literature to capture depressive symptoms in a sample is utilizing standardized self-reporting screens in the form of surveys or questionnaires. In my experiment I recorded participants' depressive symptoms by having them fill in the Patient Health Questionnaire consisting of nine simple questions (hereinafter PHQ-9). The PHQ-9 is arguably one of the most widespread tools to screen a major depressive disorder characterized by high sensitivity and specificity. Notably, the PHQ-9 is used for clinical screening in physicians' offices and hospitals in addition to being a tool for epidemiological measurement (Kocalevent et al., 2013).

The PHQ-9 consists of nine items representing nine main symptoms of depression and respondents are asked how frequently they experienced each of these over the course of past 14 days. For each symptom there are four possible responses: not at all (0 points), several days (1 point), more than half the days (2 points) or nearly every day (3 points). Thus, the range for the PHQ-9 score is 0-27 points. The PHQ-9 exhibits an 88% sensitivity and 88% specificity for the diagnosis of a major depressive disorder with a threshold of greater than or equal to 10 (Kroenke and Spitzer, 2002). Sensitivity is a probability of testing positive for depression (PHQ-9$\geq$10) when the disorder is present, while specificity is a probability of testing negative for depression (PHQ-9$<$10) when the disorder is absent. Recent applications of the PHQ-9 in the economics research as an instrument to screen depression include Bolotnyy et al. (2021), who used it as a tool to screen depression among the graduate students of Economics in the top US programs, and Ridley et al. (2020), who utilized the PHQ-9 to analyze the bi-directional relationship between poverty and depression in India.

5.2 Predicting Workers

The real-effort task subjects undertook in my experiment is a canonical matrix counting task, in which participants are required to count the number of times a particular symbol/number (e.g. \$, \$\$ or 2) appears in a 10×15 matrix of numbers and symbols. The image below portrays one such example task. On average it took participants 67 seconds to complete each counting task. I randomized the specific symbol to count in each round and for each participant. The correct response in each task was required to advance to the next task.

Example of a matrix counting task

2	2	2	2	\$	\$\$	2	\$\$	2	\$	\$\$	\$\$	2	\$\$	\$
\$\$	\$	\$\$	\$\$	2	2	2	2	2	2	\$\$	2	2	2	\$
\$\$	2	\$	2	2	\$\$	\$	\$\$	\$\$	\$	\$	\$	\$	\$\$	\$
\$	\$	\$	2	\$	2	2	\$\$	\$\$	\$\$	\$	2	\$\$	2	2
\$	2	\$	\$\$	\$\$	\$\$	\$	\$\$	2	\$	\$	\$\$	2	2	2
\$\$	\$\$	\$	\$\$	2	2	\$	\$	\$	2	\$\$	2	\$	\$	\$
\$	\$\$	2	\$\$	\$	2	\$\$	2	\$	2	2	2	2	2	\$
2	\$	\$\$	2	\$\$	2	\$	\$\$	\$	\$\$	2	\$	\$\$	\$	2
\$	\$	\$	\$\$	\$\$	\$	\$	\$\$	\$	\$	2	\$	\$	2	\$\$
2	\$	\$\$	\$\$	\$	\$\$	2	\$\$	\$	\$\$	2	2	\$\$	\$	2

How many "\$" symbols are there in the matrix?

Next I move to the details of eliciting predictions about one's future willingness to do extra work as well as actual willingness to do extra work in this experiment. The design is largely based on Bushong and Gagnon-Bartsch (2020), who in their online experiment on interpersonal projection bias elicit workers' willingness to work (hereinafter WTW) using the Becker-Degroot-Marshak (BDM) mechanism.

Actual Willingness to Work: In the given experiment, once a participant has completed all 15 mandatory counting tasks and has thereby become relatively tired, I asked them the maximal number of additional tasks they were willing to complete for a bonus of

£M, where M was 3, 5 or 7. The choice set presented to participants is a set of integers from 0 to 30 tasks in the form of a slider question in the experiment. I denote this choice by $WTW_i(£M, T)$, where T represents workers' tired state. Thus, for each participant I record their actual WTW as $\{WTW_i(£M, T)$: M = 3,5,7$\}$.

Predicted Willingness to Work: At the instant when a worker has completed 3 out of 15 mandatory tasks and is still relatively fresh and rested, I interrupt them and ask to make an incentivized prediction about their future behavior. In particular, I ask them to think about their future selves and state the maximal number of additional tasks they forecast they would be willing to do for an extra payment of £5. I denote this prediction by $\widetilde{WTW}_i(£5, F)$, where F represents workers' fresh state.

An important caveat to be added regarding predicted WTW is that the predictions made by participants are incentivized in the following way: I make it clear that if their guess is within 3 tasks of their actual future choice (i.e. $\mid \widetilde{W}_i(£5, F)$ - $W_i(£5, T) \mid \leq 3$), they receive a bonus of £0.50. This is in line with Bushong and Gagnon-Bartsch (2020), who caution that inference based on these measures should come with a caveat because by incentivizing participants' predictions an experimenter might implicitly incentivize consistency. To partially address this issue, in addition to eliciting WTW for £5, I also ask about participants' WTW for £3 and £7 to later use per pound WTW to compare predictions with actual choices, which serves as a robustness check to assess the concern about subjects' decisions stemming from choosing to be consistent.

5.3 Measuring Tiredness

One of the ingredients in the proposed effort choice model in sections 3 and 4 is what I call an inflated disutility parameter $\Gamma(e)$, or its instantaneous counterpart $\gamma(e) \geq 0$, which represents the extent to which a depressed person perceives the disutility of working to be excessively high, having already exerted effort e. In particular, I hypothesise that for some class of tasks $\gamma(e)$ is a decreasing function of e implying that the more a person has worked for, the less she perceives the task to be overly challenging. To provide suggestive evidence about the dynamics of $\gamma(e)$ over time, in the experiment both after 3 and 15 mandatory tasks I ask participants to assess how tiring the past block of work has been for them on average on a scale from 1 to 10, where 1 is "not tiring at all" and 10 is "extremely tiring" (I denote their responses as τ_1 and τ_2). Although this measure is far from being a precise estimate of subjects' instantaneous perceived marginal disutility, it does at least inform about the direction of perceived tiredness, or a partial derivative of $\gamma(e)$.

Having introduced the experimental design, the next section will marry theoretical predictions and experimental findings to provide insights into effort choice in a state of depression.

6. Results

6.1 Descriptive statistics

Table 1 contains the baseline characteristics of the sample. Having screened out those that spent unreasonably little time on completing the entire experiment, the sample size is 115. The statistics are presented for the whole sample (N=115) as well as by stratifying subjects into those with more pronounced depressive symptoms as suggested by the PHQ-9 score equal or exceeding the threshold of 10 (N=31) and those with less pronounced depressive symptoms characterized by their PHQ-9<10 (N=84).[9] The sample is characterized by a somewhat low percentage of males (27.8%) and a high percentage of whites (91.3%) with a mean age of 34.1. For just over a quarter of the sample high school is the highest education degree attained; more than half of subjects (53%) are in full-time employment and

[9]Note that for a subsequent analysis I proxy the state of depression exclusively by a running variable PHQ-9, which is an integer value between 0 and 27, inclusive.

the fraction of those not in the paid work constitutes 11.3%. Median household income is in the range £30.000-£49,999. Mean completion time of the entire experiment is 21.6 minutes.

Most of the characteristics are comparable across two groups. Somewhat surprisingly, the proportion of people not in the paid work is higher in the group with less severe depressive symptoms (14.3% versus 3.2%). Likewise, the fractions of those with a household income less than £10,000 are markedly different across two groups (15.5% versus 0%). However, these raw correlates should be taken with caution given the small sample size and rather arbitrary threshold for the PHQ-9. Figure 1 in the appendix plots the kernel density of PHQ-9 scores in the sample. It is clear that the bulk of the mass is located in the 3-6 range with a sizeable tail to the right of the threshold of 10. The probability of having PHQ-9≥10 being nearly 27% showcases this fact.

Figure 2 presents a distribution of predicted and actual WTW for £5, i.e. $\widetilde{WTW}(\pounds 5, F)$ and $WTW(\pounds 5, T)$, respectively. The distribution of the predictions about how many extra tasks to complete is Gaussian-looking and symmetric around the mean of 8-9 tasks notwithstanding a sizeable tail to the right end. In a stark contrast, the kernel density plotted for the actual choice of WTW is highly skewed to the left with most of the mass centered around 4-5 tasks. Panels (1) and (5) in table 2 support this visual evidence pointing out to a sizeable gap between the mean predicted and actual WTW for £5 (9.38 versus 6.73 tasks). This gap of 39.4% is somewhat bigger than the one found in Bushong and Gagnon-Bartsch (2020), which was around 30%, but both the direction and the magnitude are comparable.

When comparing panels (2), (3) and (4) in table 2, which illustrate the WTW for £3, £5 and £7 in per-pound terms, one can see that the actual choice for £5 is larger than per-pound choices for £3 and £7, which are both of the similar size - this is coherent with a caveat mentioned in section 5.2 about the consistency motive when stating $WTW(\pounds 5, T)$.

Indeed, the stated WTW for both £3 and £7 is lower in per-pound terms than stated WTW for £5, which is in turn lower than the predicted WTW.

Lastly, it is instructive to learn some facts about participants' self-reported tiredness levels and how they changed between the states and across depression levels. To remind the reader, each subject in the experiment was asked about their perceived level of how tiring the past block of tasks has been on average. Panels (6) and (7) of table 2 portray that subjects report getting more tired with more counting tasks, which is in line with expectations. The average response in the sample increased from $\overline{\tau}_1$=4.37 to $\overline{\tau}_2$=6.23. More important, however, is to see how this growth in the perception of tiredness correlates with depression. Panel (8) in table 2 presents a regression of the percentage growth in reported tiredness ($\%\Delta\tau = (\tau_2 - \tau_1)/\tau_1 \cdot 100\%$) on the PHQ-9 score and a vector of individual covariates. The set of variables supposedly capturing some part of individual heterogeneity in the sample and used in all regressions throughout the paper includes age, ethnicity, gender, highest attained degree, number of household members, marriage status, student status, vision issues, experiment duration and whether a participant has children, works full-time or belongs to the lowest income category (the vector denoted by X). The coefficient for PHQ-9 shows that a 10-points increase in the PHQ-9 depressive symptoms score is associated with around 4% decrease in the tiredness growth rate. What this means is that although on average participants grow more tired based on their self-reports, individuals with more severe depressive symptoms grow tired to a lesser extent. When drawing a parallel to the theoretical model, this is roughly equivalent to saying that that holding everything else constant $\gamma(e)$ decreases in effort e. This crude estimate is admittedly not capturing many confounding factors yet it does provide some suggestive evidence about how the perception of tiredness interacts with depressive symptoms.

6.2 Main experimental findings

6.2.1 Do people project their tiredness state?

First and foremost, I check the main finding of Bushong and Gagnon-Bartsch (2020) regarding intrapersonal projection bias: rested individuals project their current state onto their future selves when they are tired and, therefore, overestimate their predicted WTW. I test this hypothesis in a very simple way, similar to how Chaloupka IV et al. (2019) estimate the difference between predicted and actual probability of quitting smoking: I stack the vector of predicted WTW on top of the actual WTW to get a twice as big vector, and then regress this variable (that I call W, such that $W = [\widetilde{WTW}; WTW]$) on the indicator function that is equal to 1 for observed WTW and 0 for an incentivized prediction. The regression equation that I fit is thus:

$$W(\pounds P)_i = \beta_0 + \beta_1 \mathbb{1}\{\text{WTW}\}_i + \epsilon_i \tag{14}$$

where $i = 1, ..., 2n$ and $P = \{3, 5, 7\}$. If we expect that people overestimate their prediction of future WTW while in the fresh state, then we would expect to see $\beta_1 < 0$. Table 3 provides estimation results. From the first column we see that participants' prediction about their WTW for £5 exceeds their actual choice by 2.65 tasks on average. To assess the robustness of this estimate, I vary the dependent variable using elicited WTW for £3 and £7 - to make these actual choices comparable with the predictions I divide both by the associated amount of payment to obtain the per-pound values. Subjects' predictions of WTW per pound using these dependent variables are 0.49 and 0.62 tasks smaller than their choices. If I multiply these coefficients by 5, I will obtain the values comparable to 2.65 tasks gap observed using £5.

6.2.2 Do people in depression choose to do less work?

In order to see whether being in a comparatively more depressive state is associated with lower actual or predicted levels of effort, I fit the following regressions:

$$WTW(\pounds P)_i = \beta_0 + \beta_1 PHQ9_i + X_i'\delta + \epsilon_i \tag{15}$$

$$\widetilde{WTW}(\pounds 5)_i = \beta_0 + \beta_1 PHQ9_i + X_i'\delta + \epsilon_i \tag{16}$$

where $P = \{3, 5, 7\}$ and X is a vector of individual fixed effects. The reason for including individual covariates is to partially address a potential threat of endogeneity: variable $PHQ9_i$ is apparently not randomly assigned - and cannot be randomly assigned in principle - so that it can and probably does correlate with ϵ_i. Conditioning on the vector of observables is expected to somewhat suppress this issue.

Table 4 presents the results. From the first two columns in Panel A, one can see that a 10-point increase in the PHQ-9 score is associated with a 0.6 and 0.3 tasks decrease in subjects predicted WTW for £5 without and with individual controls, respectively. When it comes to the actual choice about WTW for £5, the effect is also negative at 0.3 tasks when including individual covariates. Panel B presents the alternative regressions using WTW(£3) and WTW(£7) as dependent variables. A 10-point increase in the PHQ-9 score is translated into a 0.1 and 0.3 tasks decrease in WTW for £3 and £7, respectively. Without controlling for individual controls, the estimates are slightly smaller in their magnitude but are not considerably different. Even though the coefficients are not economically large, they do suggest that individuals with more severe depressive symptoms tend to exert marginally less effort, which in this case is proxied by WTW.

6.2.3 Do people in depression exhibit inflated disutility projection?

Based on the evidence provided so far, it is clear that experiment subjects do exhibit projection bias whereby their current rested state affects their prediction about their future selves, which in turn leads them to overestimate their WTW. The kind of projection I am putting forward in this paper comes in conjunction with projection of the tiredness state. As has been illustrated in the previous sections, I hypothesize that people in a state of depression could additionally mispredict the extent to which their perception of the task being overly difficult would eventually die away. If so, what enters the determination of an optimal effort for a person is $\tilde{\gamma}$ and not the "true" value γ. The implication of this projection effect is that depressed people end up choosing to exert less effort compared to their healthy counterparts.

As a result, if we take that healthy participants overestimate their future WTW, then we should expect that relatively more depressed participants would overestimate it to a lesser extent driven by $\tilde{\gamma}$. This is because even at the moment of making a prediction they are overestimating the excessive difficulty of the tasks and are, hence, underestimating their WTW. To test this hypothesis, I utilize two seemingly different but practically very similar estimation approaches. For Approach A, I use the original dataset and construct an artificial variable α for each participant which represents their individual level of projection:

$$\alpha_i = \frac{\widetilde{WTW}(\pounds 5)_i - WTW(\pounds 5)_i}{WTW(\pounds 5)} \cdot 100\% \tag{17}$$

I then regress this variable on the PHQ-9 score and a set of controls:

$$\alpha_i = \beta_0 + \beta_1 PHQ9_i + X_i'\delta + \epsilon_i \tag{18}$$

to see whether $\beta_1 < 0$, i.e. if those with more pronounced depressive symptoms tend to overall project less because of $\tilde{\gamma}$ effect. Upper panel in table 5 provides the regression estimates. The slope coefficient in column 1 implies that a 10-points increase in PHQ-9 is associated with a 12.1% decrease in α. As a robustness check, I fit similar regressions using WTW for £3 and £7. From column 2 in the Approach A panel one can see that the associated effect is negative 14.1%, which is rather similar to the one in column 1. Interestingly, the estimate coming from WTW for £7 is comparatively much smaller, which I suspect might stem from the fact that this created variable gets rather noisy when per-pound WTW for £7 is subtracted from per-pound predicted WTW for £5.

To further assess the effect of projecting inflated disutility, I fit a regression using a stacked variable W and the interaction between PHQ-9 and an indicator variable for actual WTW:

$$W(\pounds P)_i = \beta_0 + \beta_1 \mathbb{1}\{\text{WTW}\}_i + \beta_2 PHQ9_i + \beta_3 \mathbb{1}\{\text{WTW}\} \times PHQ9_i + X_i'\delta + \epsilon_i \qquad (19)$$

where $P = \{3, 5, 7\}$. The coefficient of interest is β_3, which is expected to be positive under the alternative hypothesis, i.e. since β_1 has already been shown to be negative, pointing out to the fact that participants overestimate their predicted WTW, we expect that for depressed people this overestimation would be suppressed and β_3 would be of the opposite sign to β_1.

The Approach B panel in table 5 presents the regression estimates. From the first column for W(£5) one can observe that a 10-points increase in the PHQ-9 score mitigates the prediction error by around 0.5 tasks. When taking the per-pound values for W(£3) and W(£7) as dependent variables, columns 2 and 3 in the same panel of table 5 showcase that the associated effect is positive 0.2 and 0.1 tasks, respectively, that are equivalent to 1 and 0.5 tasks when adjusted to WTW(£5), and are comparable with regression estimates utilizing W(£5) as a dependent variable. This is equivalent to a decrease in prediction error of around 7.5-15%. Overall, this analysis presents evidence that participants with more severe depressive symptoms tend to exhibit some inflated disutility projection bias which has an opposite effect on the chosen effort level when compared to a more "traditional" projection bias stemming from projecting a state of tiredness.

7. Conclusion

This paper starts by presenting a simple intuitive model that showcases what factors drive the decisions about effort provision made by a person in a state of major depressive disorder. I show that in addition to usual trade-off between the marginal benefit and the marginal cost of exerting one more hour of effort in a given task, there are two additional forces guiding a dysphoric person's effort choice. First, the model integrates the feature that in a costly effort task depressed individual perceives the task at stake to be overly difficult - the effect that I title "inflated disutility" (γ). Second, the utility specification also takes into account the fact that a person in a state of depressive anhedonia might experience a flawed distortion of reward's value for her, so-called "motivational incentive deficit" (θ).

I differentiate between a depressed person's choice to supply less effort because of high γ and that because of high θ. In the spirit of Becker and Murphy (1988), whenever a current state of pain makes a task seem overly effort-demanding thus leading to under-provision of effort, I claim that this choice is optimal in the sense of maximizing one's utility similarly to a person developing her addiction to a good in the theory of rational addiction. In contrast, if this decision is fueled by a misperception of the task's benefit for a decision-maker, a dysphoric person then errs in doing so and ends up exerting less effort than in fact is optimal. A similar conclusion that people in a state of depression tend to exert less effort than mentally healthy individuals is likewise reached by integrating projection bias in the model.

Practically all the costly effort tasks are modeled using a monotonically increasing and convex disutility function. Most of these tasks additionally bear a very specific property in the context of effort choice in a state of depressive anhedonia. For a dysphoric person $\gamma(e)$ represents her instantaneous perception that the task is excessively cumbersome after having exerted effort for e hours. What I claim is that in many situations as the time of engagement in completing the task goes by, $\gamma(e)$ likely shrinks monotonically because of the therapeutic effect that becoming engrossed in activity brings for a person in depression, such that $\partial\gamma(e)/\partial e < 0$. There are plenty of task examples that have this effect: exercising at the gym, going for a jog, reading a book, painting a picture, climbing mountains with friends, etc. Depressed and projection-biased person ($\gamma > 0,\ \alpha > 0$), however, fails to fully acknowledge this and correctly evaluate the soothing effect of engaging in these tasks when clouded by her current adverse depressive state. Projecting inflated disutility leads her to exert less effort that would otherwise be optimal for a person who is not projection-biased.

This paper contributes to a narrow literature in economics on theoretical mechanisms and experimental studies describing how individuals make decisions in a state of depression. The simple model at the core of this paper can generically be applied to a wide array of scenarios involving choices in costly effort tasks. Experimental findings in this paper broadly confirm the main theoretical implications of the proposed model. In the future research it would be informative to more precisely identify and quantify the causal effect of depressive anhedonia on effort provision in the domain of real-effort tasks by increasing the sample size and running between-subjects experiments involving people with clinically diagnosed depression.

References

(1) James, S. L.; Abate, D.; Abate, K. H.; Abay, S. M.; Abbafati, C.; Abbasi, N.; Abbastabar, H.; Abd-Allah, F.; Abdela, J.; Abdelalim, A., et al. Global, regional, and

national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **2018**, *392*, 1789–1858.

(2) WHO, *Depression and other common mental disorders: global health estimates*; 2017.

(3) De Quidt, J.; Haushofer, J. *Depression for economists*; 2016.

(4) Hamermesh, D. S.; Soss, N. M. An economic theory of suicide. *Journal of Political Economy* **1974**, *82*, 83–98.

(5) O'Donoghue, T.; Rabin, M. *Risky behavior among youths: An economic analysis*; University of Chicago Press, 2001; pp 29–68.

(6) Hong, S. C.; Lee, J. People on the verge of death: evidence from impacts of celebrity suicides. *Applied Economics* **2015**, *47*, 710–724.

(7) Beck, A. T. Depression: causes and treatment. **1967**,

(8) Diagnostic and statistical manual of mental disorders, 5th Edition. **2013**,

(9) Loewenstein, G. Projection bias in medical decision making. *Medical Decision Making* **2005**, *25*, 96–105.

(10) Becker, G. M.; Degroot, M. H.; Marschak, J. Measuring utility by a single-response sequential method. *Behavioral Science 9*, 226–232.

(11) Bushong, B.; Gagnon-Bartsch, T. An Experiment on Interpersonal Projection Bias. **2020**,

(12) Brinkmann, K.; Gendolla, G. H. Does depression interfere with effort mobilization? Effects of dysphoria and task difficulty on cardiovascular response. *Journal of personality and social psychology* **2008**, *94*, 146.

(13) Cléry-Melin, M.-L.; Schmidt, L.; Lafargue, G.; Baup, N.; Fossati, P.; Pessiglione, M. Why don't you try harder? An investigation of effort production in major depression. *PloS one* **2011**, *6*, e23178.

(14) Kaufmann, M. Projection bias in effort choices. **2021**,

(15) Loewenstein, G.; O'Donoghue, T.; Rabin, M. Projection bias in predicting future utility. *the Quarterly Journal of economics* **2003**, *118*, 1209–1248.

(16) Kocalevent, R.-D.; Hinz, A.; Brähler, E. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *General hospital psychiatry* **2013**, *35*, 551–555.

(17) Kroenke, K.; Spitzer, R. L. The PHQ-9: a new depression diagnostic and severity measure. 2002.

(18) Bolotnyy, V.; Basilico, M.; Barreira, P. Graduate Student Mental Health: Lessons from American Economics Departments. *Journal of Economic Literature* **2021**,

(19) Ridley, M.; Rao, G.; Schilbach, F.; Patel, V. Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science* **2020**, *370*.

(20) Chaloupka IV, F. J.; Levy, M. R.; White, J. S. *Estimating biases in smoking cessation: Evidence from a field experiment*; 2019.

(21) Becker, G. S.; Murphy, K. M. A theory of rational addiction. *Journal of political Economy* **1988**, *96*, 675–700.

Appendix

Figures and Tables[10]

Figure 1: Distribution of questionnaire scores assessing depressive symptoms

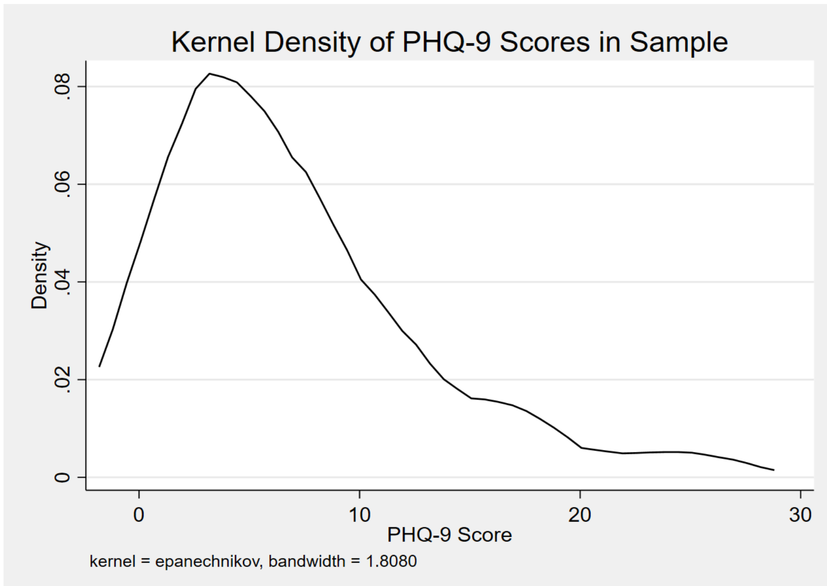
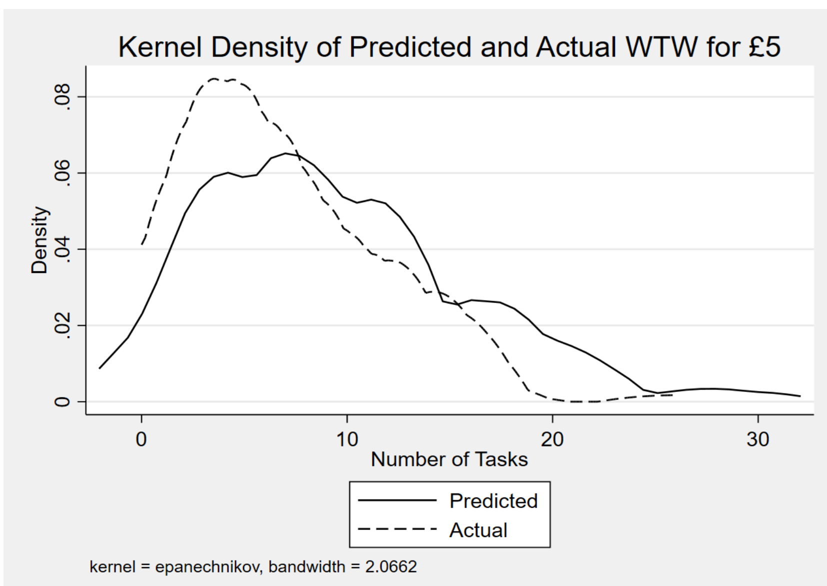


Figure 2: Distribution of elicited willingness to work



[10]The file with experiment dataset as well as the Stata Do-file to replicate all figures and tables can both be accessed at *this github repository.*

Table 1: Baseline characteristics

	All	PHQ-9≥10	PHQ-9<10	*p*-value
Demographics:				
Male (%)	27.8	19.4	31.0	0.19
White (%)	91.3	87.1	92.9	0.39
Age (mean)	34.1	32.5	34.7	0.23
Born in the UK (%)	90.4	90.3	90.4	0.99
Have children (%)	53.0	45.2	56.0	0.31
Number of household members (mean)	2.3	2.3	2.3	0.98
Highest ed. degree: high school (%)	25.2	22.6	26.2	0.69
Highest ed. degree: undergraduate (%)	40.0	45.2	38.1	0.51
Full-time job (%)	53.0	58.1	51.2	0.52
Not in paid work (%)	11.3	3.2	14.3	0.03
Household income:				
<£10,000 (%)	11.3	0.0	15.5	0.00
£10,000-£29,999 (%)	23.5	29.0	21.4	0.42
£30,000-£49,999 (%)	31.3	32.2	31.0	0.90
£50,000-£69,999 (%)	20.9	19.4	21.4	0.81
£70,000-£89,999 (%)	5.2	3.2	6.0	0.51
£90,000-£150,000 (%)	7.8	16.1	4.7	0.12
Socioeconomic status index (1-9)	5.2	5.6	5.1	0.10
Literacy difficulties, e.g. dyslexia, ADHD (%)	8.7	6.5	9.5	0.58
Student status (%)	16.5	16.1	16.7	0.95
Eyesight problems (%)	86.1	90.3	84.5	0.39
Prolific history:				
Number of approvals (mean)	124.9	146.4	117.0	0.32
Number of rejections (mean)	1.0	0.9	1.0	0.68
Completion duration, in seconds (mean)	1,296.8	1,300.9	1,295.3	0.96
Prolific score (0-100)	99.5	99.5	99.5	0.98
Number of observations	115	31	84	

Note: *p*-values are retrieved from the two-sided test of equality of means in PHQ-9≥10 and PHQ-9<10 groups. Experiment participants were restricted to UK citizens with English as a first language. The socioeconomic status is a self-reported variable representing the position on an imagined ladder with people who are best-off at the top and people who are worst-off at the bottom. Prolific score represents an approval rate and is calculated using the upper bound of the 95% confidence interval.

Table 2: Descriptive statistics. Willingness to work and proxy for tiredness.

	WTW(£5) (1)	WTW(£5)/5 (2)	WTW(£3)/3 (3)	WTW(£7)/7 (4)
Mean	6.73	1.88	1.39	1.26
	(4.90)	(1.25)	(1.14)	(0.87)
	$\widetilde{WTW}$(£5) (5)	τ_1 (6)	τ_2 (7)	$\%\Delta\tau$ (8)
Mean	9.38	4.37	6.23	-
	(6.23)	(2.29)	(2.16)	
PHQ-9 score	-	-	-	-0.41
				(1.76)
Number of observations	115	115	115	115
Individual controls	-	-	-	Yes

Note: Standard deviations are reported in brackets for variables (1)-(7) and a standard error - in brackets for model (8). The dependent variables in the upper panel represent the number of extra tasks a participant was willing to complete for the designated bonus payment; for variables in (2), (3) and (4), it is given in the per-pound form. $\widetilde{WTW}$ represents the predicted willingness to work. Variables (6) and (7) represent self-reported tiredness levels, which range from 1 to 10 points where 1 is "not tired at all" and 10 is "extremely tired". Model in (8) represents an OLS regression of the percentage change in reported tiredness level, i.e. $(\tau_2 - \tau_1)/\tau_1 \cdot 100\%$, on the PHQ-9 score and a vector of individual covariates. The coefficient for PHQ-9 represents the change in tiredness percentage growth associated with 1 point increase in the PHQ-9 score, which ranges from 0 to 27.

Table 3: Relationship between predicted and actual willingness to work

	W(£5) (1)	W(£3)/3 (2)	W(£7)/7 (3)
Behavior	−2.65	−0.49	−0.62
	(0.74)	(0.16)	(0.14)
Constant	9.38	1.88	1.88
	(0.52)	(0.11)	(0.10)
Number of observations	230	230	230
Individual controls	No	No	No

Note: The dependent variable is a stacked vector of predicted and actual WTW; for columns 2 and 3, it is given in the per-pound form. "Behavior" equals 1 for observed WTW and 0 for incentivized prediction, i.e. "Behavior" $\equiv \mathbb{1}_{\{\text{WTW}\}}$.

Table 4: The relationship between willingness to work (WTW) and depressive symptoms

Panel A	$\widetilde{WTW}$(£5)	$\widetilde{WTW}$(£5)	WTW(£5)	WTW(£5)
PHQ-9 score	−0.06	−0.03	−0.01	−0.03
	(0.10)	(0.10)	(0.08)	(0.08)
Constant	9.78	13.09	6.79	10.79
	(0.90)	(4.38)	(0.71)	(3.49)
Panel B	**WTW(£3)**	**WTW(£3)**	**WTW(£7)**	**WTW(£7)**
PHQ-9 score	0.01	−0.01	−0.01	−0.03
	(0.05)	(0.06)	(0.09)	(0.10)
Constant	4.08	7.56	8.90	14.55
	(0.49)	(2.48)	(0.89)	(4.32)
Number of observations	115	115	115	115
Individual controls	No	Yes	No	Yes

Note: The dependent variables represent the number of extra tasks a participant was willing to complete for the designated bonus payment; $\widetilde{WTW}$ represents the predicted willingness to work. The coefficient for PHQ-9 score represents the change in WTW associated with 1 point increase in the PHQ-9 score, which ranges from 0 to 27 points. Individual covariates include age, ethnicity, gender, highest attained degree, number of household members, marriage status, student status, vision issues, experiment duration and whether a participant has children, works full-time or belongs to the lowest income category.

Table 5: The relationship between projection and depressive symptoms

Approach A: Using estimated α in the original dataset	α(£5)	α(£3)	α(£7)
PHQ-9 score	−1.21	−1.41	−0.01
	(2.26)	(2.61)	(2.82)
Constant	−13.97	−10.62	−5.43
	(99.99)	(106.40)	(115.73)
Number of observations	107	111	102
Individual controls	Yes	Yes	Yes

Approach B: Using interaction in the stacked dataset	W(£5)	W(£3)/3	W(£7)/7
Behavior	−2.99	−0.59	−0.68
	(1.11)	(0.24)	(0.21)
PHQ-9 score	−0.06	−0.01	−0.01
	(0.09)	(0.02)	(0.02)
Behavior × PHQ-9 score	0.05	0.02	0.01
	(0.12)	(0.03)	(0.02)
Constant	13.44	2.87	2.69
	(2.80)	(0.60)	(0.54)
Number of observations	230	230	230
Individual controls	Yes	Yes	Yes

Note: The dependent variable in Approach A is a percentage difference between predicted and actual WTW; for columns 2 and 3, this parameter was estimated using per-pound WTW. The dependent variable in Approach B is a stacked vector of predicted and actual WTW; for columns 2 and 3, it is given in the per-pound form. "Behavior" equals 1 for observed WTW and 0 for incentivized prediction, i.e. "Behavior" $\equiv \mathbb{1}_{\{\mathrm{WTW}\}}$.

Table 6: Patient Health Questionnaire-9 (PHQ-9). Summary

	Not at all	Several days	> Half the days	Nearly every day
1. Little interest or pleasure in doing things (%)	45.22	40.87	8.70	5.22
2. Feeling down, depressed, or hopeless (%)	53.04	27.83	13.04	6.09
3. Trouble falling or staying asleep, or sleeping too much (%)	34.78	39.13	13.91	12.17
4. Feeling tired or having little energy (%)	18.26	43.48	21.74	16.52
5. Poor appetite or overeating (%)	43.48	34.78	14.78	6.96
6. Feeling bad about yourself or that you are a failure or have let yourself or your family down (%)	44.35	33.91	11.30	10.43
7. Trouble concentrating on things, such as reading the newspaper or watching television (%)	46.09	37.39	10.43	6.09
8. Moving or speaking so slowly that other people could have noticed. Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual(%)	80.87	13.91	3.48	1.74
9. Thoughts that you would be better off dead or of hurting yourself in some way (%)	85.22	9.57	2.61	2.61
Points in PHQ-9 score	0	1	2	3

Note: Each of the nine items contributes equally to constructing a PHQ-9 score, which ranges from 0 (all responses are “Not at all”) to 27 (all responses are “Nearly every day”). Item #4 had the highest average score of 1.37, while item #9 had the lowest score of 0.23 (range is 0-3).

D&PB - Experiment instructions

Start of Block: Overview and PHQ-9

Welcome!

The instructions below will provide guidance for the actual work you will do. First you will do several simple tasks and then you will predict how many more you would do for additional money after completing more mandatory tasks. Then we will also ask you about your actual willingness to do additional work.

You will earn at least £4. Depending on your ability to predict your future behavior you may earn more.

You must complete the entire session to earn any pay for this study.

The experiment session will consist of three parts:

1. You will complete a short 9-item questionnaire.
2. Then, you will do 3 mandatory tasks and predict your future willingness to do additional work for additional pay after you will have done more tasks.
3. You will do 12 more mandatory tasks and then be asked about doing additional work for additional pay.

When you see a screen saying **"THIS IS THE END OF THE EXPERIMENT"**, you will understand you have reached the end of the experiment. Please do not exit before you have seen this screen. On this last page of the experiment you will also find a Completion Code which you need to copy and enter in your Prolific account.

Page Break

Task

The task in this experiment involves counting specific characters in the matrices. You will see a matrix like in the image below:

2	2	2	2	$	$$	2	$$	2	$	$$	$$	2	$$	$
$$	$	$$	$$	2	2	2	2	2	2	$$	2	2	2	$
$$	2	$	2	2	$$	$	$$	$$	$	$	$	$	$$	$
$	$	$	2	$	2	2	$$	$$	$$	$	2	$$	2	2
$	2	$	$$	$$	$$	$	$$	2	$	$	$$	2	2	2
$$	$$	$	$$	2	2	$	$	$	2	$$	2	$	$	$
$	$$	2	$$	$	2	$$	2	$	2	2	2	2	2	$
2	$	$$	2	$$	2	$	$$	$	$$	2	$	$$	$	2
$	$	$	$$	$$	$	$	$$	$	$	2	$	$	2	$$
2	$	$$	$$	$	$$	2	$$	$	$$	2	2	$$	$	2

You will then be asked to count a specific character that is present in the image. For example, the question might read: How many '2' symbols are there in the matrix? This means you have to count how many cells with a '2' symbol there are in the matrix. The symbol that you will need to count will change in each task, so pay close attention to instructions.

Also, note that we have included two characters which are very close to each other: '$' and '$$'. These are different, so when you are asked to count '$, do not count cells with '$$'. Similarly, when you are asked to count '$$', do not count cells with '$'.

You must type in the exact correct number in order to advance to the next matrix. Counting each matrix should take about 30-40 seconds.

Page Break

ID
First things first. Please insert below your Prolific ID.
It should look something like 60fh7yff781r4pq4551b6rw2

This is to make sure we send you the right amount of money.

My Prolific ID is:

____

Page Break

PHQ-9

In this part of the study, please complete a short questionnaire on depression symptoms.

Over the last 2 weeks, how often have you been bothered by any of the following problems?

	Not at all (0)	Several days (1)	More than half the days (2)	Nearly every day (3)
Little interest or pleasure in doing things (1)	○	○	○	○
Feeling down, depressed, or hopeless (2)	○	○	○	○
Trouble falling or staying asleep, or sleeping too much (3)	○	○	○	○
Feeling tired or having little energy (4)	○	○	○	○
Poor appetite or overeating (5)	○	○	○	○
Feeling bad about yourself - or that you are a failure or have let yourself or your family down (6)	○	○	○	○
Trouble concentrating on things, such as reading the newspaper or watching television (7)	○	○	○	○
Moving or speaking so slowly that other people could have noticed. Or the opposite - being so fidgety	○	○	○	○

or restless that you have been moving around a lot more than usual (8)				
Thoughts that you would be better off dead or of hurting yourself in some way (9)	○	○	○	○

End of Block: Overview and PHQ-9

Start of Block: I1

NOW YOU WILL START THE FIRST BLOCK OF WORK WITH 3 TASKS.
WHEN YOU CLICK TO ADVANCE TO THE NEXT SLIDE, YOU WILL BEGIN.

End of Block: I1

Start of Block: I2

THIS IS THE END OF BLOCK 1.
NOW WE WILL ASK YOU SOME QUESTIONS BEFORE YOU START BLOCK 2.
CLICK TO ADVANCE TO THE NEXT SLIDE.

End of Block: I2

Start of Block: Tiredness & Prediction

Before you go to completing 12 more mandatory tasks, we will now ask you to make a prediction about your future behavior.

Specifically, we will ask you to **predict your willingness to do additional work for additional payment after you are done with all mandatory tasks**. That is, once you have completed 12 more tasks we will ask how many additional tasks you will be willing to do for some amount of extra money.

Think carefully before making your prediction on the next page. If your guess is within 3 tasks of your actual future answer, you will receive £0.50 as a bonus.

Page Break

Think about you **who has just completed all fifteen mandatory tasks**.

What do you think is the maximal number of additional tasks you will be willing to complete for extra £5?

	0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
How many additional tasks? ()	

On average, how tiring did you find the first three tasks on a scale from 1 to 10, where 1 is "not tiring at all" and 10 is "extremely tiring"?

	1 2 3 4 5 6 7 8 9 10
How tiring were the tasks? ()	

End of Block: Tiredness & Prediction

Start of Block: I3

NOW YOU WILL COMPLETE 12 MORE TASKS IN THIS BLOCK OF WORK.
WHEN YOU CLICK TO ADVANCE TO THE NEXT SLIDE, YOU WILL BEGIN.

End of Block: I3

Start of Block: I4

THIS IS THE END OF BLOCK 2.
CLICK BELOW TO ADVANCE TO THE NEXT SLIDE.

End of Block: I4

Start of Block: Tiredness & WTW

Now, we will ask you three similar questions about your **willingness to complete more tasks for additional money**.

The tasks remain the same: counting symbols in matrices.

Pay attention to the amount of bonus payment in the next questions. Please think carefully before making your choices on the next three pages!

Question #1. What is the maximal number of additional tasks you're willing to complete now for **£7**?

	0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
How many additional tasks? ()	

Question #2. What is the maximal number of additional tasks you're willing to complete now for **£5**?

	0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
How many additional tasks? ()	

Question #3. What is the maximal number of additional tasks you're willing to complete now for **£3**?

	0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
How many additional tasks? ()	

Page Break

On average, how tiring did you find the last twelve tasks on a scale from 1 to 10, where 1 is "not tiring at all" and 10 is "extremely tiring"?

	1	2	3	4	5	6	7	8	9	10
How tiring were the tasks? ()										

IMPORTANT!
COPY AND PASTE THE FOLLOWING COMPLETION CODE IN YOUR PROLIFIC ACCOUNT TO GET PAID.
DO NOT CLOSE THIS PAGE BEFORE YOU HAVE COPIED THE CODE BELOW:

8FBE2ABE

THIS IS THE END OF THE EXPERIMENT
THANK YOU FOR PARTICIPATING!

End of Block: Tiredness & WTW