

**An Online Research Environment
for Behavioural Studies into
Reinforcement Learning**

Rayo Verweij

Master of Science
Cognitive Science
School of Informatics
University of Edinburgh
2021

Abstract

In each decision we make, we need to deal with some uncertainty about the outcome of our choice. The field of Reinforcement Learning offers methods that can be used to model the effects of uncertainty on the human decision-making process, and how these effects differ in groups with specific personality traits, such as high anxiety or autism spectrum disorder. To create these models, we need a lot of behavioural data. To facilitate this data collection, this project has created an online research environment for use in the University of Edinburgh that implements the popular two-armed bandit task for investigating human decision-making. In addition, the environment includes a researchers' dashboard that gives researchers full control over the parameters of the study as well as easy access to the gathered data. The platform was tested and evaluated in a thorough pilot study and a basic RL model was created to illustrate the experiment-to-analysis pipeline.

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr Peggy Seriès, for her continuous support and feedback during this project, as well as for introducing me to the other members of her research lab. In a year like this, building and fostering some sense of community is both incredibly hard and more important than ever, and your work with the lab is very much appreciated. Secondly, I cannot give enough thanks to Filippo Ferrari, for his feedback, support, and near-constant availability if I ever needed help. I hope you will put this project to good use, but knowing you, I don't think that will be a problem. Finally, a thousand thanks to everyone who participated in the pilot study, and took the time to provide the valuable feedback that formed and validated this project. I might not be able to pay you the thousands of pounds you accumulated in the game show, but know that my appreciation for you transcends any form of tangible currency.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Rayo Verweij)

Table of Contents

1	Introduction	1
2	Background	3
2.1	Decision-making and uncertainty	3
2.2	Anxiety	5
2.3	Autistic Spectrum Disorder	6
2.4	Measuring decision-making	7
2.5	Models of reinforcement learning	9
3	Project Goals and Design	11
3.1	Project goals	11
3.2	Experimental design	12
3.3	Platform requirements	13
4	Platform Architecture and Implementation	16
4.1	Hosting	16
4.2	Database	17
4.3	Dashboard	18
4.4	Experiment	21
5	Evaluation	26
5.1	Qualitative feedback	27
5.2	Quantitative modelling	31
5.2.1	General statistics	32
5.2.2	Implementing an RL model	34
6	Discussion	37
	Bibliography	40

A Participant information sheet and consent form	46
B Pilot questionnaire results	50
C Installation guide	56
D Adding questionnaires	59

Chapter 1

Introduction

The world is an uncertain place: whatever action you take, or decision you make, knowing the exact outcome for sure is virtually always impossible. While this keeps life interesting, it can also induce anxiety, especially if the results of your decisions might cause harm and you're trying to avoid that.

In Informatics, studying how an actor has to make decisions to achieve positive outcomes and avoid the negative ones is the subject of the field of Reinforcement Learning, and besides teaching computers how to do this, insights from this field can be applied to humans as well. A successful computational model that abstracts human behaviour allows us to poke at it, changing parameters to see what results we get, and possibly lead us to new insights about ourselves. Several successful computational models that abstract the behaviour of different groups of people allow us to compare them, to not just understand ourselves but also each other. In case of psychiatric or developmental disorders, specifically, knowing more about how individuals make the decisions that they do moves us from describing symptoms to describing causes, which can lead to not only better understanding but also better treatment.

Creating these models of human decision-making can only be done if we have a lot of data on human decision-making, however, as that is what we want to base the models on. Gathering a lot of data is hard, and that is where this project comes in. We have built a comprehensive online research environment that implements a web-based version of the two-armed bandit task, a popular tool for measuring human behaviour, as well as a database that securely stores all gathered data and a researchers' dashboard that gives researchers full control over the variables of the experiment without having to dive into the code.

In this document, I will describe how this platform came to be. In Chapter 2, we

will take a look at how one would build a model of decision-making, how it would be applied to anxiety and autism spectrum disorder, and how one would gather data for such a model. In Chapter 3, we will set out the goals and requirements of this project, and discuss exactly what kind of behavioural experiment we have implemented. In Chapter 4, we will go into detail about how the platform was built, and in Chapter 5, we will see the results of a pilot study that was conducted to test the platform. Finally, in Chapter 6, we will discuss the overall state of the platform as it stands at the time of this writing, and look at some recommendations for how to use it going forward.

Chapter 2

Background

We will start by taking a look at the work that has been done so far to create models of human decision-making. First, we will discuss the general theory behind these models, and how this theory adapts to different groups of people. Second, we will examine the types of experimental designs used to gather data about human decision-making. Finally, we will look at some specific models in detail, and see how they use the data collected in behavioural experiments. Combined, this will form the basis of both this project's implementation and our approach to evaluating it.

2.1 Decision-making and uncertainty

How do we make the decisions that we make? In our modern understanding, decision-making is usually formalised as an application of Bayes' theorem in the brain [1]. In a Bayesian model, the probability that a certain outcome happens based on some data depends on both the *likelihood* that the observed data would occur if the outcome did indeed happen as well as any *prior* beliefs about the outcome, that is, a *bias* towards it. When we learn new words, we have to decide to assign certain meanings to certain sounds or characters by using both our estimate of how likely it is that that meaning actually fits the word as well as our prior beliefs of how language works [2]. When we play tennis and our opponent serves, we have to decide on when to swing by combining our sensory estimate of how fast we think the ball is going as well as our prior knowledge of how serves usually play out [3]. This process is universal, and happens across the brain.

Conveniently for our existence, the process does not stop when the decision is made: usually, we compare our expected outcome with the actual outcome and update

our prior beliefs if necessary. This is one of the major ways in which we learn things. Specifically, *Reinforcement Learning* (RL) describes how actors learn and change their decision-making processes over time in response to receiving rewards or punishments for their actions [4, 5]. Computational models of RL, which we are working towards here, can then be used not only to try and teach machines how to learn, but also to describe and analyse human behaviour.

Two basic parameters sit at the core of these RL models: the *learning rate* α and the *inverse temperature* β [6, 7]. The learning rate, usually a value of $0 < \alpha < 1$, is a weight representing to what extent recent feedback is used to update the estimated outcomes of the different actions. That is, higher values of α lead to faster updating of one's predictions. The inverse temperature, usually a value of $0 < \beta < 10$, represents to what extent the estimated outcome of an action is used in deciding which option to choose. Very high values of β mean the actor almost always picks the option which is estimated to lead to the better outcome, whereas lower values of β allow for more exploration of different options. This is known as the *exploration-exploitation trade-off*. For humans, it has been suggested that as we grow up and get older, we get better at adapting learning rates to different situations, whereas at the same time, inverse temperature rises as we become less likely to explore and 'take risks' [7].

Usually, learning to make sound decisions is not as easy as knowing when and how to integrate the data at our disposal, however, as often this data is incomplete. As we noted: the world is an uncertain place. Of course, the entire reason why biases are a necessary part for the Bayesian model to work and why one might not always want to maximise inverse temperature is that in the face of uncertain outcomes, it is useful to be able to use prior experience, and it is good to sometimes try the less-valued outcome, just in case the situation around us has changed.

We can distinguish two types of uncertainty: expected and unexpected [8, 9, 10]. We speak of *expected uncertainty* when the decision-maker is aware that the outcomes of their action might vary; for example, when picking a colour on a roulette wheel, we know that our choice has a certain probability of giving a reward. In case of the roulette wheel, we can precisely calculate what each probability is, but often, there is some degree of noise in the possible outcome probabilities. Following [11], we will call this noise *unpredictability*. If a decision's outcome is very unpredictable, it will have little to say about any future decisions. As such, if someone detects that a series of decisions is very unpredictable, their learning rate will go down [11, 12].

On the other hand, *unexpected uncertainty* refers to changes in the environment,

say, if the casino suddenly decides to change the rules of the game while you're playing. To describe how and how often the environment changes, we will use the term *volatility* [11, 13, 14], though it has also been known as *hazard rate* [15]. In a landmark study, Behrens et al. [16] found that humans are able to accurately assess volatility and adapt their decision-making accordingly. In contrast to unpredictability, a high level of volatility means prior beliefs need to be updated more often and thus a higher learning rate.

In some cases, there might be other types of uncertainty to account for—for example, *perceptual uncertainty* arises if one is not certain of one's sensory experiences and has reason to question the perceived data itself [10]. While we will briefly touch upon this when discussing the accessibility of our research experiment in Section 4.4, it is not relevant to the mathematical models we are building here.

With the necessary terminology defined, it becomes interesting to look at how estimations of these variables differ between different groups of people. For example, we've already seen the current theories on how learning rate and inverse temperature change with age, and in addition, it has been suggested that old age weakens the ability to process information about unpredictability [15]. For this project, we are specifically interested in how the integration of information about unpredictability and volatility changes in individuals with traits of anxiety and/or ASD, for reasons that we will explore next.

2.2 Anxiety

Typically, we distinguish between two types of anxiety: *state anxiety*, a temporary condition as a result of dealing with an adverse situation in the moment, and *trait anxiety*, a long-term condition and personality trait affecting one's daily life [17, 18, 19]. In addition, anxiety disorder is one of the most common psychiatric disorders, affecting an estimated 7-8% of humans worldwide [20].

The cause of anxiety (in general, not anxiety disorder specifically) is, of course, uncertainty. Not being certain about the outcomes of our actions is what causes us to worry about them, sometimes to the point where our anxiety 'takes over' and we worry too much. Not only does uncertainty lead to anxiety, anxiety also influences the processing of that uncertainty back. In what way exactly is what we want to find out here.

For behavioural experiments, various questionnaires have been developed that al-

low participants to self-report their level of anxiety, of which Spielberger's State-Trait Anxiety Inventory [19] is the most commonly used [21]. The STAI consists of two parts of twenty questions each, one assessing state anxiety (form Y-1) and the other assessing trait anxiety (form Y-2). Note that the STAI is not a diagnostic tool, as participant self-scoring can be inaccurate and there is no trained professional present to evaluate and compare the results. The STAI merely gives an indication of roughly where on the spectrum the participant falls.

Several avenues of research into the influence of anxiety on the decision-making process have been explored. Recall that learning rate, the weight given to new information, should increase in environments with high volatility. Browning et al. [13] confirmed these findings for individuals with low trait anxiety, but found that this updating process is impaired in high trait anxiety, speculating that a deficiency in the ability to use current environmental information to correct assess actions' potential outcomes could be a major cause for developing anxiety disorders.

This is slightly nuanced by Aylward et al. [22], who found that anxious individuals actually have a higher learning rate specifically for potential punishments. This would mean that not only is most new information given less weight than in an optimal Bayesian model, there is an additional fixation on potential negative outcomes.

Pulcu and Browning [9] speculate that similar deficits could be present in the processing of expected uncertainty in anxiety, that is, assessing the unpredictability of the outcomes. No empirical research has yet been done to measure whether this is present or not and there is almost no literature looking at the interplay of expected and unexpected uncertainty in anxiety. Ultimately, answers to these questions could help us determine whether misestimation of uncertainty is simply a symptom of anxiety, or one of its root causes.

2.3 Autistic Spectrum Disorder

Autism Spectrum Disorder¹ describes a wide range of related developmental disorders exhibiting similar but differently presenting deficits in social communication and preferences towards repetitive behaviours [24]. While a proper diagnosis of ASD is an involved and expensive process [25], for the purposes of short behavioural experiments,

¹We wish to acknowledge that within the autism community, there is no single agreed way to describe ASD [23]. For the purposes of this project, we will use the terminology as defined by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, or DSM-5.

Baron-Cohen et al. [26] developed the Autism-Spectrum Quotient (AQ), a 50-item questionnaire that assesses to what extent the participant displays traits of ASD, with reasonable accuracy [27].

Investigating exactly how traits of ASD impact decision-making is important as individuals with ASD often report both more issues with as well as higher reluctance to making decisions than neurotypical controls [28]. Due to the nature of the disorder's most prominent symptoms, however, the vast majority of research into decision-making and ASD has traditionally focussed on the use of emotional context in the process [29, 30].

One major complication in trying to uncover more fundamental differences decision-making is that frequently, individuals with ASD also display higher trait anxiety [31], making it difficult to distinguish exactly where the cause of any effects lie. For example, South et al. [32] suggested that children and young adolescents with ASD were focussed on avoiding punishments rather than seeking rewards in a decision-making task more so than their neurotypical peers, but as these effects are similar to what we've seen about anxiety, it is unclear whether they are directly caused by ASD.

At the same time, and in contrast to what we saw in anxiety, Lawson et al. [14] showed that adults with ASD have a higher learning rate in highly volatile environments, therefore placing a lot more focus on environmental changes compared to their prior beliefs than neurotypical controls. This starts to get at the kind of information we need to understand how to update our behavioural models to apply to ASD. Just as with anxiety, however, there has been no research yet into any potential differences in assessing unpredictability.

2.4 Measuring decision-making

So far, we have seen what we need to build an accurate RL model of human decision-making, and to what extent the current literature has or has not already done this. Next, we are going to discuss the methodology used to gather data for these models.

The most common instrument to measure human decision-making in the above studies has been the *bandit task* [12, 13, 22]. The classic version of the task imagines an actor having to choose between different slot machines in a casino (sometimes called "one-armed bandits", leading to the name *multi-armed bandit task*) in a limited amount of time while unsure of exactly how the probabilities that a machine will give a reward change over time. The challenge, of course, is to gain as high a reward as possible,

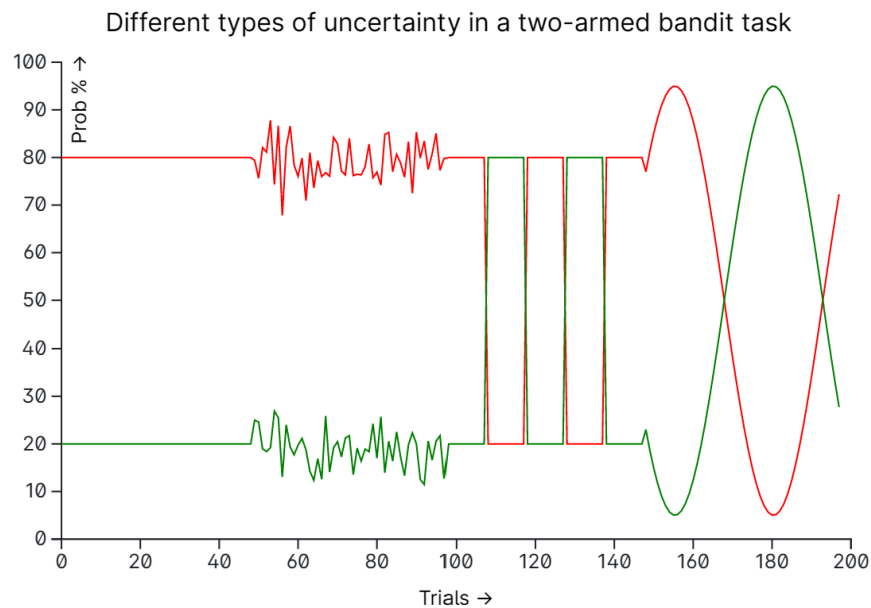


Figure 2.1: An example two-armed bandit task, with four different types of trials. For the first 50 trials, no uncertainty is present, as the probabilities are static. In the second block, while the environment is stable, unpredictability is represented as Gaussian noise in the probabilities. In the third and four blocks, no unpredictability is present, but the environment is volatile. First we see change point volatility, where the probabilities swap a certain number of times, and secondly we see smooth volatility, where the probabilities go back and forth over time.

for which the actor needs to optimise learning rate—how do you weigh new versus old information—and inverse temperature—do you stay with the option that seems to work or do you try another one that may be better [33, 34].

While originally cast as a machine learning problem, Acuña and Schrater [35] showed that human behaviour on a multi-armed bandit task can be described using Bayesian models, and Speekenbrink and Konstantinidis [36] formalised a version of the task designed for behavioural experiments that for the first time included expected uncertainty. In summary, for a set number of trials a participant has to choose one of several options. Each option has the probability of giving a reward and the goal is to maximise gains. Unpredictability is present in that the participant never exactly knows how high the probability of getting a reward is, and volatility is present in that the probabilities of getting a reward changes over time. Both types of uncertainty are illustrated in Figure 2.1. By creating multiple environments, we can see how participants respond in each of them, and compare them to the optimal Bayesian model [12].

2.5 Models of reinforcement learning

When we have then collected data on how human participants perform on these tasks, there are two types of computational models that we can build. Neurobiological models use this data to examine how information is processed at a cortical level to try and understand more about the hardware of the brain [33, 37]. Here, we focus on models of behaviour, to see how different people respond to different environments on a macro-scale.

For simplicity, let's focus on a bandit task with only two arms. In a basic RL model [5, 6, 38], we want to give both options a score of how good they are, $V_1^{(t)}$ and $V_2^{(t)}$, where t represents the current trial. We set the initial values $V_i^{(0)} = 0$ and can calculate the score for each following option as follows:

$$V_i^{(t+1)} = V_i^{(t)} + \alpha \times (r^{(t)} - V_i^{(t)}) \quad (2.1)$$

where $r^{(t)}$ is the reward that we got that trial (1 for receiving a reward, 0 for not receiving anything, -1 for receiving a punishment, if present) and α is the learning rate parameter, described above. The value of $r^{(t)} - V_i^{(t)}$ is also known as the *prediction error*.

Then, at each trial, we can model which choice our actor will make by calculating a *softmax* function for each option, which defines the probability P that option i becomes choice c :

$$P(c = i)^{(t)} = \frac{\exp(\beta \times V_i^{(t)})}{\exp(\beta \times V_1^{(t)}) + \exp(\beta \times V_2^{(t)})} \quad (2.2)$$

where β is the inverse temperature parameter.

The challenge in creating these models, then, is to find values of α and β that best describe the type of participant you are training on. We do this by *minimising the log-likelihood* of the parameters for a sequence of choices C by the participant, as follows:

$$\text{NLL} = - \sum_{c \in C} \log(P(c)) \quad (2.3)$$

Broken down, we calculate the likelihood that a set of parameters works for a series of choices, we take the likelihood's logarithm to make the outcome of the likelihood function easier to work with, and feed that into an optimisation function that does this for a bunch of different sets of parameters in order to find the best one. We take the *negative log likelihood* specifically because optimisers in statistical packages, such as MATLAB's `fminunc` [39] or SciPy's `optimize` [40], tend to work by minimising rather than maximising their arguments.

While constructing this model, it is important to perform a process called *parameter recovery*. In parameter recovery, you use the parameters you chose to generate some simulated data, fit a new model to this simulated data, and look at the correlation between the old and new parameters. If they are similar, we can be confident that the model-fitting process went well, and that the parameters can be recovered independently—that is, that the parameters influence the model regardless of each other.

When we are confident about our parameters, we can finally compare them between different groups of people. If there are any significant differences, this may give us specific insights into how the decision-making process changes across groups.

Chapter 3

Project Goals and Design

As we saw in the last chapter, researchers are still in the process of figuring out the influences of different disorders on how uncertainty is handled in the decision-making process. Specifically, at the time of this writing, Filippo Ferrari of the Seriès lab at the University of Edinburgh is investigating unpredictability and volatility processing in anxiety.

Gathering the human data necessary for this research is a complicated process. To invite enough participants to a physical laboratory is not only very time-consuming but also rather inadvisable due to the current COVID-19 pandemic. The preferable option, therefore, is to gather data online. This project aims to facilitate that, by building an online research environment for behavioural studies into reinforcement learning.

3.1 Project goals

The goal of this project was to build an online research platform that

1. implements a two-armed bandit task for gathering data into human decision-making;
2. saves and processes the gathered data for easy access by the researchers; and
3. allows researchers to easily create different variations of the task that can be assigned to different groups of people,

for the immediate benefit of the research of Filippo Ferrari and as a template for others looking to implement their own behavioural experiments.

We will first take a look at the requirements for the experiment itself, before describing the requirements of the platform as a whole.

3.2 Experimental design

Per the requirements of Filippo Ferrari's research, we want to implement a two-armed bandit task where both unpredictability and volatility can change over time, such as described in Section 2.4 and Figure 2.1. The experiment consists of two parts. First, after reading the Participant Information Sheet and signing the Consent Form, the participant is asked to fill out a personality questionnaire, such as the STAI or the AQ (see Sections 2.2 and 2.3). After this, the decision-making task begins.

While the past decade has seen a huge rise in experiments being conducted online and for these types of project they are now usually seen as a valid alternative to in-person testing [41, 42], there are still some disadvantages. The most pressing one for us is that online, the researcher has no control over the participant's environment. They can very easily be distracted, cannot ask any questions to the researchers, and have not even met the researchers, which might mean they feel less motivated to perform well [43]. In addition, the task we are asking them to do is very repetitive and likely to be boring to at least some of the participants, which might decrease performance even further.

To combat this, several previous implementations of the bandit task provided their participants with the narrative that they were in a casino-setting, giving rewards in terms of (fake) money, and encouraging participants to try and uncover the casino's changing strategies to try and beat them at their own game [22, 33]. Originally we planned on writing a similar narrative, but changed it to that of a game show early in development, to avoid any negative connotations that participants might have with casinos and gambling in general.

Thus, when starting the decision-making task, the participants are told that they are participants in a game show, where they continuously have to choose between two doors, each with their own colour. Each door has a certain probability of giving a fake monetary reward, and their goal is to gain as much money as possible. However, so they are told, the game hosts will change their strategy over time, so the participants will need to keep paying attention to the types of rewards they are receiving. An early mock-up of what a trial would look like is shown in Figure 3.1. The exact number of trials, the reward probabilities of each door and how they change over time, and which questionnaire is given at the start should be configurable by the researchers, and we will detail each required variable in the next section.

Participant recruitment also needs to happen online. Nowadays, there are several



Figure 3.1: A rough mock-up of a task trial that was created in the exploratory phase of the project. A participant chooses the yellow door and receives a reward, adding to their total at the bottom. After we settled on the game show narrative, the design of the doors was kept throughout development, even though other elements changed.

large platforms available that allow researchers to specify a demographic to market their study to, of which we chose Prolific [44], as it came recommended by multiple independent studies [45, 46]. Prolific is a huge boon to us as it completely abstracts the participant sign-up and reward process, not only saving us time, but also ensuring complete anonymity for the participant. When they select our study to participate in, they are redirected to our experiment along with a unique ID, and after they finish the experiment, they are returned to Prolific which then takes care of payment. In addition, Prolific allows for setting some restrictions on who is able to take part, for example allowing us to only take participants that have a large enough screen to properly complete the study (and are not on their phones).

One final thing to mention about the experimental design is that several other studies supplement their behavioural data with pupillometry for a different perspective on participants' responses [13, 14]. As we are conducting the experiment online, it is sadly not possible for us to collect this type of data.

3.3 Platform requirements

To build a platform that achieves the goals above, three parts are needed: the experiment, a database to store the collected information, and a front-end for the researchers

to use so when they want to retrieve the data or create variations of the study, they don't have to dive into the actual code. We call this part the Dashboard.

In this section, I'll lay out the requirements for the platform as established at the start of the project. In the next chapter, we will see what the final product looks like, and in Chapter 5, we will reflect upon these requirements and see how well the requirements have been satisfied.

The Database must...

- Store all participant data, anonymised, on a secure University server where only the relevant researchers have access.

The Experiment must...

- Seamlessly integrate with Prolific to allow recruited participants to take the experiment and be compensated for doing so;
- Offer the Participant Information Sheet and Consent Form in clear language, and approved by the Informatics Ethics Committee;
- Implement a variety of questionnaires, including at least the STAI and AQ, and an easy way for researchers to add their own;
- Implement the two-armed bandit task with support for all variables specified below;
- Include a short training section for the bandit task to familiarise participants with the task procedure;
- Be designed to be as engaging as possible through the decision-making task, with an overview of the total reward gains on each trial and graphics to support the game show-narrative;
- Be accessible to as wide a range of participants as possible, adhering to the current Web Content Accessibility Guidelines of the W3 World Wide Web Consortium [47] and usable for participants with visual impairments such as colour blindness, dyslexia, or motion sickness, and specific input requirements such as keyboard-only users;

- Be performant, with short loading times and clear feedback if the platform does need to load;
- Completely guarantee participants' privacy and be compliant with GDPR.

The Dashboard must...

- Offer a list of all participants of the experiment and a way to download the data associated with them;
- Allow researchers to create different variations of the experiment and generate a URL for each of them that can be linked to Prolific;
- Allow for selecting which of the implemented questionnaires appears at the start of the experiment;
- Allow for determining how many different blocks of trials each experimental condition contains and whether they appear in a set or random order;
- For each block of trials, allow for determining its length in trials, the amount of fake money that is given as the reward, the base reward probabilities for each door, whether or not Gaussian noise is applied to the reward probabilities, and if applied, the standard deviation of that noise;
- For volatile blocks of trials, allow for determining how often the probabilities change between the two doors, and in case of smooth volatility, set a minimum and maximum probability;
- Serve a visualisation of the reward probabilities over time for each created condition (similar to the graph in Figure 2.1);
- Be easy and unambiguous to use and navigate for all users;
- Be performant, with short loading times and clear feedback if the platform does need to load;
- Only be accessible to authorised researchers;
- Be well-documented and adaptable for any similar projects besides Filippo Ferrari's research.

Chapter 4

Platform Architecture and Implementation

The research environment is composed of three, relatively distinct components: the Database, the Dashboard, and the Experiment itself. I will start with a brief overview of where they are hosted before moving on to discuss each of them in detail.

The code for both the Dashboard and the Experiment is attached. For a guide on how to install the application, see Appendix C. For detailed information on exactly how the code works, comments are provided throughout the codebase.

4.1 Hosting

To ensure secure storage for the gathered data and compliance with privacy regulations, the entire application is hosted on the University's servers instead of a commercial platform. While this imposes several limitations on the kinds of technology that can be used, it is a necessity, especially since we are collecting data on people's mental health.

The School of Informatics offers several options for hosting content, the most secure of which is their sweb ("secure web") service. While most of the School's hosting services use one shared ID when the server accesses a user's files, on the sweb service, each user is given their own ID, meaning the application is safe from (accidental) tampering by others on the University network. All content is hosted at https://sweb.inf.ed.ac.uk/<user_id>/.

The only option in terms of server technology on the sweb service is to use the School's Apache server and to interface with it using PHP. This does not always play

nicely with the JavaScript-based Experiment and Dashboard front-end described below, though as we will see, it does mean that the functionality for creating the front-end and interfacing with the database are neatly separated.

4.2 Database

The School of Informatics' only database service uses PostgreSQL, an advanced relational database system based on the SQL paradigm. One convenient aspect of PostgreSQL is that it includes a `json` datatype, which simply allows any JSON object to be stored as a value. Given the Experiment records its data in JSON format, this means that the experimental data can be copied into the database with barely any need for processing.

The experimental data is stored in a the *experiments* table containing three fields: the primary key `id` (type `variable character`; length 128) which stores the participant IDs; `timestamp` (type `timestamp without time zone`) which stores the time at which the record was inserted; and `data` (type `json`) which stores the entire data object as constructed by the Experiment.

By default, using SSL to secure the connection between an application and the School's PostgreSQL service is not allowed by Informatics Computing Support, meaning it is possible for third parties to listen in when data is sent from the Experiment to the database or from the database to the Dashboard. At my request, however, to ensure the security of our data, SSL was enabled. The Installation Guide in Appendix C stresses that anyone who is installing the application themselves ask the same.

Only the experimental data is stored in the permanent database. The different experimental conditions created on the Dashboard are not stored on the server, but locally, in the browser's `localStorage`. This means that every user can create their own conditions without the fear of one being accidentally changed by another person. It does also mean that any created conditions are lost if the user switches browsers or clears their `localStorage`, but as conditions are easy to create, this should not be a large issue. It is advised, however, to keep an overview of one's created conditions separately from the application.

4.3 Dashboard

The Dashboard contains two areas: one for creating conditions (Figure 4.1) and one for viewing the collected data (Figure 4.2). It is written in ReactJS augmented with TypeScript. ReactJS [48] is a popular JavaScript library for writing modular user interfaces and TypeScript [49] is an extension to JavaScript that adds a typing system to the language, for easier debugging. Both libraries are open-source and widely used and they were chosen because I was already highly familiar with them before starting the project. Additional technologies used include Sass [50], an extension of CSS that allows for more flexibility; Bootstrap [51], a library of pre-built CSS components that I used as the basic building blocks of the UI; and Axios [52], a JavaScript library that facilitates making HTTP calls. Each of these libraries is fully open-source as well.

To access the Dashboard, the user needs to login with their DICE account through the School's Cosign service. This prevents access from anyone who is not part of the School of Informatics. Because of a quirk in Cosign, when a user logs in to the Dashboard, opens a second browser window, closes the browser window that contains the Dashboard, and then navigates to the Dashboard again, Cosign might 'remember' that the user previously logged in and provide access to the Dashboard without actually logging in the user. In this case, the Conditions page is fully accessible, but the Data page will display an error as it cannot retrieve the data if the user is not properly authenticated. If this happens, simply reloading the page will prompt Cosign to ask the user to log in again.

On the Conditions page, each condition is listed with options to change its variables and copy the URL to the experiment, as well as a graphical preview of the experiment's timeline. As mentioned, all conditions are stored in the browser's localStorage and can therefore always be accessed through the browsers developer tools for debugging purposes. General variables that need to be set for each condition include the choice of starting questionnaire, the number of blocks, the door colours, and whether or not the blocks should be presented in random order. Six different door colours are available for selection. For each block, the user can then select its type (training, static, change point volatile, or smooth volatile), the number of trials, the reward amount in British pounds, the reward probabilities for each door in percentages, and the standard deviation of the noise applied to the probabilities of each trial, also in percentages. Volatile types include the option to set the number of changes in the probabilities and, lastly, the smooth volatile type also includes the option to set a minimum and maximum for the

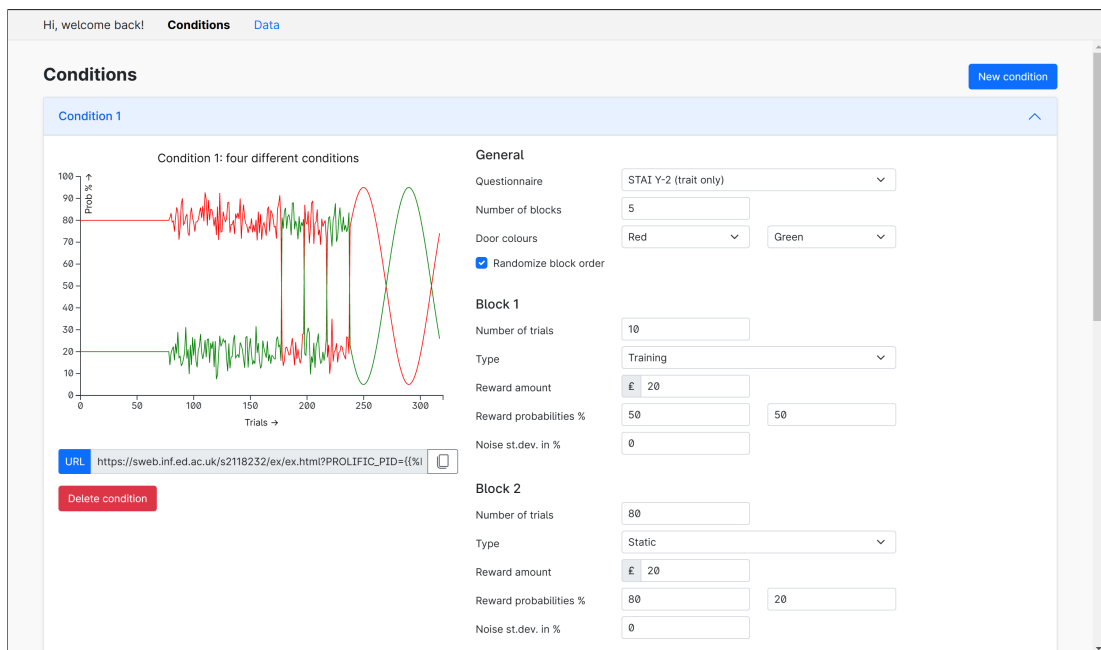


Figure 4.1: A screenshot of the final Dashboard, on the Conditions page. Conditions are presented in a accordion-style list where each item can be opened and collapsed. On the left, a visualisation of the condition is present, as well as the URL to be copied. On the right, each variable can be set.

reward probabilities. Taken together, these include every variable needed as described in Section 3.3, and their exact implementation in the Experiment is described in more detail in Section 4.4 below.

Each variable input is cleaned and checked thoroughly to ensure only valid input is accepted. The number of blocks can range between 1-9, reward amounts, probabilities, and noise levels have to be numbers between 0-99, and the number of trials per block can vary between 1-999. These ranges have to be strict as these numbers are used to create a unique URL for each condition.

A condition URL consists of three parts: the base address, the Prolific variables, and the condition variable. The base address is `https://sweb.inf.ed.ac.uk/<user_id>/ex/ex.html/`. Next, Prolific assigns three variables: `PROLIFIC_PID`, the unique participant's ID, `STUDY_ID`, the study's ID, and `SESSION_ID`, the session's ID. Of these three, the participant's ID is the only one relevant to us, as it becomes the key for each data record in the database. The other two are only needed for Prolific to function. When testing, `STUDY_ID` and `SESSION_ID` can be assigned any value, and `PROLIFIC_PID` can be assigned a value of "R" to make the Experiment generate a random, 12-digit alphanumeric ID.

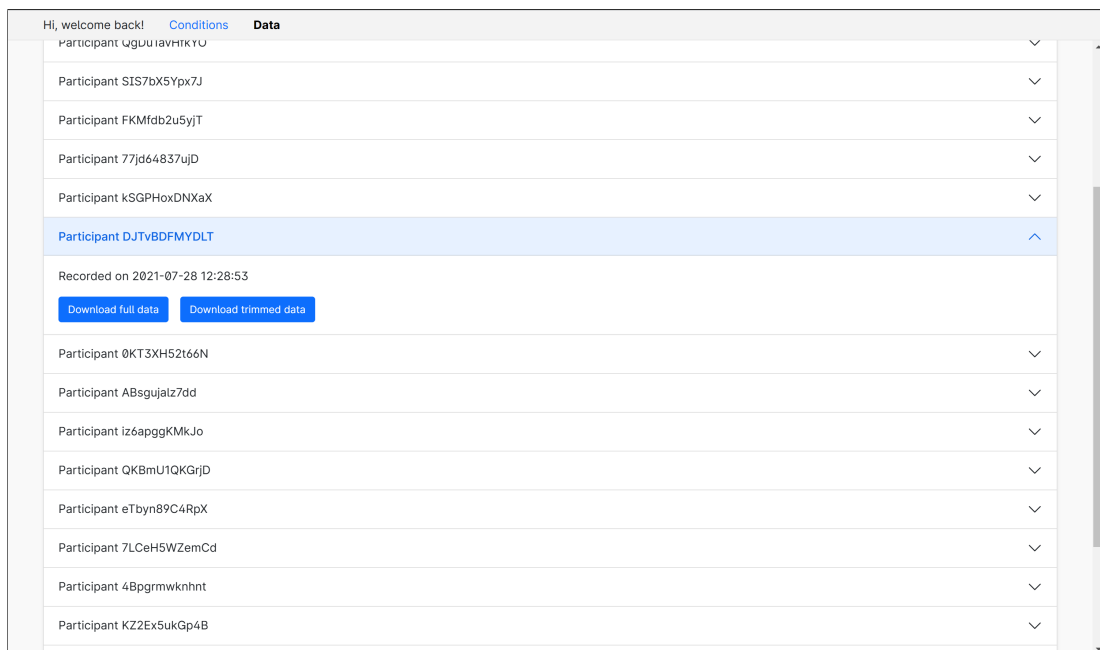


Figure 4.2: A screenshot of the final Dashboard, on the Data page. Selecting a participant shows the options of downloading the full or the trimmed data object.

The condition variable, simply called C in the URL as to not attract too much attention to it, is a long string containing each variable as described above. The first five characters contain the general variables: the questionnaire type number, the number of blocks, two letters that each are the first letter of one of the door colours, and a 0 or 1 for whether the block order is randomised or not. After that, a string of number is added for each block, separated by dashes. For example, the string 080320802010040595 represents a block of 080 trials, of type 3 (volatile), with a reward amount of 20 pounds, base probabilities of 80 and 20 percent, a noise standard deviation of 10 percent, 04 changes in probabilities (in the case of the smooth volatile condition, maxima of the sinusoid), and a minimum and maximum probability of 05 percent and 95 percent, respectively.

The variables are encoded in the URL like this as both Prolific and the Experiment make extensive use of URL variables to record data. Someone with extensive knowledge of the experiment will be able to deduce all variables from looking at the URL, but it should be impossible to decipher for someone who has never seen it before. In addition, this method makes it easy to quickly change conditions while testing the platform.

The graphical visualisation of the condition is constructed using the open-source JavaScript data visualisation library D3.js [53] and is updated in real-time as the vari-

ables are changed. The option to randomise the block order does not affect the graph, and it will list all blocks except for any training blocks in the order that they were created. Random noise is generated based on the standard deviation variable that can be set for each block, and the graph allows a researcher to instantly see what effects the noise level they picked will have. The title of the graph can be edited, so that a researcher writing a report can simply take a screenshot of this graph to use. In fact, Figures 2.1 and 5.1 were both created with this function.

The Data page connects to the Database to list all data that has been gathered so far. Under the hood, instead of the page requesting data from the server itself, the React code uses the Axios library to make a HTTP request to a separate PHP script that requests the data. This two-step process means that the Dashboard will render regardless of whether or not it was able to make a connection to the Database. This is highly beneficial in cases such as the Cosign quirk described earlier, as now, the Data page can list that an error has occurred (with detailed error reporting being logged to the console), as opposed to the application simply crashing without any indication of what went wrong.

Each record has two options for accessing the data: downloading the full JSON object or downloading a trimmed version. jsPsych, the library used for the Experiment described below, saves virtually everything to a JSON object called `data`, including “fake trials” such as the Participant Information Sheet and the stimulus for each trial, which in our case is simply the HTML code describing the doors. The clean option provides a JSON object with only the data relevant to analysis: reaction times, doors chosen, which of those were correct, probability values, whether a reward was given, the total reward at the end, the total time spent on the experiment, and the answers to the questionnaires.

Advanced operations on the data, such as renaming or deleting records, are not provided on the Data page as they are not expected to be needed very often. In case an advanced operation is necessary, the user can access the PostgreSQL database directly from any DICE machine as described in the Computing Support manual.

4.4 Experiment

The Experiment is written using the open-source JavaScript library jsPsych [54], which allows its users to define an experiment as an array of trials of different types. In jsPsych, every screen a participant sees is a trial, including for example the welcome

For the first part of the experiment, a number of statements which people have used to describe themselves are given below. Read each statement and then select the appropriate option to indicate how you *generally* feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

I worry too much over something that doesn't really matter.

almost never sometimes often almost always

I feel inadequate.

almost never sometimes often almost always

I read carefully, and therefore select 'almost always'.

almost never sometimes often almost always

I am a steady person.

almost never sometimes often almost always

I feel rested.

Figure 4.3: A screenshot of the questionnaire screen of the Experiment, seen here with form Y-2 of the STAI, which assesses trait anxiety.

screen and the Participant Information Sheet. jsPsych allows the developer to attach complex logic and data to each trial and automatically takes care of common functions such as measuring reaction time. While jsPsych experiments run in a web browser, instead of an application native to, for example, a lab computer, research has shown that there is no significant difference in data collection between the two [55, 56], and multiple studies mentioned earlier have already successfully deployed jsPsych [41, 43].

When the Experiment is loaded, the script parses the URL generated with the Dashboard described above to extract all the variables. Using a unique, invisible type of trial provided by jsPsych, it then loads the images of the doors that will be needed for the experiment. Preloading the images at the start of the experiment means that there will be no loading screens or performance lag once the experiment has started. When the Experiment has loaded, the participant has to press a key to start, at which point the starting time is recorded.

The Experiment starts with the Participant Information Sheet and participants are able to download and save it as a PDF if they so wish. After this, the initial questionnaire is presented (Figure 4.3). Four different questionnaires are implemented: a placeholder, the AQ, the full STAI, and a partial STAI that only includes the half about trait anxiety. The placeholder does not actually contain any questions, and can be used to avoid having to fill out a questionnaire over and over when testing the platform.



Figure 4.4: A screenshot of the trial screen of the Experiment, after the participant has made their choice and the door has opened to reveal a reward.

Appendix D includes a guide on how to add more questionnaires to the application.

The interesting stuff happens during the game show. After a brief explanation, the participant gets a continuous stream of game trials as described in Section 3.2 (Figure 4.4). The participant has to use the “F” key to select the left door and the “J” key to select the right door. These keys were chosen because on most keyboards, these keys have a small bump on them, making them easily identifiable.

Between each block, the participant is shown a pause screen, which encourages them to take a brief break if they would like to. Blocks are presented either in the order that they were created in the Dashboard or, if the option is selected, in a random order. The only exception to this is that a block of type Training is always put first, regardless of its position on the Dashboard. (If, for some reason, there are multiple Training blocks, only one of them gets pulled to the front. It’s not advised to have multiple Training blocks, however.)

Behind the scenes, the core of this part of the Experiment is a function called `game_trial`. It takes in a `block` object containing all variables of a block; a randomly chosen `correctResponse` of either “F” (left) or “J” (right), which determines on which side of the screen the door with the higher reward probability will be shown; a boolean `swapped` which indicates when the reward probabilities should be swapped in a change point volatile environment, and two `precalculated` values which contain

the reward probabilities in a smooth volatile environment. First, this function adds noise to the probabilities if a noise level is set and swaps them if `swapped` is true. Secondly, it randomly decides for both doors whether a reward will be given if they are selected, based on their reward probabilities. Finally, it generates the trial for jsPsych, putting the doors in their positions and “hiding” the reward messages behind them.

Indeed, this means that a door’s success is determined not when the participant selects one, but *at the very start of the experiment*, when all trials are generated. This is a result of how jsPsych operates, as it is not possible to directly reference and update any HTML elements from jsPsych. To enable the door opening animations I was able to manually edit the source code of jsPsych to add a short delay when the participant makes a selection and to hard-code the CSS changes that enable the animation, but to actually update the trial’s logic to change when a participant makes a choice (which is necessary to determine in the moment whether they will receive a reward or not) was not possible. In practice, this means that if a participant opens their browser’s developer tools and navigates to the appropriate spot, they can see exactly what will happen when they select either door. While it is probably safe to assume that participants won’t do this (and if they do, a 100% success rate can easily be filtered out of the results), it does constitute an architectural flaw as a result of using jsPsych.

For now, back to the timeline. In essence, the game show is generated by iterating over each block and calling `game_trial` as often as there are trials in the block. For training and static environments, it simply calls the function with a random `correctResponse` each time. For a change point volatile environment, it uses the number of changes variable to divide the block into a series of smaller blocks of equal length, alternately setting `swapped` to true or false. For a smooth volatile environment, the base reward probabilities as defined in the Dashboard aren’t sufficient, as here the probabilities vary as a function of time. Therefore, we first generate two arrays as long as the number of trials, one for each door, containing evenly spaced out values on the sinusoid starting at the base probability, changing direction as often as set with the number of changes variable, and translated to fit the minimum and maximum variables, all set on the Dashboard. The base algorithm for calculating these values was written by Filippo Ferrari and adapted by me for JavaScript—see the attached code for the exact implementation. Having generated these arrays, we then iterate over them, feeding each value as `precalculated` into `game_trial` to produce the jsPsych trial.

Finally, between each block, we insert a pause trial that allows the participant to take a breather. At the end of the experiment, we display the participant’s total reward

in a congratulatory message, before saving the data and sending them back to Prolific. In the pilot study, we also included a second questionnaire at the end with questions about the experiment, as detailed in Section 5.1.

Saving the data happens similarly to retrieving the data on the Dashboard, that is, by making a HTTP call to a separate PHP file that encodes the data and connects to the Database. This means that even if something goes wrong with saving the data, the participant is still sent back to Prolific to receive their reward, instead of having the experiment crash on them.

Several efforts have been made to make the Experiment as accessible as possible. Text is presented in a reasonably large font size and with high contrast, and the door colours were chosen to be contrasting even for participants with colour blindness. While usually an animation plays when a door opens to reveal whether a reward will be given, if a participant has disabled animations in their operating system or browser (that is, if the flag `prefers-reduced-motion` is set), no animations are shown and the reward simply appears. Finally, even on screens that require confirmation with a button press, the entire Experiment is navigable using only a keyboard, if for some reason a participant cannot use a mouse or does not have one present. In terms of the computational models we are creating, this all serves to minimise any chance of perceptual uncertainty arising, which is otherwise hard to control given that the experiment is conducted remotely. These functions, in addition to some minor changes made as a result of feedback detailed in Section 5.1, should mean that the Experiment is accessible to at least the vast majority of potential participants.

Chapter 5

Evaluation

When all main functions described in the previous chapter were implemented, we held a pilot study to test the platform. Other members of the Seriès lab as well as some direct friends were invited to take part. They did not participate through Prolific and as such did not receive any compensation; however, we did ensure their anonymity by generating a random ID for each of them when they started the experiment. The pilot study was certified according to the Informatics Research Ethics Process, RT number 6014, and the complete Participant Information Sheet and Consent Form are attached in Appendix A.

For this pilot study we used 5 blocks of trials: first a training block, after which plain static, noisy static, noisy change point volatile, and noisy smooth volatile blocks were presented in random order. The training block lasted for 10 trials and had 50-50 reward probabilities. The other blocks all lasted for 80 trials with 80-20 base reward probabilities and a standard deviation of 10% for the blocks with noise. The change point volatile environment had 3 changes, whereas the smooth volatile one had 4, with a minimum base probability of 5% and a maximum of 95%. Finally, door colours were pink (80) and green (20), reward amounts were set at £20 per correct trial, and the questionnaire for this pilot study was the trait-focused part of the STAI. A visualisation of the study condition can be seen in Figure 5.1.

We received responses from 17 participants in total, enough for plenty of feedback and up from the initial goal of 8-10, as most people that were reached out to decided to take part, which we highly appreciated. The complete set of data gathered during the study is included in the extra materials, attached. During the pilot study, we collected two types of data: qualitative comments by means of a questionnaire at the end of the experiment, and the quantitative results from the experiment itself. In this chapter,

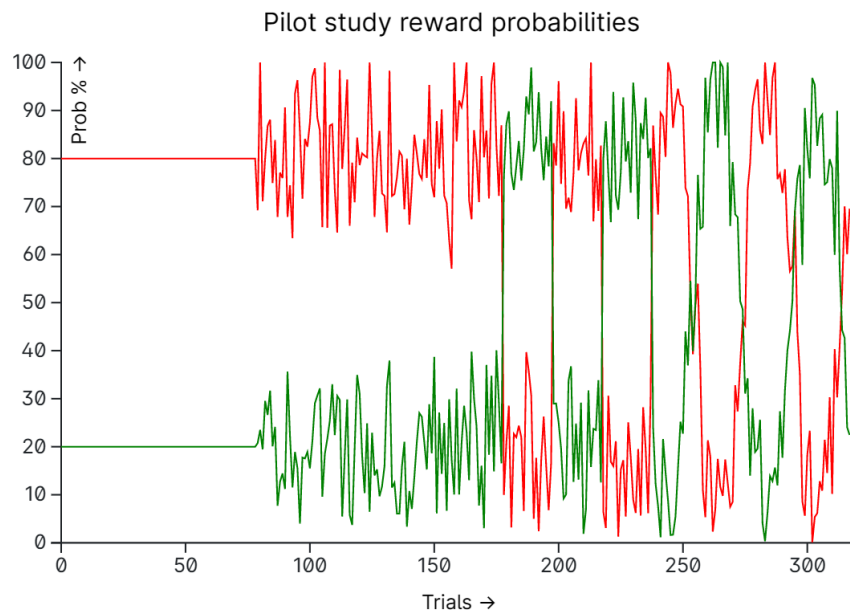


Figure 5.1: A visualisation of the reward probabilities in the pilot study over time, not including the training block. The order of blocks was randomised, so not each participant got the blocks in this exact order. Noise is randomised on the spot during the experiment; the noise pictured here is purely illustrative.

we will first discuss the qualitative data to see how the experiment was perceived by participants and what changes were made in response to them. Secondly, we will look at the quantitative data to illustrate how other studies can use this in order to answer their research questions.

5.1 Qualitative feedback

Unique to this pilot study was a second questionnaire, at the very end of the experiment, asking six questions about the participants' experiences. I will briefly summarise the responses to each question as well as discussing some extra feedback that some of the participants volunteered. All responses are included in the attached data as well as listed in Appendix B for easy access.

How did you feel during the game show? (E.g. excited, bored, indifferent, etc.)

All participants fell clearly into one of four groups. Five participants wrote they were engaged by the game. One wrote, "At first I was indifferent about it, but then I felt quite intrigued by it and eager to discover whether I chose the right door throughout

the entire game,” whereas another “tried to get in the head of whoever made this” to find an optimal strategy. One specifically mentioned it “was quite exciting seeing the money building”, even though they knew the money was fake.

Seven participants indicated that while they started out interested, they lost interest as the game went on. One mentioned they lost interest after the second round, one during the final round, and the others did not indicate when they started to get bored. Specifically, one writes “it felt like there was no way of knowing which door was more likely (...) so it felt like I was randomly guessing” - though this person in a subsequent answer also mentioned that they did not know whether probabilities were linked to door colour or position, which indicates that either the instructions were unclear or that they did not read them. Another person, convinced that the probabilities only changed after a pause trial, suggested the graphics should be replaced in each block (“[for example] a window instead of a door”) to keep things more interesting.

Finally, three participants simply indicated that they were bored throughout the experiment, with no other comment, and two said they felt indifferent.

Did you have any issues while doing the experiment? Were the instructions clear? Is there anything that would have made the experience better?

No major issues were reported, a solid sign that the experiment works without bugs or crashes across a range of browsers and devices, and the few issues that were raised were all related to quality-of-life. Two participants mentioned they would have liked to see a countdown saying how many trials were left. One would have liked the “F” and “J” keys below the doors to be bigger and another did not like that they had to use the keyboard at all, and wanted to be able to click on the doors. Ironically, that was the original implementation of the experiment (as can be seen in the design concept of Figure 3.1), but we switched to keyboard input as it tends to reduce reaction times.

One salient piece of feedback I should note is that one participant indicated that as a non-native speaker of English they did not know some of the words used in the STAI questionnaire, in this case, *turmoil* and *inadequate*. No direct changes were made to the questionnaire in response to this one person, but it is important to mention that any researcher using this platform should consider whether to only recruit native and/or near-native English speakers for their experiment, and whether this decision could have any influence on their study’s outcome.

How long did you (approximately) spend on the entire experiment? Did it feel long, short, or fine?

The initial goal was for the 320-trial experiment to not take longer than 30 minutes, so it was heartening to see that virtually every participant reported a time of 10-15 minutes, with only one writing 20. The major difference lied in how they perceived this time, with 3 reporting it as short, 3 as fine, and the remaining 11 as long. Note that while the Participant Information Sheet they received mentioned the experiment would take up to 30 minutes, their invitation to participate stated it was more likely to be around 15, based on my own testing before starting the pilot. Depending on how well they read either, participants might have had different expectations of how much time it would take them before starting. Regardless, given that most who ended up bored indicated they started to lose interest in the second half of the experiment, perhaps a total number of trials of 240 or 270 would work better.

How did you decide which door to choose? Did you use a specific strategy? Did your strategy change over time?

By far the most popular strategy, explicitly mentioned by 6 participants, was to continuously change doors every time they got 2 or 3 fails in a row. Some decided, likely because of the distribution in the static blocks, that the pink door gave more rewards and stuck with that one for most of the experiment. Some did not decide on a strategy at all, one only ever picked the green door “with no particular reason”, and one decided to keep their cards close to their chest: “Yes, I use a strategy. No, I haven’t changed it because it works”.

There is an interesting range of strategies here, with an equally interesting range of total reward outcomes. For more on this, see the section on quantitative modelling, below.

The probabilities of getting a reward associated with each door changed throughout the experiment. (Roughly) how often do you think they changed?

The main reason for this question, especially in combination with the above one, was to see whether people noticed that there were different blocks and whether they would update their strategy after each pause trial. Of course, this is hard to notice, as any noisy or volatile blocks have a lot of internal variation already (and, given that three out of four blocks were noisy, technically the correct answer to this question would be

241).

Only two participants specifically mentioned they thought the probabilities updated after the breaks. Most thought they changed 3-5 times (but did not mention the breaks), a few thought 2-3 times each block, and some outliers included “every 8 trials” and “every 20 [trials]”. In general, comparing the responses of this question with the previous one, it seems that most participants did not meaningfully update their strategy in between blocks.

Is there anything else you’d like to share about your experience?

Most had nothing to share, some wrote a thoughtful message of encouragement, one caught a typing mistake in the Participant Information Sheet, and a few had some comments that are worth sharing.

One good suggestion was to include a diagram of a trial in the instructions, to make the explanation clearer. Given that a few were either confused with the instructions or did not read them very closely, this could be a useful way to increase performance.

Two participants - one in an answer to this question and (probably) another in a private conversation - found out that the browser’s developer tools can be used to cheat on the experiment, as described in Section 4.4. As almost all participants in the pilot study were sourced from the School of Informatics, it is not unexpected that some would use their skills to take a closer look at the inner workings of the experiment, and an average participant is probably less likely to find this out. Still, it is important to be reminded that this technically constitutes a flaw in the design.

Implemented changes

Several changes to the platform were made after the pilot study, directly inspired by the feedback that was received. Besides the questionnaire discussed above, important feedback was also gathered from Dr Seriès and Filippo Ferrari on a continuous basis and from a particularly exhaustive email from Nikitas Angeletos Chrysaitis, one of the participants in the pilot study, for which we’d like to thank him.

In the experiment preamble, the Participant Information Sheet was updated to reflect that the experiment lasts roughly 20 minutes instead of 30—5 minutes more than what most participants reported it took to allow for some leeway. In addition, each of the implemented questionnaires had an attention check added for easy filtering of lackluster responses.

The explanation of the experiment was given a major overhaul with clearer text, a larger focus on the importance of door colour, and a graphic of what a trial looks like to reduce chances of confusion.

As for the trials themselves, the duration of the door opening animation was shortened from 1 second to 0.5 seconds to reduce waiting times and potentially take away some of the boredom. Doors in the training block are now assigned random colours that are different from the colours used in the main trials, to reduce the chance of introducing a bias during training. Finally, the images of the doors were made slightly smaller so the trials fit better on smaller screen sizes, whereas the “F” and “J” keys below each door were updated to be slightly larger and easier to read.

Within the time frame of the project, it was not possible to act on every idea and piece of feedback, and some pieces were rejected. For a wider discussion of some of the feedback, considerations on how to make the experiment more exciting, and recommendations for future features, see Chapter 6.

5.2 Quantitative modelling

In Chapter 2, we discussed the theory of implementing an RL model of human behaviour based on the data gathered from experiments like these. To illustrate the pipeline of going from collecting raw data to interpreting results, in this section, we will walk through some basic statistical analysis of the pilot study data as well as an implementation of the Rescorla-Wagner model of learning.

Before we start, it must be stressed that the data in this section is of no statistical significance whatsoever and does not constitute actual scientific results. Not only is a sample size of 16¹ not large enough to draw any conclusions, but also, as we already discussed, not every participant was properly engaged throughout and as this was a pilot study not all might have answered truthfully on the initial questionnaire. However, this analysis is not without merit. Firstly, it was an important exercise for me as developer to identify exactly what data needed to be saved, and to make sure that the data that was collected was compatible with the models created by Filippo Ferrari. Secondly, it serves to illustrate the relevance of the entire project, which we so far have only talked about in a theoretical capacity.

¹Of the 17 responses, one was not included in the quantitative analysis as they only ever picked the green door, rather than following the instructions of the experiment. This strategy did not work particularly well for them—their total reward was £2540, a clear outlier from the range of £3200–4760 the others found themselves in.

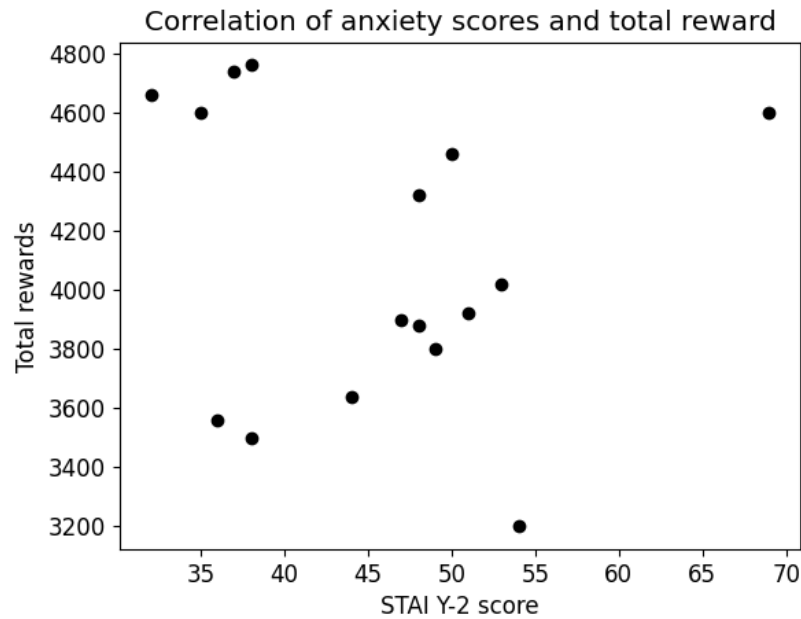


Figure 5.2: A scatter plot of the anxiety score and total reward of each participant. Pearson’s correlation coefficient r is -0.14 ($P = 0.60$), indicating no significant correlation between these values. However, if the participant with anxiety score 69 is seen as an exceptional outlier and removed from the equation, the correlation is somewhat stronger, at -0.45 ($P = 0.09$). If this correlation were to also show when more data is present (and with $P < 0.05$), this would mean that on average, participants with lower anxiety scores perform slightly better.

5.2.1 General statistics

The participants’ STAI Y-2 scores ranged from 32 to 69, with an average of 45.6 and a standard deviation of 9.1. This is high, as the average for a general population tends to lie around 10 points lower [19]. Reasons for these high scores among the pilot population might include the large proportion of participants that was writing their dissertations, the ongoing COVID-19 pandemic, or simply untruthful completion of the questionnaire.

The total rewards achieved ranged from £3200 to £4760 over 320 trials, with an average of £4097.50 and a standard deviation of £482.64. This means that on average, the participants received a reward in 64% of the trials. As can be seen in Figure 5.2, between these 16 participants, there was no clear correlation between anxiety score and total reward.

In terms of response time, one participant took a 38-minute break during one of

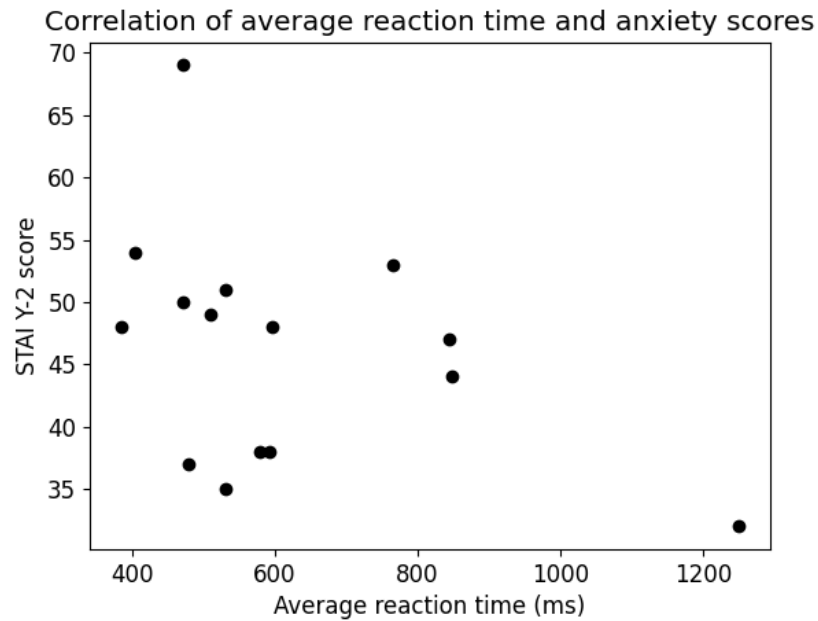


Figure 5.3: A scatter plot of the average reaction time and total anxiety score of each participant. Pearson's correlation coefficient r of the 14 participants *not* including the outlier at 1250ms is -0.15 ($P = 0.60$), indicating no correlation.

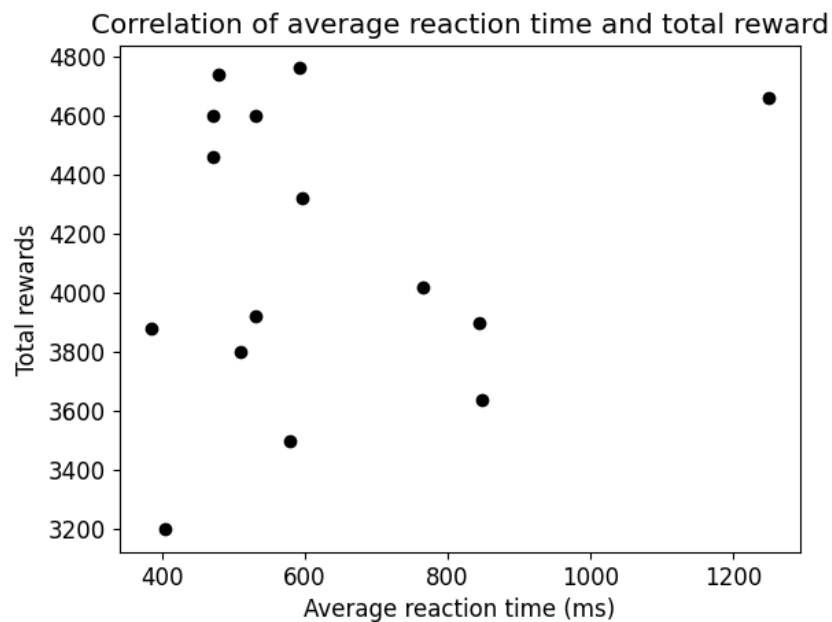


Figure 5.4: A scatter plot of the average reaction time and total reward of each participant. Pearson's correlation coefficient r of the 14 participants *not* including the outlier at 1250ms is -0.14 ($P = 0.64$), indicating no correlation.

the decision trials instead of one of the scheduled breaks, which caused their average response time to be so high they are not included in the following evaluation. As for the remaining 15 participants, the average response time was 617 milliseconds, with a standard deviation of 219 milliseconds. Given that this is not very low, it can be argued that at least on average, this is more evidence that participants were taking a brief moment to look at the positions of the doors and actively decide which one to pick, instead of pressing a random key as quickly as possible.

As can be seen in Figures 5.3 and 5.4, there are no clear correlations between average reaction time and either anxiety scores or total rewards, respectively. If these are also absent in larger sets of data, it would suggest that there is little influence of taking more or less time to decide on one's performance, nor that individuals with higher STAI scores take longer or shorter than others.

5.2.2 Implementing an RL model

Finally, we can look at the results of feeding this data in an RL model. The main benefit of using a behavioural model over using general statistics such as those in the section above is that a good model allows us to componentise a participants' behavioural patterns into easily interpretable parameters. Recall from Section 2.1 the core parameters α , the learning rate, and β , the inverse temperature. Differences in these between conditions and groups allow us to actually quantify how traits such as anxiety and ASD affect learning (α), the trade-off between exploration and exploitation (β), as well as derived parameters, such as response to feedback (whether α changes more or less after rewards or errors).

Here, we will implement the Rescorla-Wagner model. Originally created as a model of classical conditioning to account for the blocking effect [57], it can serve us well as a simple example model of learning [5, 38]. Refer to Section 2.5 for a theoretical introduction to the procedure. All code for implementing the Rescorla-Wagner model was provided by Filippo Ferrari, while the code for importing the data from the JSON files into the model, performing the statistical tests, and visualising the results was written by me. Who wrote what exactly is delineated in the attached code itself.

There are three main elements to building and analysing the model. First, usually before the data collection process has even begun, it is important to perform parameter recovery, as discussed in Section 2.5. When data has been collected, we need to find the values of α and β leading to the model that best captures their behaviour. Finally, we

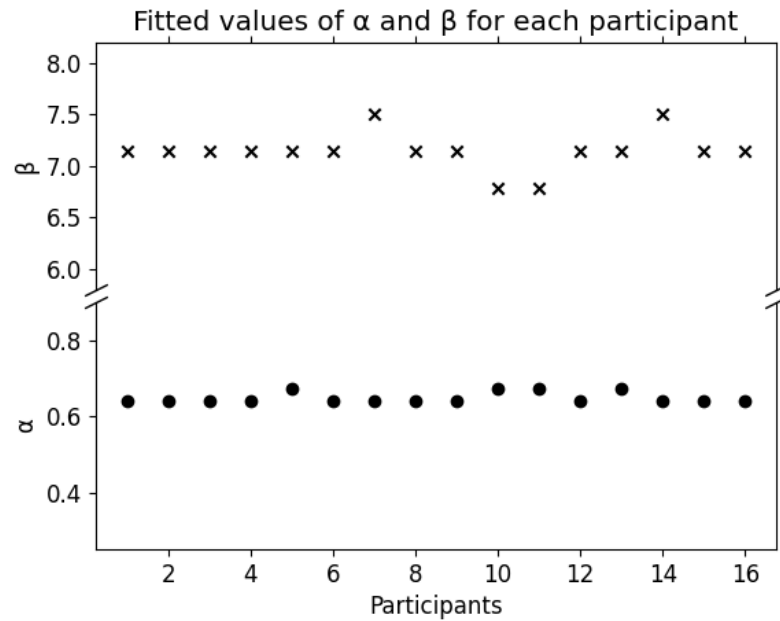


Figure 5.5: Fitted values for α and β for each of the 16 participants. Most values stayed at their initial estimates of 0.64 and 7.14, but a few diverted.

can divide the participants into different groups, such as low-STAI and high-STAI or low-RT and high-RT, and compare the parameters in each via standard statistical tests. Due to the tiny and uncertain nature of the data set, running these types of analyses on the pilot study data ends up being mostly nonsensical, but as an example, we will perform the second part, model fitting.

To find the optimal parameters for each participant, we define a function that calculates the negative log likelihood that a given set of parameters fits the data, pick some initial values for α and β to start off with, and feed this into a gradient-based optimisation function that automatically tries to find which parameters have the highest likelihood to fit. In this case, the initial value for α was 0.64 and β was 7.14, based on an exhaustive grid search by Filippo Ferrari with the aim to find the set of values that were giving the highest amount of rewards, using data from [12].

The calculated values can be seen in Figure 5.5. As can be seen, most of the values did not stray far from the initial estimates. Note that, for performance reasons, the optimisation function in this example was run with relatively large steps in the gradient, leading to the values seeming to ‘snap’ to the initial values and certain points around those. When a large data set is present, this function should be run with higher precision, to enable the researcher to tease apart finer differences between participants.

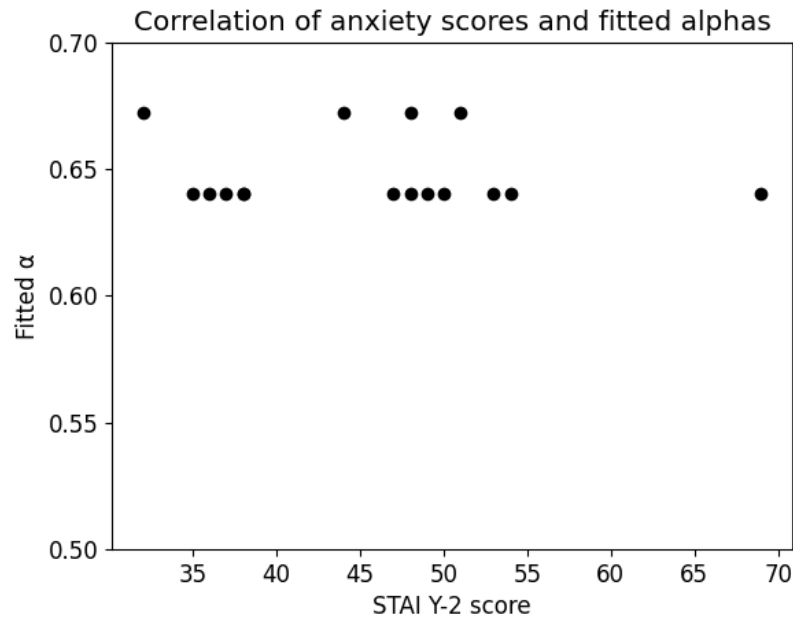


Figure 5.6: A scatter plot of the anxiety scores and fitted values of α related to them. No meaning can be attached here due to the small nature of the data set, but this serves as an illustration of the types of visualisations that the research platform enables.

Finally, it is also possible to investigate the correlations between α and β and the different performance metrics and questionnaire scores, as illustrated in Figure 5.6. The correlation between the anxiety score and α , for example, would be able to give us an idea of whether higher level of anxiety have any effect on (or at least are correlated with) different rates of integrating new information.

Chapter 6

Discussion

This project has delivered a comprehensive platform for building different variations of the two-armed bandit task and distributing them to potential participants. As mentioned, Appendix C includes a guide on how to deploy it and Appendix D includes a guide on how to add new questionnaires. While we mostly discussed questionnaires assessing traits of anxiety and ASD in this project, these are only two of many conditions that might influence the decision-making process in one way or another. Variations of the bandit task that we implemented here have been used in settings from depression [58] to schizophrenia [59] and the existence of this platform as a template for others to work with will hopefully enable many more of these kinds of investigations within the University.

Qualitative evaluation showed that the platform fulfills all the requirements outlined in Section 3.3 with no outstanding features or remaining bugs, and Chapter 4 has already discussed the various measures for security and privacy and optimisations for accessibility and performance. The modular nature of the attached codebase as well as its thorough documentation should open the door for others to create any variations they might need.

The biggest remaining question is that of engagement. As we saw in the results of the pilot study, the narrative and graphics of the game show were sufficient to keep most participant at least reasonably entertained, but this is an area where much can still be improved. Let's look at a few options.

The easiest option to tweak is the length of the study. Thankfully, the pilot's 320-trial run did not take as long as initially expected, and since the pilot, the length of the door animations has been shortened, meaning the total study is already slightly shorter as well. In addition, it is up to the researcher to decide whether 4 blocks of 80

trials are really necessary, and if it is not enough to only have 3 different variations of uncertainty, or else blocks of 60 or 70 trials. This is not a question that I can answer here, as it depends on the exact models used to analyse the data, but it is important to think about as it is very possible that a shorter study may take away much of the current feeling of boredom.

A second option, which includes editing the Experiment code rather than simply updating variables in the Dashboard, would be to give participants something concrete to work towards. This can be as simple as providing a countdown of how many trials are left, or a more involved system that rewards participants based on their performance. For example, the researcher could calculate the expected reward of each block, and at each pause moment in the task tell the participants how their score compares to the ‘average’ score. Alternatively, Prolific allows for paying participants a bonus based on performance, so the promise of an extra monetary reward might provide an extra boost to motivation that keeps participants engaged.

These kinds of reward systems need to be carefully considered, however, as they can easily backfire. If a participant sees that their score is far below the expected score, they might completely lose interest and stop caring altogether, and especially when working with highly anxious individuals, the experiment should not cause any distress or discomfort. My main recommendation would be to first do some testing with a shorter experiment, and to only consider more impactful options if a significant number of participants still indicate a low level of engagement.

A final suggestion we received in the feedback for increasing engagement was to have the experiment be held in fullscreen mode, to remove distractions on the participant’s device. Sadly, while jsPsych does include an easy to implement way to enter fullscreen mode, this does not work in the Safari browser, as Safari does not allow keyboard input in fullscreen. While it is technically possible to disallow Safari users from taking the experiment, and telling them to copy and paste the URL into a new browser, there is no guarantee that they have multiple browsers installed, nor should we want to introduce more friction in the participation experience. For now, we will have to rely on the participants themselves to tune out any distractions as much as possible.

It’s worth mentioning a few features unrelated to engagement that were not implemented due to time constraints but might be worth considering for future versions of the platform. jsPsych requires the developer to choose between mouse or keyboard input for each trial, but given that one respondent indicated that they would prefer to use mouse input, it might be worth it to build a custom type of trial that allows for

both. While the experiment explanation has been written to be as clear as possible, and the training block allows participants to become familiar with the procedure, one could additionally implement one or two questions about the procedure to check that the participants are familiar with the rules they were told. An example of this could be *True or false: the doors may shift in position, but to figure out which one gives more rewards, I should focus on the colour* as this is a piece of information that could only be gathered through actually reading the instructions, and not by doing the training.

Finally, while the second questionnaire at the end of the experiment was written specifically for the pilot study, it might be interesting to keep it in some form for any actual studies. Gathering data on the participants' attitudes on a continuous basis could identify any flaws that our pilot missed, or even provide some new insights on the decision-making process itself, as participants reflect on their own performance. More discrete and less descriptive questions than we used above would additionally allow for a more graphical representation of the data. A second questionnaire does of course increase the running time of the study, but might nonetheless surface valuable data.

The Dashboard, while comprehensive and ticking off all requirements set at the start of the project, offers some areas for expansion as well. More complex or follow-up studies might want to include the option to give punishments as well as rewards, which is currently not supported, or to vary the reward amount over time, which is now fixed for each block. On the Data page, for large-scale studies, it might prove useful to include different sorting and filtering options. An entirely different project could probably be made of a web-based analysis tool, that implements some of the analysis that we did in Python in Section 5.2 as a part of the Dashboard instead.

Software-based tools such as these are never finished, but we have to draw the line somewhere. We have fulfilled the requirements outlined in Section 3.3 and performed a pilot study that was ultimately even slightly larger than originally planned. Now, real data can be gathered. And after that, who knows! The world is an uncertain place, after all.

Bibliography

- [1] Peggy Seriès. *Computational psychiatry: A primer*. MIT Press, 2020.
- [2] Fei Xu and Joshua B Tenenbaum. Word learning as Bayesian inference. *Psychological review*, 114(2):245, 2007.
- [3] Konrad P Körding and Daniel M Wolpert. Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244–247, 2004.
- [4] John P O’Doherty, Sang Wan Lee, and Daniel McNamee. The structure of reinforcement-learning mechanisms in the human brain. *Current Opinion in Behavioral Sciences*, 1:94–100, 2015.
- [5] Richard S Sutton, Andrew G Barto, et al. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998.
- [6] Kentaro Katahira. The relation between reinforcement learning parameters and the influence of reinforcement history on choice behavior. *Journal of Mathematical Psychology*, 66:59–69, 2015.
- [7] Kate Nussenbaum and Catherine A Hartley. Reinforcement learning across development: What insights can we draw from a decade of research? *Developmental cognitive neuroscience*, 40:100733, 2019.
- [8] Angela J Yu and Peter Dayan. Uncertainty, neuromodulation, and attention. *Neuron*, 46(4):681–692, 2005.
- [9] Erdem Pulcu and Michael Browning. The misestimation of uncertainty in affective disorders. *Trends in Cognitive Sciences*, 23(10):865–875, 2019.
- [10] Rasmus Bruckner. *Decomposing the influences of uncertainty on learning*. PhD thesis, 2021.

- [11] Payam Piray and Nathaniel D Daw. Unpredictability vs. volatility and the control of learning. *bioRxiv*, 2020.
- [12] Massimo Silvetti, Eliana Vassena, Elger Abrahamse, and Tom Verguts. Dorsal anterior cingulate-brainstem ensemble as a reinforcement meta-learner. *PLoS computational biology*, 14(8):e1006370, 2018.
- [13] Michael Browning, Timothy E Behrens, Gerhard Jocham, Jill X O'Reilly, and Sonia J Bishop. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature neuroscience*, 18(4):590–596, 2015.
- [14] Rebecca P Lawson, Christoph Mathys, and Geraint Rees. Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, 20(9):1293, 2017.
- [15] Matthew R Nassar, Rasmus Bruckner, Joshua I Gold, Shu-Chen Li, Hauke R Heekeren, and Ben Eppinger. Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, 7(1):1–13, 2016.
- [16] Timothy EJ Behrens, Mark W Woolrich, Mark E Walton, and Matthew FS Rushworth. Learning the value of information in an uncertain world. *Nature neuroscience*, 10(9):1214–1221, 2007.
- [17] Christian Grillon. Models and mechanisms of anxiety: evidence from startle studies. *Psychopharmacology*, 199(3):421–437, 2008.
- [18] Pollyana Caldeira Leal, Tiago Costa Goes, Luiz Carlos Ferreira da Silva, and Flavia Teixeira-Silva. Trait vs. state anxiety in different threatening situations. *Trends in psychiatry and psychotherapy*, 39:147–157, 2017.
- [19] Charles Donald Spielberger. Manual for the state-trait anxiety inventory. *Consulting Psychologist*, 1970.
- [20] Amanda J Baxter, Kate M Scott, Theo Vos, and Harvey A Whiteford. Global prevalence of anxiety disorders: a systematic review and meta-regression. *Psychological medicine*, 43(5):897–910, 2013.
- [21] Valentina Rossi and Gilles Pourtois. Transient state-dependent fluctuations in anxiety measured using STAI, POMS, PANAS or VAS: a comparative review. *Anxiety, Stress & Coping*, 25(6):603–645, 2012.

- [22] Jessica Aylward, Vincent Valton, Woo-Young Ahn, Rebecca L Bond, Peter Dayan, Jonathan P Roiser, and Oliver J Robinson. Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature human behaviour*, 3(10):1116–1123, 2019.
- [23] Lorcan Kenny, Caroline Hattersley, Bonnie Molins, Carole Buckley, Carol Povey, and Elizabeth Pellicano. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4):442–462, 2016.
- [24] Catherine Lord, Mayada Elsabbagh, Gillian Baird, and Jeremy Veenstra-Vanderweele. Autism spectrum disorder. *The Lancet*, 392(10146):508–520, 2018.
- [25] Catherine Lord, Traolach S Brugha, Tony Charman, James Cusack, Guillaume Dumas, Thomas Frazier, Emily JH Jones, Rebecca M Jones, Andrew Pickles, Matthew W State, et al. Autism spectrum disorder. *Nature Reviews Disease Primers*, 6(1):1–23, 2020.
- [26] Simon Baron-Cohen, Sally Wheelwright, Richard Skinner, Joanne Martin, and Emma Clubley. The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of autism and developmental disorders*, 31(1):5–17, 2001.
- [27] KL Ashwood, N Gillan, J Horder, H Hayward, E Woodhouse, FS McEwen, J Findon, H Eklund, D Spain, CE Wilson, et al. Predicting the diagnosis of autism in adults using the Autism-Spectrum Quotient (AQ) questionnaire. *Psychological medicine*, 46(12):2595–2604, 2016.
- [28] Lydia Luke, Isabel CH Clare, Howard Ring, Marcus Redley, and Peter Watson. Decision-making difficulties experienced by adults with autism spectrum conditions. *Autism*, 16(6):612–621, 2012.
- [29] Benedetto De Martino, Neil A Harrison, Steven Knafo, Geoff Bird, and Raymond J Dolan. Explaining enhanced logical consistency during decision making in autism. *Journal of Neuroscience*, 28(42):10746–10750, 2008.
- [30] Suzane Robic, Sandrine Sonié, Pierre Fonlupt, M-A Henaff, Nathalie Touil, Giorgio Coricelli, Jérémie Mattout, and Christina Schmitz. Decision-making in a

- changing world: a study in autism spectrum disorders. *Journal of autism and developmental disorders*, 45(6):1603–1613, 2015.
- [31] Christina Boulter, Mark Freeston, Mikle South, and Jacqui Rodgers. Intolerance of uncertainty as a framework for understanding anxiety in children and adolescents with autism spectrum disorders. *Journal of autism and developmental disorders*, 44(6):1391–1402, 2014.
- [32] Mikle South, Paul D Chamberlain, Sarah Wigham, Tiffani Newton, Ann Le Couteur, Helen McConachie, Laura Gray, Mark Freeston, Jeremy Parr, C Brock Kirwan, and Jacqui Rodgers. Enhanced decision making and risk avoidance in high-functioning autism spectrum disorder. *Neuropsychology*, 28(2):222–228, 2014.
- [33] Nathaniel D Daw, John P O’Doherty, Peter Dayan, Ben Seymour, and Raymond J Dolan. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–879, 2006.
- [34] John C Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2):148–164, 1979.
- [35] Daniel Acuña and Paul Schrater. Bayesian modeling of human sequential decision-making on the multi-armed bandit problem. In *Proceedings of the 30th annual conference of the cognitive science society*, volume 100, pages 200–300. Washington, DC: Cognitive Science Society, 2008.
- [36] Maarten Speekenbrink and Emmanouil Konstantinidis. Uncertainty and exploration in a restless bandit problem. *Topics in cognitive science*, 7(2):351–367, 2015.
- [37] Ben Seymour, Nathaniel D Daw, Jonathan P Roiser, Peter Dayan, and Ray Dolan. Serotonin selectively modulates reward value in human decision-making. *Journal of Neuroscience*, 32(17):5833–5842, 2012.
- [38] Robert C Wilson and Anne GE Collins. Ten simple rules for the computational modeling of behavioral data. *Elife*, 8:e49547, 2019.
- [39] The Mathworks Inc. Find minimum of unconstrained variable function, August 2021. Retrieved from <https://uk.mathworks.com/help/optim/ug/fminunc.html>.

- [40] The SciPy community. Optimization and root finding, August 2021. Retrieved from <https://docs.scipy.org/doc/scipy/reference/optimize.html>.
- [41] Jesse Chandler and Danielle Shapiro. Conducting clinical research using crowd-sourced convenience samples. *Annual review of clinical psychology*, 12, 2016.
- [42] Claire M Gillan and Nathaniel D Daw. Taking psychiatry research online. *Neuron*, 91(1):19–23, 2016.
- [43] Samuel Zorowitz, Yael Niv, and Daniel Bennett. Inattentive responding can induce spurious associations between task behavior and symptom measures. *PsyArXiv*, 2021. <https://doi.org/10.31234/osf.io/rynhk>.
- [44] Prolific. Prolific: online participant recruitment for surveys and market research, May 2021. Retrieved from <https://www.prolific.co>.
- [45] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70:153–163, 2017.
- [46] Stefan Palan and Christian Schitter. Prolific. A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17:22–27, 2018.
- [47] Ben Caldwell, Michael Cooper, Loretta Guarino Reid, Gregg Vanderheiden, Wendy Chisholm, John Slatin, and Jason White. Web content accessibility guidelines (WCAG) 2.0. *WWW Consortium (W3C)*, 290, 2008.
- [48] Facebook Open Source. React—a JavaScript library for building user interfaces, May 2021. Retrieved from <https://www.reactjs.org>.
- [49] Microsoft Corp. TypeScript: typed JavaScript at any scale, May 2021. Retrieved from <https://www.typescript.com>.
- [50] The Sass team. Sass: Syntactically awesome style sheets, August 2021. Retrieved from <https://sass-lang.com/>.
- [51] The Bootstrap team. Bootstrap: the most popular html, css, and js library in the world, August 2021. Retrieved from <https://getbootstrap.com/>.
- [52] Matt Zabriskie. Axios, August 2021. Retrieved from <https://axios-http.com/>.

- [53] Mike Bostock. D3.js - data-driven documents, August 2021. Retrieved from <https://d3js.org/>.
- [54] Joshua R de Leeuw. jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior research methods*, 47(1):1–12, 2015.
- [55] Joshua R de Leeuw and Benjamin A Motz. Psychophysics in a web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, 48(1):1–12, 2016.
- [56] Benjamin E Hilbig. Reaction time effects in lab-versus web-based research: Experimental evidence. *Behavior Research Methods*, 48(4):1718–1724, 2016.
- [57] Robert A Rescorla. A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Current research and theory*, pages 64–99, 1972.
- [58] Katia M Harlé, Dalin Guo, Shunan Zhang, Martin P Paulus, and Angela J Yu. Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making. *PloS one*, 12(10):e0186473, 2017.
- [59] Flurin Cathomas, Federica Klaus, Karoline Guetter, Hui-Kuan Chung, Anjali Raja Beharelle, Tobias R Spiller, Rebecca Schlegel, Erich Seifritz, Matthias N Hartmann-Riemer, Philippe N Tobler, et al. Increased random exploration in schizophrenia is associated with inflammation. *npj Schizophrenia*, 7(1):1–9, 2021.

Appendix A

Participant information sheet and consent form

What follows is the complete Participant Information Sheet and Consent Form as integrated in the experiment.

Participant Information Sheet

- Project title: Decision making in uncertain environments
- Principal investigator: Dr Peggy Seriès
- Researchers collecting data: Filippo Ferrari, Rayo Verweij
- Funder (if applicable): N/A

This study was certified according to the Informatics Research Ethics Process, RT number 6014. Please take time to read the following information carefully. You should keep this page for your records.

Who are the researchers?

Dr Peggy Seriès, Filippo Ferrari, and Rayo Verweij.

What is the purpose of the study?

This study aims to gather insight into the human decision-making process in uncertain environments. Specifically, this study aims to examine how this process changes in individuals with personality traits broadly related to or commonly associated with

anxiety, autism, or depression. For this purpose, the study gathers experimental data through a decision-making task, which can be used to inform computational models of human decision-making.

Why have I been asked to take part?

You have been invited to take part in the study because you signed up for this study on Prolific.

Do I have to take part?

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason, by closing the experiment in the browser. Your rights will not be affected, and any progress you made while doing the experiment will not be saved.

What will happen if I decide to take part?

You will take part in an online behavioural experiment lasting around 20 minutes. In the first part of the experiment, you will be asked to fill out a brief questionnaire asking about different personality traits that you might display. In the second part of the experiment, you will be asked to play a decision-making game, in which you have to constantly choose from two options with potential rewards in order to maximize your total gains.

While the questionnaires ask questions about personality traits that may be broadly associated with specific psychological conditions, they do not represent diagnostic tests.

You only need a keyboard in order to participate in the experiment. No audio, video, or other device data will be recorded. To keep the collected data consistent, the experiment can only be done on a laptop or desktop screen, and not on a mobile phone or tablet.

Compensation.

You will be paid through Prolific for your participation in this study.

Are there any risks associated with taking part?

There are no significant risks associated with participation.

Are there any benefits associated with taking part?

You will be compensated for taking part as mentioned above.

What will happen to the results of this study?

The results of this study may be summarised in published articles, reports and presentations. As Prolific assigns you an anonymised Participant ID, it is impossible for both us and anyone else to identify you. Your data may be archived for a minimum of 2 years.

Data protection and confidentiality.

Your data will be processed in accordance with Data Protection Law. All information collected about you is anonymous. Your data will be referred to by the unique Participant ID as assigned through Prolific. Your data will only be viewed by the researcher/research team Dr Peggy Seriès, Filippo Ferrari, and Rayo Verweij.

All electronic data will be stored on a password-protected encrypted computer, on the School of Informatics' secure file servers, or on the University's secure encrypted cloud storage services (DataShare, ownCloud, or SharePoint).

What are my data protection rights?

The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance with Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

For general information about how we use your data, go to: edin.ac/privacy-research.

Who can I contact?

If you have any further questions about the study, please contact the lead researcher, Dr Peggy Seriès, at pseries@exseed.ed.ac.uk.

If you wish to make a complaint about the study, please contact inf-ethics@inf.ed.ac.uk. When you contact us, please provide the study title and detail the nature of your complaint.

Updated information.

If the research project changes in any way, an updated Participant Information Sheet will be made available on <http://web.inf.ed.ac.uk/infweb/research/study-updates>.

Consent

By proceeding with the study, I agree to all of the following statements:

- I have read and understood the above information.
- I understand that my participation is voluntary, and I can withdraw at any time.
- I consent to my anonymised data being used in academic publications and presentations.
- I allow my data to be used in future ethically approved research.

Button labelled "I agree".

Appendix B

Pilot questionnaire results

To give a complete overview, these are all responses to the evaluation questionnaire at the end of the pilot study, as discussed in Section 5.1. To keep the responses intact, they were not edited, but as such include the occasional typing or grammar mistake.

How did you feel during the game show? (E.g. excited, bored, indifferent, etc.)

- “It was quite exciting seeing the money building. It was a decent amount of money to get excited about, despite knowing that we wouldn’t actually receive it.”
- “Slightly bored”
- “Excited at the beginning but started to get bored towards the end.”
- “Indifferent.”
- “indifferet”
- “First excited, then started to get very bored aferwards.”
- “I felt slightly bored towards the end of the game (i.e. after the second break). I was trying to figure out a strategy... don’t think I managed to. lol”
- “After a few attempts I was already a little bored by the repetition of the task.”
- “I was very concentrated at the beginning and enjoyed choosing the doors but during the last round I got a bit bored.”
- “At first it was exciting but it got boring very quickly and then it lasted too much.”
- “interested, focused on understanding the mechanics of the game in order to get as many rewards as possible”
- “I was a little bored towards the end, and sometimes it felt like there was no way of knowing which door was more likely - no clear rules- so it felt like I was randomly guessing (potentially what you wanted?)”
- “Quite excited, the questionnaire was fairly short and the game did not last for too long.”

- “Excited at the beginning, but then I started losing interest as the same task repeatedly came. My suggestion might be to change graphics between the tasks, like in the second part it could be a window instead of a door or another secret door, or anything what is different. Also, it could emphasize that it is a different round with different probabilities.”
- “At first I was indifferent about it, but then I felt quite intrigued by it and eager to discover whether I chose the right door throughout the entire game.”
- “Excited! I had fun. I tried to get in the head of whoever made this, tried to see if there was logic or a pattern that repeated itself”
- “Bored - once I understood what was happening then it was fairly boring”

Did you have any issues while doing the experiment? Were the instructions clear? Is there anything that would have made the experience better?

- “The instructions were very clear.”
- “No issues, the instructions were very clear.”
- “Everything okay!”
- “Instructions were very clear.”
- “I’d be great to know how many games I have left :)”
- “No issues.”
- “Everything was clear.”
- “Everything was pretty clear”
- “Yes, I think the instructions were clear. I don’t remember seeing a sentence saying how long it was gonna be in terms of amount of rounds/decisions etc. (I know it said 30 mins) which I would’ve found nice (because time estimates are usually not very precise based on my experience).”
- “No issues.”
- “it was too long and repetitive”
- “The instructions were clear, but I was a bit confused as to whether the value of the door would change depending on the COLOUR of the door or the POSITION of the door (left/right) (green/pink), or whether it was a combination of both? But again, maybe you’re not meant to know this”
- “No issues, the instructions were clear, however, I would consider showing larger key icon below the door to make it clearer which key opens which door. Although maybe it’s deliberately small and showing extra information would be a distraction?”

- “Firtsly, there were some words I did not know. If the audience will not be only native speakers, it would worth considering replacing the following words with easier ones, or explaining them: turmoil, inadequate.
Secondly, I did not really like the ”press a key” concept. Maybe it had a reason, but I felt strange that after pressing ”Submit” several times I had to press ”C” to continue. Same with the doors, I would prefer cllicking on them, if possible.
A tiny thing, but some page was scrollable even there was nothing in the bottom. It can be annoying or misleading for someone. I noticed this on these pages: last page of the experiment, the previous page before the doors.”
- “I didn’t have any issue while doing the experiment and the instructions were super clear.”
- “No, all good”
- “The instructions just before were a bit confusing but once I did the trial it made sense”

How long did you (approximately) spend on the entire experiment? Did it feel long, short, or fine?

- “It felt just right, not too long or too short.”
- “15 minutes. It felt quite repetitive, but shorted than I expected.”
- “A bit long.”
- “Possibly a little too long.”
- “fine”
- “Spent about 10-15 minutes. It felt long.”
- “Around 15 minutes. It felt slightly long towards the end. Perhaps a progress bar would have helped ?”
- “It felt quite long, but not too much (I think almost 15 minutes)”
- “15ish minutes probably & it felt a tad long at the end (as I said in question box 1, I got a bit bored during the last round) but it was alright. :)”
- “13 minutes, it felt vey long.”
- “it feel long”
- “Spent around 10 minutes, it started to feel quite long”
- “About 15 minutes. Did not feel very long.”
- “It felt long in the end. My recommendation is in the first answer (changing figures).”
- “I think I spent 20 minutes on the entire experiment and it felt fine.”
- “short I would say. A bit repetitive after a while”
- “15mins - a little bit long but okay”

How did you decide which door to choose? Did you use a specific strategy? Did your strategy change over time?

- “Initially I chose between right and left, but then quickly realised it was based on colour. So I kept choosing the door with the colour that kept giving rewards. If I received no reward for 2 or 3 goes, I would switch and try the other door for one or two turns.”
- “Initially I explored the two doors and then I noticed (or assumed) that one door was the ‘good one’ and the other was the ‘bad one’. I then stuck with it until they got reversed. So when I felt that the current door wasn’t providing the rewards of a ‘good door’ anymore, I changed to the other one. I think that happened when I got 2 Xs in a row, but that wasn’t a conscious choice on my part.”
- “I tried a door 2 or 3 times... if I got 2 or 3 X in a row, I’d change my alternative. I did this from trial 2 onwards.”
- “I tried to figure out the underlying probabilities of reward associated with each door.”
- “I was exploiting until I found no reward for 2/3 clicks at the same door.”
- “I tried several strategies (based the rules on color, sides, place changes) but they didn’t seem to work.”
- “I tried to choose a strategy - which also changed throughout trials. Didn’t hack it though!”
- “I just tried to stick with the green door with no particular reason”
- “I started by choosing a random door and when that door was unsuccessful twice in a row (or unsuccessful, successful, unsuccessful) I switched to the other door. If it was successful enough I kept choosing the door.”
- “I stuck with the colour of the door that gave rewards (even if the doors swapped places). If I didn’t get a reward for 3 consecutive turns by choosing that door, I changed my strategy to choosing the door of the other colour.”
- “Yes, I use a strategy. No, I haven’t changed it because it works”
- “I just picked the color which seemed to give more rewards, but when there wasn’t much reward I would switch. I started with more complicated strategies, such as swapping the color of my choice if the doors flipped position on two consecutive trials.. but I soon found this too confusing and lost track”
- “It seemed like there is a different pattern in each round which I tried to understand as I played. For some rounds it was quite easy to figure out the pattern and almost always gain the reward, for others I couldn’t figure it out. I was more willing to change strategy in situations where I felt I lose very often.”
- “At the beginning I tried to be as logical as possible and think to all possible strategies (like figuring out when the rewarding door EXACTLY changes). Later I realised that it might be probabilistic, thus, I started choosing the door I believed had the higher probability to be rewarding. In the end, when I felt the probabilities are very close to each other, I tried to be a bit more

random (choosing the one with higher probability usually, and for every 5th click choosing the one with less).”

- “At first I tended to focus more on the side (left or right) rather than on the color of the door: it was a matter of finding a sort of algorithm; I would often press 3 times on the same side regardless of the color and then switch. However, as the game progressed, I focused more on the color of the doors and I would select the pink ones more often, since I found out that it was more likely that the money was there.”
- “At first i just decided on the green because i like the color. But after a while, it felt as though the pink had better chances so I just started chosing the pink by default.”
- “When it started I would check which it was and then I would continuously pick that colour as it stayed that colour mainly apart from a few swaps - so rather than keeping on swapping back and forth I would stay on that colour as I knew it was the majority”

The probabilities of getting a reward associated with each door changed throughout the experiment. (Roughly) how often do you think they changed?

- “Perhaps two or three times during each trial.”
- “Every 8 trials?”
- “Around 5.”
- “Twice a trial?”
- “every 20 clicks”
- “I didn’t really notice the changes.”
- “I think they changed after each break...”
- “I think it changed 3 or 4 times (I’m not sure)”
- “~10 times”
- “Maybe 3 times or so.”
- “Every time the game had a break”
- “I think they maybe changed at least more than four times.. but I am not too sure.. I wasn’t that aware of the changing probabilities to start with”
- “It felt like the probabilities were quite stable throughout each round, possibly with minor adjustments 2-3 times in each round. However, I based on my decisions primarily on the color of the door and most of time did not use different technique for left and right door.”
- “I would say it changed only between rounds. At least, that is what I felt. I would say 4-5 overall.”
- “I would say 3-4 times. The probabilities of getting a reward with each door were quite low during the first ”match” and they increased as the game progressed.”
- “I dunno. 4- 5 times?”
- “I’m not sure - as then I felt it changed back quite quickly”

Is there anything else you'd like to share about your experience?

- “The interactive opening of the doors was nice to see, and the slight delay between choosing a door and seeing whether there was a reward created a certain amount of anticipation that was quite exciting.”
- “No”
- “It seemed to me that the purple door was almost always the door with the highest reward probability.”
- “Nothing in particular.”
- “good job Rayo!:)”
- “.”
- “Cool experiment - I'd be interested to know more about it!”
- “No thank you”
- “I was a bit confused about the sentence “I feel pleasant.” (It might just be a personal problem because English isn't my native language but it seems to be very similar to content or happy (which were both in there as well)?).
Also - just a super tiny thing: “Your right of access can be exercised in accordance Data Protection Law” is missing a “with” or something in front of Data (I think; I could be wrong though haha).
Oh and I really liked the simplistic design of it, I think it's important to not have any distractions on the screen!”
- “No”
- “It was fun, but too long.”
- “Potentially have the instructions with diagrams of what the experiment will look like! I can't remember there being any”
- “Very interesting experiment, I have never done this kind of combination of two activities. I would be keen to see the results of this study as well as find out what do my results say about me.”
- “Firtsly, it would have helped me if I knew that a probability is associated with each door. Somehow I thought that the rewarding door is changing based on a deterministic heuristic, and I wanted to figure this heuristic out. Maybe it is only my mathematical mind, and it would be obvious for anyone else.
Also, if the money will be some kind of real money, it would be extremely important to hide the answer in the source code of the page. Cheating was very easy.”
- “No,I have nothing more to add.”
- “No”
- “N/A”

Appendix C

Installation guide

This guide should be able to take you from downloading the attached code to running the experiment with Prolific. Keep in mind, though, that platforms and technologies change, and that over time, this guide will become outdated, too. If you run in any trouble, feel free to reach out to me, or to submit a ticket to the Informatics Computing Help Desk at <https://computing.help.inf.ed.ac.uk/help-desk> - they've been incredibly helpful throughout this project with setting up the required infrastructure.

Step 1 - Request server space. Contact Computing Support through the link above, explaining the purpose of your study and that you require the following:

- Access to the PostgreSQL database
- Your own domain on the sweb service with 100MB of storage
- SSL enabled for accessing the database from the sweb service

100MB of storage is more than you'll ever need but trivial for Computing Support to arrange, and it's good to have a little spare.

Step 2 - Set up the database. Once you have the above, we'll start by setting up the database. Access a DICE terminal by either using a lab computer or SSH'ing into one from your own system and run the following commands:

- `psql -h <database_name> which is either pgteach or pgresearch, depending on which you were assigned to`
- `CREATE TABLE experiments (id varchar(128), timestamp timestamp, data json);`

- `quit` to quit the PostgreSQL environment

While this is the bare minimum necessary to set up the database, it would be useful to familiarise yourself with both SQL as a paradigm and the University's implementation at <https://computing.help.inf.ed.ac.uk/postgresql>.

Step 3 - Set up the web space. With the database set up, we'll move on to the websites. First, familiarise yourself with the University services we'll be using:

- <https://computing.help.inf.ed.ac.uk/informatics-filesystem>
- <https://computing.help.inf.ed.ac.uk/sweb>

Follow the instructions on the University pages to install AFS on your system and make sure you're able to use Kerberos to access this space. When you access it, you should be able to see two folders: 'data' and 'web'. We'll only be using the 'web' folder.

Inside the 'web' folder, a file named `test.html` should be present. Follow the instructions on the University pages to make sure you're able to view `test.html` in a web browser. If you can, you're all set.

Step 4 - Set up the experiment. Copy the files inside the attached 'experiment' folder to a location on your system. Inside, find the file named `write_data.php` and edit the first few lines with your database credentials. Then, create a new folder called 'ex' inside the AFS 'web' folder and copy all files from 'experiment' to 'ex'.

Step 5 - Set up the dashboard. If you are on your own system, install Node.js (<https://nodejs.org/>) - on DICE this is already installed. When this is done, copy the files inside the attached 'dashboard' folder to a location on your system. Open a terminal inside your local folder and run `npm i` to install all dependencies. Then, find the file at `public/db/connect.php` and edit the first few lines with your database credentials, similar to the step above.

Go back to the root folder of the project and run `npm run build`. This will take a minute or so and compiles the project. When this is done, create a new folder called 'dashboard' inside the AFS 'web' folder and copy all files from the 'build' folder to 'dashboard'.

By default, any user with a DICE login can access the Dashboard. To restrict access to only specific users, edit the `public/.htaccess` file as specified at <https://computing.help.inf.ed.ac.uk/cosign>.

Step 6 - Test your setup. Everything should now be in order. Try to access the Dashboard at <https://sweb.inf.ed.ac.uk/<you>/dashboard/index.html>. You should receive a prompt from Cosign to log in. Do so, and the Dashboard should load.

Play around with it for a bit and create a condition. Copy the URL with the copy button and paste it in a new tab. Before pressing Enter, change `{{PROLIFIC_PID}}` to 'R' (without quotes; to generate a random ID), and `{{SESSION_ID}}` and `{{STUDY_ID}}` to a value of your choosing. The experiment should load, and when you've finished the experiment, the data should appear in the Data tab of the Dashboard.

Step 7 - Get ready for Prolific. At the bottom of the `ex.html` file in the 'ex' folder, uncomment line 779, to redirect participants back to Prolific upon completing the experiment. This should be everything that you need to do to get set up for Prolific.

Create a Prolific account and follow their instructions. When they prompt you to provide a study URL, simply copy and paste in the URL from the Dashboard without making any changes.

Congratulations, you're now all set up. Enjoy your science!

Appendix D

Adding questionnaires

This guide will walk you through adding more questionnaires to the platform, without being very familiar with React, TypeScript, or jsPsych. This process will be a lot easier and less prone to errors if you are (especially with jsPsych), but it should work even if you are not.

Step 0 - Set up the code. Follow the instructions in Appendix C to set up the code on your machine and make sure that everything is working.

Step 1 - Add the questions and scale. In the Experiment codebase, find the ‘questionnaires’ folder. Inside, you will find three files containing the questions and scale of the AQ, STAI Y-1, and STAI Y-2, respectively. Copy one of these and edit them with the questions and scale of your questionnaire. In this case, the structure of the files should speak for themselves.

Step 2 - Edit the Experiment. In `ex.html`, we need to make a few additions.

- In the HTML `<head>`, import the script file you made in Step 1.
- Starting from line 291 (marked by a “Questionnaire” header comment) is the code that selects the appropriate questionnaire based on the variable that is set. Copy the AQ if-statement (*not* the STAI ones, as they have some custom logic to account for their two options), and update the values in the object inside with the values of your questionnaire. Don’t forget to update the if-condition itself with a new number (most likely 4, if this is the first questionnaire that you are adding).

When this is done, copy `ex.html` and the new questionnaire file to their respective locations on the AFS space.

Step 3 - Edit the Dashboard. Finally, the questionnaire should be selectable in the Dashboard. Find the file `src/conditions/Condition.tsx` and open it. Around line 225, there is a series of `<option>` tags. Add a new one with the value you entered in the if-condition in Step 2 and the title of the questionnaire.

Run `npm run build` in a terminal and copy the files from the 'build' folder to your 'dashboard' folder on the AFS space.

This is all that should be required to add a new questionnaire.