THE UNIVERSITY
of EDINBURGH

Experiments on Behaviour and Decision Making in Health

Hannah A. Behrendt

A thesis presented for the degree of
Doctor of Philosophy

School of Economics
University of Edinburgh
2018

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text. Chapter 3 of this thesis is co-authored with Dr. Yonas Alem (University of Gothenburg), Professor Michèle Belot (University of Edinburgh, European University Institute), and Anikó Bíró (University of Edinburgh, Corvinus University of Budapest). I made substantial contributions to the third chapter, including to the origin of the research question, trial design, implementation, data cleaning, data analysis, and writing the paper. This thesis has not been submitted for any other degree or professional qualification.

(Hannah A. Behrendt)

To my family and to Richard.

Acknowledgements

I am sincerely grateful to Professor Michèle Belot for her friendship, guidance, and encouragement throughout my research. I would like to thank the faculty and staff at the Edinburgh School of Economics for their support. In particular, I am grateful to my second supervisor Professor Ed Hopkins for helpful discussions. Kohei Kawamura and Ana Nuevo-Chiquero also provided helpful feedback while managing the ethics approval process at the Edinburgh School of Economics. Ivan Salter at the Behavioural Laboratory at The University of Edinburgh (BLUE) provided invaluable practical support in administering the experimental sessions. I am also grateful to many of my fellow PhD students for their friendship and advice.

I would like to thank the Behavioural Insights Team for hosting me during much of this research. I am especially grateful to Michael Hallsworth, Michael Sanders, Pantelis Solomon, Alex Tupper, and David Nolan for their helpful feedback. The GP referrals trial wouldn't have been possible without Hannah Burd's expert project management.

I am also grateful to many NHS staff who helped implement my field experiments. In particular, I would like to thank the staff at East Kent Hospitals University Foundation Trust, especially Julie Pearce, Penelope Searle, and Alisha Siddle, as well as Thariea Whisker and Alistair Martin from Canterbury and Coastal CCG.

Regarding chapter 3, I am grateful to my co-authors Dr. Yonas Alem, Professor Michèle Belot, and Dr. Anikó Bíró. We thank participants of the Winter Workshop on Labour Economics at the University of Amsterdam, seminar audiences at CREED (University of Amsterdam), CESS (University of Oxford), NYUAD, Innsbruck and University of Gothenburg for constructive comments.

I gratefully acknowledge financial support from the Edinburgh School of Economics and the Scottish Institute for Research in Economics (SIRE) through the Early Career Research Grant. For chapter 3, financial support from the Swedish Research Council, Formas, through the program Human Cooperation to Manage Natural Resources (COMMONS) at the Environmental Economics Unit, Department of Economics, University of Gothenburg is gratefully acknowledged.

Finally, I am deeply indebted to my best woman, Anne Hilger, for her unwavering friendship, for reading my code, my drafts, and my worries. This thesis is dedicated to my family and my husband Richard - your love and support sustains me every day.

Lay Summary

Research in judgment and decision-making has identified numerous ways in which human decisions are likely to be biased, deviating systematically from the behaviour one would expect if humans were fully rational. Suboptimal decision-making as a result of these biases imposes large costs on individuals and society. In this thesis I investigate experimentally how decision-making can be improved, focusing on the health domain, where errors ultimately become a matter of life and death.

Chapter 1: Providing feedback has been shown to be an effective way to change behaviour across several domains, including energy use and the workplace. However, there is little evidence on the effect of providing feedback in high stakes environments, where people make risky decisions with potentially catastrophic losses. In the first chapter I conduct a field experiment that provides feedback and information in such a context - emergency healthcare. Patients who had recently attended an Emergency Department (A&E) were randomly allocated to either receive a follow-up letter with information on alternatives to A&E or no letter (usual care). Overall, I do not find that the letter had an effect on the re-attendance rates of patients. However, the letter seems to have different effects on re-attendance, depending on the patient's age. I explore several reasons that might explain this.

Chapter 2: Policy interventions drawing on insights from behavioural sciences are increasingly popular, but relatively little attention has been paid to the extent to which the effects of repeated behavioural interventions are sustained over time. I study this question through a field experiment in the English National Health Service (NHS). The intervention consists of changing the salience of waiting time on the clinicians' e-Referral Service (ERS) interface. I find that putting a simple alert against services with high waiting times leads to a 35 percent reduction in the share of referrals to these services. The effect of the intervention is sustained over time and does not vary with prior referral habits. A small, low-cost intervention increasing the salience of waiting time to clinicians has a powerful and sustained effect on the choices of their patients.

Chapter 3: Behavioural attitudes toward risk and time, as well as behavioural biases, are thought to be important drivers of unhealthy lifestyle choices. While the first two chapters of this thesis take behavioural biases as given, the third chapter makes a first attempt at exploring the possibility of training the mind to alter these attitudes and biases, in particular relating to health-related behaviours. We conduct an experiment with 139 participants, around half of whom receive a four-week mindfulness training. We find strong evidence that mindfulness training reduces perceived stress, but only weak evidence of its impact on behavioural traits and health-related behaviours. Our findings have significant implications for a new domain of research on training the mind to alter behavioural traits and biases that play important roles in lifestyle.

Abstract

Research in judgment and decision-making has identified numerous ways in which human decisions are likely to be biased, deviating systematically from the behaviour one would expect if humans were fully rational 'Econs', maximizing their utility with perfect self-control. Suboptimal decision-making as a result of these biases imposes large costs on individuals and society. In this thesis I investigate experimentally how decision-making can be improved, focusing on the health domain, where errors ultimately become a matter of life and death.

Chapter 1: Providing feedback has been shown to be an effective way to change behaviour across several domains, including energy use and the workplace. However, there is little evidence on the effect of providing feedback in high stakes environments, where people make risky decisions with potentially catastrophic losses. In the first chapter I conduct a field experiment that provides feedback and information in such a context - emergency healthcare. Understanding the drivers of patients' decision-making relating to their demand for healthcare services and how they can be directed to the most appropriate services at any given time is a challenge for health systems across many countries. This trial aimed to reduce avoidable Emergency Department (A&E) attendances by sending a personal feedback letter to people who recently attended an Emergency Department but whose health concerns could have been dealt with elsewhere. Patients were randomly allocated to either receive a follow-up letter with information on alternatives to A&E or no letter (usual care). Overall, I do not find a statistically significant difference between the re-attendance rates of patients who did and did not receive the letters. However, the effect of the intervention interacted significantly with patient age, especially in men. I develop a conceptual framework that explores possible explanations for these heterogeneous effects.

Chapter 2: Policy interventions drawing on insights from behavioural sciences are increasingly popular and have been successfully applied across a number of different policy areas. However, little attention has been paid to the extent to which the effects of repeated behavioural interventions are sustained over time. In the second chapter I study this question through a natural field experiment in the English National Health Service (NHS). The intervention consists of changing the salience of waiting time on the clinicians' e-Referral Service (ERS) interface. Waiting longer can negatively affect patients' health gains from receiving treatment, so referring patients to services with shorter waiting times can be of benefit. I find, through a stepped wedge trial, that putting a simple alert against services with high waiting times leads to a 35 percent reduction in the share of referrals to these services. The effect of the intervention is sustained over time and does not vary with prior referral habits. A small, low-cost intervention increasing the salience of waiting time to clinicians has a powerful and

sustained effect on the choices of their patients.

Chapter 3: Behavioural attitudes toward risk and time, as well as behavioural biases such as present bias, are thought to be important drivers of unhealthy lifestyle choices. While the first two chapters of this thesis take behavioural biases as given, the third chapter makes a first attempt at exploring the possibility of training the mind to alter these attitudes and biases, in particular relating to health-related behaviours, using a randomized controlled experiment. The intervention we consider is a well-known psychological technique called "mindfulness", which is believed to improve self-control and reduce stress. We conduct an experiment with 139 participants, around half of whom receive a four-week mindfulness training, while the other half are asked to watch a four-week series of historical documentaries. We find strong evidence that mindfulness training reduces perceived stress, but only weak evidence of its impact on behavioural traits and health-related behaviours. We do not see that engagement with mindfulness training is correlated with behavioural characteristics such as impulsiveness and impatience. Our findings have significant implications for a new domain of research on training the mind to alter behavioural traits and biases that play important roles in lifestyle.

Contents

Introduction

Several decades of research in judgment and decision-making have highlighted the many different ways in which human behaviour deviates systematically from the behaviour one might expect from a perfectly rational, utility maximizing 'Homo Economicus'. Bounded rationality, willpower, and self-interest lead to various 'anomalies' in decision-making, requiring a richer model of human behaviour[1]. In 2009, Milkman et al. (2009) argued that the time had come, given the greater understanding of human behaviour and the different behavioural biases, to focus on researching how human decision-making could be improved. Since then, research on strategies for changing behaviour has not only increased, but also been integrated into policy-making in many countries (Whitehead et al. (2014) found 51 countries with central state-led policy initiatives drawing on behavioural sciences).

In this thesis I contribute to this growing area of research, focusing on the health domain where sub-optimal decision-making creates large costs for individuals and society. These may relate to both lifestyle behaviours and use of healthcare services more broadly; I consider each of these in turn.

Economists tend to think about people's lifestyle choices as a series of decisions involving trade-offs between immediate benefits and discounted future costs. Eating too much unhealthy food, smoking, drinking alcohol, and failing to exercise enough are but some examples of these kinds of choices. The trade-offs also involve risk - indulging in these behaviours affects the likelihood of ill-health. When analyzing these kinds of lifestyle choices, risk and time preferences therefore play a key role in standard economic models. Assuming that humans are perfectly rational and have perfect self-control, such models would also tend to support a view that, with humans rationally maximizing their utility, the case for policymakers to intervene is unclear. Loewenstein et al. (2012) note that conventional economics, therefore, can justify interventions aimed at changing such behaviours only in cases where there are externalities, i.e. where the actions of an individual impose costs on others.

However, research in behavioural economics has shown that people in fact exhibit a range of "behavioural biases". It is the prevalence of these biases and the not-fully-rational human behaviour resulting from them that opens up the possibility of policy interventions to help address costs that people impose not on others, but on themselves, i.e. internalities (Herrnstein et al., 1993). Present bias - a bias toward immediate gratification, for instance, has been shown to exist in various different contexts, such as in people's food choices, where people consume fewer calories when they order their

[1]Many of these are described in the so-called 'Anomalies' series of papers by recent Nobel Laureate Richard Thaler in the Journal of Economic Perspectives.

food in advance (VanEpps et al., 2016).

Not only do these biases exist, when given the chance to overcome them people often choose to do so. Thaler and Benartzi (2004), for example, allow workers to pre-commit to future savings increases, exploiting the fact that is easier to commit to doing something costly in the future than in the present. If people's behaviour often does not reflect how they themselves would like to behave, the case for policy interventions to address this becomes clear.

People's lifestyle choices are an important risk factor in many non-communicable diseases, often referred to as "lifestyle diseases". Worldwide, the World Health Organisation estimates that the majority of deaths in 2012 (around 68%) were due to non-communicable diseases (NCDs) including cardiovascular diseases, cancer, chronic respiratory diseases and diabetes (World Health Organisation, 2014). Looking more closely at the underlying risk factors, Forouzanfar et al. (2015) find that behavioural risk factors alone account for more than 16% of total DALYs (disability-adjusted life-years) as well as an additional 13% in interaction with other risk factors.

The economic costs of poor health-related behaviours are also high. In the United States, obesity is associated with higher absenteeism from work (Cawley et al. (2007) estimate an annual cost of $4.3 billion) as well as lower work productivity while working (Gates et al., 2008). In the United Kingdom, cardiovascular disease alone cost the UK economy £29.1 billion in 2004. Productivity losses due to mortality and morbidity accounted for 23% while informal care made up 17% of costs, with the remaining 60% comprising of healthcare related costs (Luengo-Fernandez et al., 2006). A more recent study concluded that "the largest economic burden to the NHS is due to poor diet and that much of this food-related burden is due to overweight and obesity"(Scarborough et al., 2011).

Alongside the increasing burden of lifestyle-related diseases on health services, how we use these services is also an important question. Timely access to the right treatment is key, yet health systems do not always deliver this. In England, each additional week of patients' waiting for hip or knee replacement surgery, for example, is associated with health losses estimated at £11.1m and £11.5m respectively (Nikolova et al., 2015). Patients do not always receive treatment in the most appropriate setting, with around 20% of admissions to Accident and Emergency (A&E) departments made for known conditions that could be managed in other settings (National Audit Office, 2013). Looking at A&E attendances, NHS England's Urgent and Emergency Care Report (NHS England, 2013) notes that 40% of attendances result in discharge without requiring any treatment at all. Of course this does not mean that all of these attendances were inappropriate, however, it is likely that at least some of these patients' health concerns could have been addressed in other settings. While the problems outlined above are complex, the question arises whether behavioural interventions can improve health outcomes, both individually and for society as a whole, by helping people to make better use of the health services available to them.

The preceding paragraphs have outlined the potential for changes in human behaviour to impact on the provision of health services and human health. Each chapter of this thesis focuses on a different aspect of behaviours related to health outlined above. Chapters 1 and 2 of this thesis look at the demand and supply of health services, while chapter 3 turns to health-related 'lifestyle' behaviours. The chapters also differ in the

approach to changing behaviour.

In order to outline these approaches, I draw on the distinction between two systems of cognitive functioning (Stanovich and West, 2000). System 1 refers to people's intuitive system, where decisions are fast and automatic, while System 2 makes reasoned judgments in a way that is slower and effortful. Milkman et al. (2009) argue that there are then broadly two different ways in which research on biases and errors in decision-making can be used to improve decision-making: helping people to move from System 1 to System 2 when this would improve decision-making on the one hand and 'leveraging' System 1 to facilitate better decisions on the other. Loewenstein et al. (2012) note that "the same errors that trip people up can also be used to help them". People's present bias, for example, can be harnessed to their advantage by programmes offering small, immediate rewards for behaviours such as smoking cessation or weight loss (see for example Volpp et al. (2009) or John et al. (2011)).

Chapter 1 of this thesis might be considered as exploring the effect of a more classical policy intervention - providing information and feedback to people about their behaviour, but in a way that is both timely and personalised. As we shall see, changing behaviour that is likely driven more by the automatic and emotional System 1 can be difficult, especially in a high-stakes decision such as whether or not to attend an A&E department.

Chapter 2 then presents an intervention that can be considered as 'leveraging' System 1 through a 'nudge' - changes in the choice architecture that still preserve individuals' freedom of choice. As noted by Thaler and Sunstein (2008) "small and apparently insignificant details can have major impacts on people's behavior. A good rule of thumb is to assume that "everything matters". In many cases, the power of these small details comes from focusing the attention of users in a particular direction." I find that a small change, increasing the salience of waiting time on the e-referrals interface, indeed has a relatively large impact on where patients are referred to by General Practitioners for further investigation or treatment. Importantly, these effects appear to be sustained over time.

Finally, chapter 3 explores the possibility of shifting the power from System 1 to System 2. Psychologists (e.g., Hofmann et al. (2008)) have argued that self-control could play a key role in determining which of the two prevails. We test whether mindfulness training can shift the balance of power between the two systems and affect lifestyle behaviours related to risk and time preferences. While the intervention appears to be effective at reducing participants' feelings of stress, the effects on decision-making related to health are not entirely clear. More research is needed to understand how such techniques can be used to help people make decisions that support their long-term goals.

Since the publication of *Nudge* (Thaler and Sunstein, 2008), policy interventions based on behavioural economics have become increasingly popular. Understanding how such interventions can be applied successfully is therefore more important than ever before. This thesis builds on existing work in the field and tries to shed light on a key question: how can behaviour be changed in a way that is sustained over time. It does this through three experiments related to decision making in health - a context where making the right choice can literally be a matter of life or death.

Chapter 1

Information and Feedback in a High-Stakes Environment: a Field Experiment in Emergency Care

1.1 Introduction

Providing people with timely, specific feedback on their actions can help maintain or change future behaviour (whichever is intended). A series of trials run by the energy company Opower in the US, for example, finds that receiving personalised feedback on past energy use compared to one's neighbours can reduce overall energy usage by $2-4\%$ (Allcott, 2011). However, evidence on the effect of feedback and information provision in what might be called a 'high stakes' environment - decisions that involve risk and potentially catastrophic losses - remains limited. In this chapter I study patients' decisions to attend an Accident and Emergency Department (A&E) to help shed light on this question.

Drivers of patients' decision making relating to their demand for health services are poorly understood. Across many countries, directing patients to the most appropriate health services at the right time is a challenge. However, policy interventions aiming to address this problem often lack a solid evidence base and robust evaluation is rare. Reducing avoidable attendances at Emergency Departments has become a major policy priority in England. Demand for these services is high, putting pressure on National Health Service (NHS) trusts and their staff. NHS England's Urgent and Emergency Care Report (NHS England, 2013) reports that 40% of patients who attend A&E are discharged without requiring any treatment at all and a National Audit Office report (National Audit Office, 2013) finds that around 20% of admissions are for known conditions that could be managed effectively outside of emergency care - by primary, community or social care services. 'Avoidable' attendances at A&E clearly have many causes. However, a systematic review of studies investigating non-urgent emergency department attendances found that more than 80% of patients reported that they believed their symptoms were urgent and could not wait for treatment (Uscher-Pines et al., 2013). Therefore, I hypothesize that this gap between perceived and actual

urgency plays an important role in 'avoidable' A&E attendances.

I conduct a field experiment aimed at addressing this gap by sending a personal feedback letter to people who recently made such an 'avoidable' attendance at A&E. During the trial period, patients whose attendance met the criteria for inclusion in the trial were randomly allocated to either receive a follow-up letter with information on alternatives to A&E (treatment) or no letter (control). I do not find a statistically significant difference between the re-attendance rates of patients who did and did not receive the letters. However, further exploratory analysis finds that the treatment effect interacts with age - especially in men. In order to develop a better understanding of the role of information and feedback in this high stakes context, I develop a conceptual framework of feedback and patient behaviour that accounts for several possible reasons why feedback may not change behaviour in a high stakes decision such as whether or not to attend A&E.

This is the first paper to provide evidence about the role of feedback in the high stakes decision of whether to seek emergency healthcare. There are many different contexts in which the effect of providing feedback has been studied, including for example the impact of performance feedback (Falk and Ichino, 2006; Kandel and Lazear, 1992). Receiving information about a previous 'avoidable' A&E attendance should give a patient information that could inform future decisions to attend A&E. Patients should thus become increasingly accurate in determining when it is appropriate to attend A&E. One might think of this as being similar to how performance feedback, for example, can help workers develop an increasingly accurate understanding of their performance and underlying ability. I contribute to the literature by developing theories of why feedback provision may or may not be successful in changing behaviour in a high stakes context, which could have important implications for the development of similar interventions in the future.

The rest of this chapter is structured as follows: Section 2 outlines related literature. Section 3 lays out the experimental design and procedure, describes the participant sample, and outlines the intervention and empirical strategy while section 4 presents descriptive statistics. I present results in section 5 and discuss possible explanations for the observed heterogeneous treatment effects in section 6.

1.2 Related Literature

This chapter contributes to the literature on information and feedback provision in a high stakes environment by studying decision-making in a health context. The intervention aims to reduce 'avoidable' attendances at East Kent Hospitals University NHS Foundation Trust (EKHUFT) by sending a timely personal feedback letter to patients who recently attended A&E, but whose health concerns could have been dealt with elsewhere. I consider an attendance at A&E 'avoidable' if the visit was coded as requiring 'No Treatment No Significant Investigation', indicating that other services or indeed self-care would have been appropriate for the patient's health concerns.

There are many factors that could explain patients' behaviour. One hypothesis, which forms the basis of public information campaigns, is that people do not have relevant information about alternative health services or lack an understanding of which services are appropriate for health concerns of different severity and acuteness.

A systematic review of studies investigating this question finds that several factors

are associated with non-urgent emergency department attendances (Uscher-Pines et al., 2013). Based on their review of the literature, the authors construct a conceptual framework to illustrate patients' decision making. In this framework, upon experiencing symptoms, the patient considers different options (ranging from doing nothing, self-care, or visiting a primary care physician, to attending the emergency department). On the basis of the studies reviewed in their research, the authors identify six factors that likely play a role in making this decision: perceived severity of symptoms, beliefs and knowledge about treatment options, convenience, access/availability, cost, and advice/referrals.

The review notes that four studies focusing on non-urgent emergency department attendances found that more than 80% of patients reported that they believed their symptoms were urgent and could not wait for treatment (Uscher-Pines et al., 2013). A survey of 279 patients in a university emergency department, for example, found that the most commonly reported reasons for choosing to attend included the belief they would receive better care there, the urgency of their medical complaint, and immediacy (Northington et al., 2005). Patients who cited urgency as the reason for their attendance thought that the symptoms they were experiencing required immediate attention and feared that failing to attend the emergency department would lead to an adverse outcome (Northington et al., 2005).

The intervention I present can be thought of as addressing two of the factors identified by Uscher-Pines et al. (2013). First, it provides information about alternative treatment options, i.e. individuals may gain new knowledge about alternative services. Second, it provides feedback on the severity of the symptoms experienced by the patient on their last attendance.

1.2.1 Information Provision

There is limited evidence of the effect of targeted information provision on emergency department visits. One trial in the US, focusing on paediatric care, provided families receiving Medicaid with information about the importance of primary care provision and supported them in making appointments (Grossman et al., 1998). The intervention showed promising short-term effects, decreasing emergency department attendances by 11.1% over the six months following the intervention, but no longer-term effects were found. In another trial nearly 7,000 households receiving Medicaid were randomly chosen to receive a booklet with information about common non-urgent conditions and how to receive treatment outside of emergency care departments (Rector et al., 1999). The study found that the intervention did not significantly reduce emergency department attendances. It is unclear, however, to what extent the results of studies in the US are relevant to the UK context where care through the NHS is free at the point of use.

In the NHS there have been numerous attempts at providing general information to NHS users about the alternatives to A&E departments, for example the regional 'Choose Well' campaign, which aims to educate people about which health services are appropriate in which situations. Like many public health campaigns, it provides general information that is not linked to individual behaviour.

In contrast to these previous trials and policies, my experiment presents information in a more concise and targeted way in order to avoid information overload. It aims to

provide patients who have recently made an avoidable visit to A&E with timely and simple information about what to do if an urgent care problem arises again in the future. Compared to other possible interventions, this particular trial offers a 'light-touch' intervention at relatively low cost.

Rather than just providing information, the trial also gives patients feedback on their previous attendance. Even when a patient's attendance at A&E is considered to be 'avoidable' by medical professionals, this may often not be communicated to patients.

1.2.2 Feedback

This is the first study to provide evidence about the impact of feedback on patients' decisions to seek emergency healthcare - a high-stakes situation. Feedback provision has been studied across a variety of different contexts. In the economics literature one obvious example is the impact of performance feedback (see for example Kandel and Lazear (1992); Falk and Ichino (2006)). However, in my case the feedback is given privately to patients, so is therefore perhaps more closely related to studies with private performance feedback provision (Azmat and Iriberri, 2010; Blanes i Vidal and Nossol, 2011). When a patient is provided with feedback on whether their visit to A&E was appropriate, this should also give her information about the need to attend A&E in the future. A Bayesian updater should develop increasingly accurate knowledge of when it is appropriate to attend A&E, similar to how performance feedback should allow workers to develop increasingly accurate knowledge of their ability or performance. Despite receiving feedback, however, it is well-documented that people are over-confident about their ability or performance (Burks et al., 2013) even when this has negative financial consequences for them (Camerer and Lovallo, 1999; Malmendier and Tate, 2008). It is unclear, however, to what extent this can be considered as indicating biased information processing.

My paper is perhaps more closely related to the limited literature on educational investments and choices and the role that subjective beliefs about academic ability have in this context. Similar to the health context I study, school choice is an infrequent choice with high stakes. Several studies using observational data have documented the role of students' beliefs about their own ability in predicting choices of college major or dropout decisions (Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2014, 2012). More recently, several field experiments have studied the effect of feedback. In a field experiment conducted in Malawi, Dizon-Ross (2016) generates exogenous variation in beliefs by providing parents with information about their children's academic performance and analyzes the effect this has on educational investments. Bobba and Frisancho (2016) conduct a field experiment that provides feedback on academic performance, reducing the gap between students' expected and actual performance. This leads to a closer alignment between students' school track choice and academic skills. Developing a better understanding of how feedback can shape behaviour in a high-stakes context such as the decision to seek emergency healthcare has important policy implications. I contribute to this literature by outlining the circumstances which may play a role in determining whether or not feedback is successful in changing behaviour.

1.3 Experimental Design

The field experiment was a randomized controlled trial with simple randomization at the individual attendance level. I analyze hospital patient records to determine whether the patients who received the letter are less likely to make an avoidable repeat visit in the future. Ethics approval was obtained for this trial from the NHS Research Ethics Committee (REC reference: 14/YH/1280).[1]

1.3.1 Sample

The field experiment was run in partnership with East Kent Hospitals University NHS Foundation Trust (EKHUFT) and Canterbury and Coastal Clinical Commissioning Group (CCG). The participant pool consisted of individuals who attended EKHUFT Emergency Departments or Centres during the trial period. This included the full Emergency Departments at William Harvey Hospital (WHH, Ashford) and the Queen Elizabeth The Queen Mother Hospital (QEH, Margate), as well as the Emergency Care Centre at Kent and Canterbury Hospital (KCH, Canterbury).

Each week during the trial period, EKHUFT identified those attendances in the last seven days that could have been dealt with elsewhere. Discussions with EKHUFT identified a process for classifying attendances as avoidable, using the "No Investigation No Significant Treatment" code[2]. Full inclusion and exclusion criteria are summarized in Table 1.1. A DBS (Disclosure and Barring Service) tracer was run on patients identified as participants in the trial to ensure accuracy of patient information prior to sending the letter.

Table 1.1: Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
A&E Attendance at EKHUFT	No fixed abode
'No Investigation No Significant Treatment'	No UK or valid address
(recorded as HRG4 code VB11Z)	Helicopter mode of arrival
18 years of age and older	Departure method 'BID'/
Successful DBS trace	'DID'(deaths)
WHH, QEH and KCH sites	
Local initiator is 'self'	

1.3.2 Intervention

The intervention aimed to reduce 'avoidable' re-attendances at Emergency Departments and Centres by sending letters to individuals who recently attended, but whose health issues could have been dealt with elsewhere. It provided clear information about NHS health services most appropriate for non-emergency health issues and was produced in consultation with Canterbury and Coastal CCG and staff at EKHUFT. Alongside a letter, patients also received a 1-page leaflet with more detailed information about

[1]The trial was pre-registered on the ClinicalTrials.gov Protocol and Results Registration System (NCT02334189) as well as the AEA RCT registry (AEARCTR-0000600).

[2]The Healthcare Resource Group (HRG) is used by the NHS to group patient events that are thought to consume a similar amount of resource. VB denotes Emergency Medicine, and VB11Z stands for 'No Investigation No Significant Treatment'

the available health services in the area - both intervention materials are included in Appendix 1B.

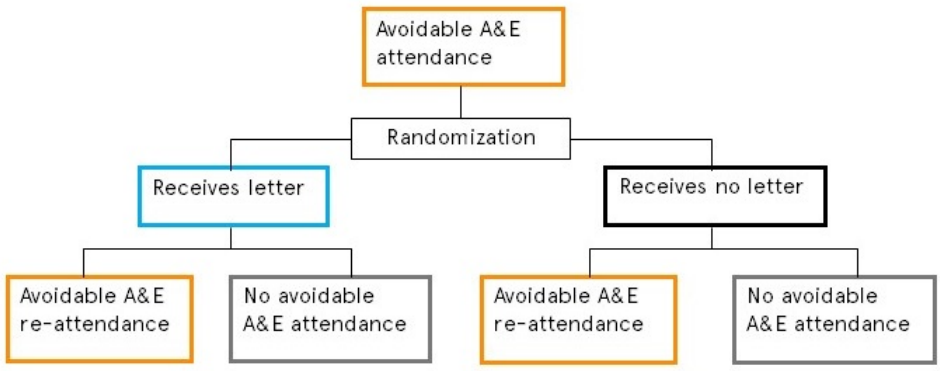
To make the letter as effective as possible, it included simple, clear messaging. Research shows that simplifying messages can significantly increase response rates to communications; breaking down complex goals into simpler, easier actions can be particularly effective (The Behavioural Insights Team, 2014). The letter was addressed to the patient and was signed by a senior medical professional, as both personalisation and the use of a high profile messenger can help to make communications more effective (see for example Durantini et al. (2006), The Behavioural Insights Team (2014)).

It was not possible to give patients direct feedback about their previous attendance, i.e. telling them that they should not have attended A&E - instead the letter refers to the recent attendance ("our records show you recently attended") and simply adds that there are "easy, local ways to get medical advice, without having to wait in A&E". Thus the letter did not criticize the individual for attending A&E and avoided being too closely tailored to the recipient's specific circumstances in order to keep the intervention simple and avoid the risk of error. The feedback patients receive is therefore relatively subtle. However, informal interviews with hospital staff revealed that patients rarely receive any feedback about their attendance. Medical staff often lack the time to provide this kind of feedback and education when patients are discharged from A&E; instead, patients are often reassured that they did the right thing by attending - just to be 'on the safe side'.

1.3.3 Experimental Procedure

The following analysis is based on hospital attendance data during the 38 week trial implementation period, including attendances from January 3rd 2015 to September 19th 2015, as well as follow-up data on attendances until February 29th 2016.

Figure 1.1: Trial flowchart



Each attendance was randomly assigned to either the treatment (letter) or control group (no letter). Ideally, the randomization process would have ensured that patients re-attending during the trial period would have either received a letter each time they attended or never. However, this was not possible given the logistics of implementation. On the same day every week during the intervention, the hospital created an Excel file of all patients who made an avoidable Emergency Department visit in the last 7 days. The data file contained patients' names, addresses, and salutations, together with the unique hospital identifying number. Patient contact details were copied by

hospital staff into an Excel sheet that contained a randomization mechanism, randomly assigning patients to either receive a letter, or receive no letter. The Excel sheets with the randomization outcomes were locked to prevent tampering. Figure 1.1 illustrates the trial process.

1.3.4 Empirical Strategy

I estimate the following linear probability model on the sample of patients who attended during the trial period:

$$Y_i = \alpha + \beta \; Letter_i + \Gamma \Psi_i + \epsilon_i \tag{1.1}$$

where Y_i is a binary indicator of whether an individual i re-attended during the intervention and follow-up period, *Letter* is a dummy variable equal to 1 for individuals in the treatment group, Ψ_i is a vector of individual characteristics such as gender, age, area of residence, and hospital site attended, and ϵ_i is an error term. I also present results using probit and logit models in the appendix.

1.4 Descriptive Statistics

During the trial period there were 11,592 qualifying attendances at the Emergency Departments of the three EKHUFT sites included in this trial (i.e. attendances coded as requiring 'No Investigation No Significant Treatment' and meeting the additional criteria outlined above). Out of these attendances, 52.45% were by women, and the average age at attendance was 41 (patients were only included in the trial if they were at least 18 years old).

Table 1.2: Characteristics of patients in treatment and control groups

	[1]		[2]		[3]		[4]
	Total		**Treatment**		**Control**		**Diff**
Variables	Mean	SD	Mean	SD	Mean	SD	Mean
Age	41.476	18.749	42.001	19.205	41.006	18.320	-0.994***
Female	0.525	0.499	0.534	0.499	0.518	0.500	-0.017*
Area	1.183	1.061	1.168	1.054	1.196	1.067	0.028
Hospital	1.896	0.778	1.885	0.780	1.906	0.777	0.021
Observations	10889		5145		5744		

* $p < 0.1$,** $p < 0.05$, *** $p < 0.01$.

Over the trial period 5,469 attendances were randomly assigned to the treatment group. Table 1.2 presents information on covariates from patients in the treatment and control groups, showing some imbalance on age and to a lesser degree gender - variables I control for in my analysis. Treatment and control groups were split evenly across the three hospital sites and geographical areas, though there are fewer attendances in total at William Harvey Hospital (Ashford) than at the other two sites. At each site the majority of qualifying attendances were made by people living relatively close to the hospital.

Over the 38 week course of the trial, 93.94% of 'avoidable' attendances (10,889) were first attendances during this period - i.e. this is the number of patients included

in the trial (it is of course possible that these patients made previous 'avoidable' attendances before the trial period). A further 4.87% of attendances (565) were second attendances by these patients, and 0.69% (80) were third attendances. The remaining 56 attendances were made by very few patients, including one patient who accounted for 21 'avoidable' attendances.

1.5 Results

Comparing 'avoidable' re-attendance rates of patients who received the letter (treatment group) with those of patients who did not (control group), I do not find a statistically significant difference in the re-attendance rates between the two groups.

Table 1.3: Impact of the Intervention on Re-Attendance Rates

VARIABLES	(1) Linear Probability Model (LPM)	(2) LPM with controls
Treatment	0.122	0.152
	(0.430)	(0.429)
Age		-0.0372***
		(0.0107)
Female		0.181
		(0.430)
Hospital and area controls		Yes
Constant	5.223***	5.787***
	(0.294)	(0.747)
Observations	10,889	10,889
R-squared	0.000	0.004

Robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Note that coefficients and standard errors are all multiplied by 100 in Tables 1.3 and 1.4 to facilitate interpretation of coefficients as percentage points - i.e. the re-attendance rate in the control group is about 5.2 percent. As I cannot be sure that patients opened the letters they were sent, these are intention-to-treat (ITT) estimates.

The dataset received from the hospital contains only few control variables - age, gender, 3-digit postcode area, and hospital site. It would have been interesting to investigate whether there are differential treatment effects depending on prior attendance history, education, or deprivation, but the limited nature of the dataset makes this impossible. Therefore, I focus on the two demographic variables I have access to: age and gender. Previous research has found that these variables are associated with differences in use of health services, so it seems appropriate to explore whether this is the case in my experiment.

The results of the main analysis show that older patients are less likely to make an 'avoidable' re-attendance than younger people, while gender does not appear to be significant. I conduct further exploratory analysis, which finds evidence of a differential impact of the intervention depending on patient age. Adding an interaction term for treatment with age, the coefficient on this interaction variable is positive and small, but statistically significant for men only. It appears to be the case that the intervention backfires slightly for older men while reducing re-attendances as intended in younger men.[3] The following section develops a conceptual framework of the underlying mechanisms of the intervention that is consistent with these findings.

Table 1.4: Impact of the Intervention on Re-Attendance Rates by Age and Gender

	(1)	(2)	(3)
VARIABLES	All	Women	Men
Treatment	-1.355	-0.311	-2.789*
	(1.036)	(1.468)	(1.459)
Age	-0.0551***	-0.0557***	-0.0540***
	(0.0146)	(0.0208)	(0.0205)
Treatment x Age	0.0364*	0.00561	0.0769**
	(0.0213)	(0.0293)	(0.0312)
Hospital and Area dummies	Yes	Yes	Yes
Constant	6.616***	6.234***	7.112***
	(0.861)	(1.221)	(1.211)
Observations	10,889	5,722	5,167
R-squared	0.004	0.005	0.005

Robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Another way to consider reattendance is to look at how long it takes for people to reattend. In order to consider this question I undertake a survival analysis to determine for each patient's first attendance how long it takes for them to reattend (a 'failure event' in the model). Figure 1.2 shows the Nelson-Aalen cumulative hazard estimates, depicting for both the treatment and control groups what proportion of patients have reattended within a given number of days since their first attendance in the trial period. Results from the semi-parametric Cox proportional hazard model are presented in Appendix 1A. The graphs also illustrate how the intervention seems to impact men differently across age groups. The first panel of Figure 1.2 shows the estimates for the full sample (men and women), while the other three panels depict the results for men, split into three age groups. For men in the youngest age group the graph shows that more patients in the control group reattend over time, while the opposite can be observed in the older age groups.

[3]I report results from a model exploring the interaction of treatment and gender in the Appendix. The interaction term of treatment and female is negative, but not statistically significant.

Figure 1.2: Survival Analysis

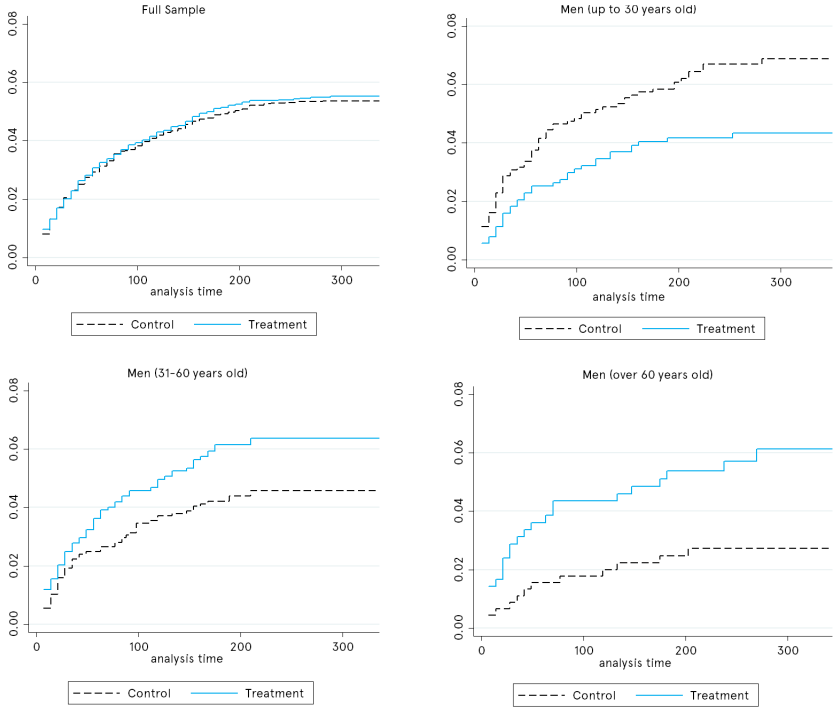


Table 1.5 reports estimates from the Cox proportional hazard model for the cohorts shown in Figure 1.2. This confirms that the intervention works as intended for young men, while significantly backfiring for men over 60. It is worth noting that the significance levels are sensitive to the exact age cutoff chosen - the coefficients are not statistically significant when the sample is split by age tertiles[4] (see columns 5-7), though the sign and magnitude of the coefficients are consistent across different ways of splitting the sample. Equivalent graphs and tables for women are shown in Appendix 1A, confirming that there are no significant interactions of treatment and age for women.

Table 1.5: Cox Proportional Hazard Model - Full Sample and Age Splits (Men Only)

	(1) Full sample	(2) Men up to 30	(3) Men 31-60	(4) Men over 60	(5) Men low tertile	(6) Men mid ttl.	(7) Men top ttl.
Treatment	0.0281	-0.444**	0.306*	0.764**	-0.338	0.222	0.328
	(0.0839)	(0.205)	(0.181)	(0.355)	(0.217)	(0.203)	(0.233)
Hospital & Area Dummies	yes	yes	yes	yes	yes	yes	yes
Observations	10,883	1,937	2,352	874	1,724	1,784	1,655

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[4]Tertiles were calculated for the full sample (men and women); the lowest tertile includes ages 18-28, the middle tertile 29-48, and the top tertile those over 48. Graphs showing the survival analysis by age tertiles are included in Appendix 1A.

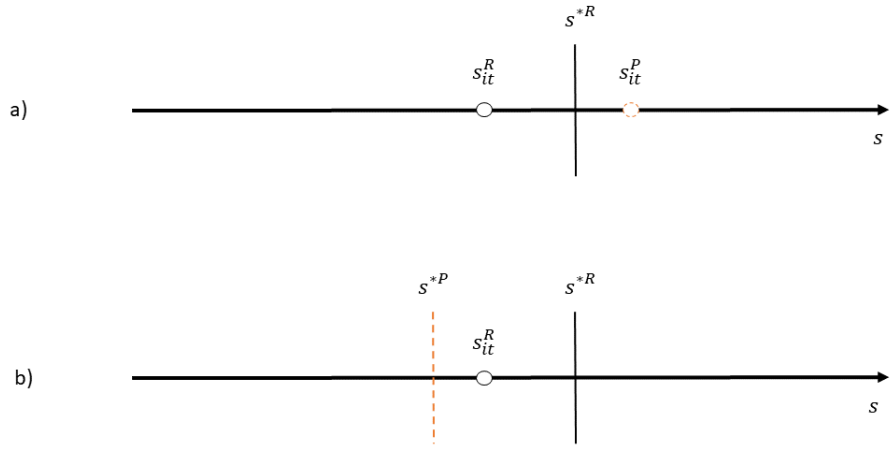
1.6 Discussion and Conclusion

While the intervention did not reduce the re-attendance rate of patients who visited A&E overall, there is evidence that the treatment interacted with patient age, working in the intended direction for younger patients, but backfiring in older patients. This particularly appears to be the case for men. In this section, I develop a simple conceptual framework consistent with this finding and present several hypotheses that might account for the observed differential treatment effect.

In this trial I focused on a patient's first re-attendance upon receiving a feedback letter. Reflecting this, assume there are two periods denoted by $t = \{1, 2\}$ in which patients experience a health problem. Patients inappropriately attend A&E at time $t = 1$. They then receive (or don't receive) information and feedback, after which they experience another health problem at $t = 2$. In this simplified framework, suppose that health concerns can be characterized on a scale of severity or urgency, where attending A&E is appropriate above a certain cutoff point, s^{*R}. The urgency of patient i's medical concern at time t is denoted by s_{it}. When patients have imperfect information, this can lead to two distinct reasons for 'inappropriate' attendances.

For a patient to decide to attend A&E, the perceived level of urgency must be greater than or equal to the perceived cutoff, i.e. $s_{it}^P \geq s^{*P}$. For this attendance to be 'inappropriate' the real level of urgency must be less than or equal to the real cutoff, i.e. $s_{it}^R \leq s^{*R}$.

Figure 1.3: Perceived and real urgency - scale and cutoffs



First, as shown in panel a) of Figure 1.3, a patient's perceived urgency s_{it}^P could be greater than the real urgency of their medical concern (s_{it}^R), leading them to believe that the degree of urgency is greater than the minimum level required for attending A&E. Second, as illustrated in panel b), it could be the case that while patients perceive the urgency of their medical concern correctly, they believe that the minimum level of urgency required for attending A&E is lower than it really is: $s^{*P} < s^{*R}$. Of course these two mechanisms are not mutually exclusive - patients could under- or overestimate one or both of these factors.[5]

[5]For example, a patient might inappropriately attend if they both underestimate the severity of their concern, but underestimate the cutoff by even more (e.g. they think it is appropriate to attend

The letter intervention provides patients with information about local healthcare services for non-urgent medical needs, and about the appropriateness of different healthcare services for different health problems. For patients who previously had mistaken beliefs about the location of the cutoff point s^{*R}, providing them with this information should help correct these beliefs.

The letter also provides patients with feedback about the urgency of their first attendance, i.e. that $s_{i1}^R < s^{*R}$. However, beliefs about the cutoff point are quite different from those about health concerns: the cutoff point should be considered as being fixed over time. In a dynamic setting where patients repeatedly experience health concerns, the real cutoff point is always in the same place. In contrast, health concerns are not always identical, so learning about the location of the previous period's health problem is only imperfectly informative about future health concerns. As noted in the literature review above, research on non-urgent emergency department attendances finds that the majority of patients making such attendances perceive their medical concerns to be urgent. Furthermore, as the field experiment took place in a real world medical setting, the feedback is fairly subtle as it was impossible to directly tell patients that they should not have attended A&E. It is possible that including more direct feedback could increase the size of the treatment effect, although implementing this in a clinical context would likely never be feasible.

Using this framework, I explore several hypotheses that might explain why the intervention appeared to have a differential impact on various groups. The first hypothesis is related to the issue illustrated in panel b) of Figure 1.3 - questions relating to knowledge about health services and how to access them. The second hypothesis relates to knowledge about health more generally and about a person's ability to correctly identify the severity of their health concern, as shown in panel a). Finally, the third hypothesis relates to how people behave in the face of uncertainty about how severe their health concern is. I outline the different hypotheses and present some evidence from the literature supporting them.

Hypothesis 1: Differences in knowledge of and access to health services. The intervention provides patients with information relating to health services and their appropriate use. There is evidence that young men are less likely to access health services such as primary care and the urgent telephone helpline NHS 111, though it is not entirely clear whether this is due to a lack of knowledge or other reasons. Wang et al. (2013) document a gender gap in General Practitioner (GP) consultation rates, which is 32% lower for men than it is for women (accounting for reproductive-related consultations makes the gap smaller, but does not close the gap completely). The greatest gender gap is observed amongst patients aged between 16 and 60. Similarly, evaluations of the telephone urgent health advice services NHS 111 and its predecessor NHS Direct both found that men were less likely to use the services than women were (Larner, 2009; The University of Sheffield, 2012). Certainly, given their lower use of these services, young men will be less familiar with them. Therefore it may be the case that young men are the only group who receive previously unknown information through the intervention.

Hypothesis 2: Differences in knowledge of health concerns. It might be the case that it is the subtle feedback that is only effective for young men. There is evidence of significant gender differences favouring women in many areas of health

A&E with very minor complaints).

knowledge, including common illnesses and treatment (Beier and Ackerman, 2003). If young men have a lower baseline of knowledge, feedback might have more impact on their future behaviour.

Hypothesis 3: Differences in risk attitudes. It may be the case that all groups receive new information through the intervention. However, perhaps young men respond more to the information because they are less risk averse and therefore more likely to try the alternatives suggested in the letter, rather than going straight to A&E next time they have a similar health concern. The evidence on gender differences in risk attitudes is not wholly conclusive. Dohmen et al. (2011) find that women are more risk averse than men both in general measures of risk, but also in context-specific measures, including health. The authors also note an interesting differential impact of age on willingness to take risks for men and women. For men the willingness to take risks decreases steadily as they get older. Women however experience a rapid decrease from their teens to age 30. Willingness to take risks then remains flat before decreasing again from their mid-50s (the authors note that this is not solely driven by childbirth). However, a recent survey of the experimental literature on risk taking and gender overall found that "the absence of conclusive results makes further research necessary to properly identify when and why males are more risk tolerant than females" (Filippin and Crosetto, 2016). Gender differences appear to correlate with certain features of studies' risk elicitation methods (e.g. the presence of a safe option as well as fixed probabilities). It is also possible that there are no differences in risk attitudes, but young men are more likely to experience health concerns that are easier to correctly identify as being non-urgent or are less "risky". I do not have data on the health concerns that patients presented with at A&E, so am unable to investigate this further.

It is worth noting that I cannot exclude the possibility that there were differential opening rates for the intervention letter across different groups. All estimates should be considered as intention-to-treat (ITT), as we cannot be sure that the letters were opened. However, given that they were sent by the hospital to people who had recently attended, the expectation was that people would want to know what was in the letter. In addition, the letter was sent to individual patients, but the information may have been shared, for example with other household members. It was not technically possible to randomize treatment at the household level, so it may be the case that the treatment effect of the intervention is underestimated due to possible spillovers. More broadly, there are interesting open questions relating to the role of social influences and networks in decisions relating to demand for healthcare. Unfortunately I do not have data about the living arrangements and relationship status of participants in this field experiment, otherwise it would be interesting to investigate whether the observed differential treatment effects for men are related to these (e.g. does the intervention only work for single men).

Of course these are not the only possible explanations for the observed effects. Thinking about the feedback component of the intervention, if patients experience health shocks in each period, feedback about a past health problem is only informative about current health concerns to the extent that the health problems are similar across time periods. The conceptual framework outlined above also does not capture that patients may in fact simply not remember the feedback and information they received. While the design of the intervention aimed to provide timely feedback, shortly after a patient had attended A&E, the moment that this information would need to be used may in fact be weeks or months ahead in the future. This is what Rogers and Frey

(2015) refer to as the intervention-behaviour lag. Finally, the choice to attend A&E may simply be an emotional, fast, and automatic 'System 1' decision for many people - fundamentally difficult to shift through information and feedback.

Another question that would be interesting to explore further is the potential of spillovers to other health services. Given the lack of alternative Emergency Departments in the immediate area of the study, it is unlikely that the intervention would have increased demand at A&E departments not included in my data. To the extent that such spillovers divert future attendances from A&E to more appropriate services such as NHS 111 or GP practices, this could be considered part of the intended effect of the treatment, leading patients to be treated in more appropriate and economical settings. Of course, the intervention could potentially also generate additional demand for other health services as an unintended consequence. Unfortunately, I am not able to link the data used in this study to data recording patients' use of other health services.

This field experiment was the first to provide timely tailored information and feedback to patients in the high-stakes context of emergency care. As outlined above, there are several possible explanations for the observed differential treatment effect of the intervention. Further research is needed to gain a better understanding of the drivers of A&E attendances in different populations and the best ways to ensure patients receive care in the settings most suited to their health concerns.

Chapter 2

Sustained Change - a Field Experiment in Healthcare

2.1 Introduction

The extent to which the effects of a behavioural intervention are sustained is of interest to both academics and policy makers, yet there are only few previous studies addressing this issue and none that look at the behaviour of professionals in the workplace. I undertake a unique natural field experiment in the English National Health Service (NHS) to study this question.

The intervention I exploit consisted of changing the salience of waiting time on the clinicians' e-Referral Service (ERS) interface. Waiting longer can negatively affect patients' health gains from receiving treatment. Controlling for pre-surgery health, Nikolova et al. (2015) find that waiting time has a small but statistically significant negative impact on the health gains from hip and knee replacement surgery. They note that each year an additional week of waiting is associated with health losses worth £11.1m and £11.5m for hip and knee replacement patients in England respectively.

In this experiment I evaluated simple changes to the results screen visible to General Practitioners (GPs) searching for a service to refer their patient to through the ERS. In the ERS referral process, GPs create a shortlist of services for patients. They or their practice staff then either book one of these services for the patient or send the list to the patient, who books the service of her choice out of the shortlisted options. The intervention consisted of a simple change to the ERS system: adding a red 'limited capacity' alert to services with very high waiting times. It was implemented at a hospital trust keen to trial this as a way to help reduce demand for their services in 4 clinical specialties with low capacity. The evaluation was conducted through a stepped wedge design and the limited capacity messages were added to a new specialty every fortnight. The order in which the clinical specialties crossed from the control to the treatment group was randomized in advance of the trial.

The experiment demonstrated that putting a simple alert against services with limited capacity led to a 35 percent reduction in the share of referrals to these services. The effect of the intervention was sustained over time and did not vary with prior referral habits. I argue that the underlying mechanism causing this reduction is the increased salience of waiting time for the treated services. A possible alternative ex-

planation (also considered alongside salience by Chetty et al. (2009) in their study of consumer taxation) could be that agents lack information. However, a separate randomized controlled trial (in a different location) found that providing GPs with similar information about limited capacity in a letter did not lead to a statistically significant change in referrals. In this case, waiting time information was made available, but not necessarily at the time of referral, strengthening the case for salience. This trial is the first to rigorously test the sustained impact of information provision and choice architecture on referral behavior in a real-world setting.

This chapter contributes to three distinct strands of literature in economics. The first is the literature on longer term effects of behavioural interventions and the impact of these interventions when they are repeated over time. An important question for the longer term effectiveness of my intervention is whether its impact starts to decay over time as referrers get used to seeing the alerts, or whether its effects are sustained, i.e. there are marginal benefits to continued treatment, as discussed by Rogers and Frey (2015). There are only few studies that address this question and none that look at the behaviour of professionals in the workplace. In a health context, for example, Altmann and Traxler (2014) find that sending reminders repeatedly to patients seems to neither strengthen nor weaken their effects.

Second, the literature related to rational inattention and boundedly rational consumers. Taylor and Thompson (1982) describe salience as referring to "the phenomenon that when one's attention is differentially directed to one portion of the environment rather than to others, the information contained in that portion will receive disproportionate weighing in subsequent judgments"(Taylor and Thompson, 1982). The effects of salience on consumer behaviour have been documented in various different contexts, including commodity taxes (Chetty et al., 2009), highway tolls (Finkelstein, 2009), electricity consumption (Gilbert and Zivin, 2014), social security benefits (Brinch et al., 2017), and law enforcement (Dur, 2017). Neuroeconomic research provides a neurobiological foundation for the role of salience in choice. Fehr and Rangel (2011) note that directing attention towards certain attributes should increase the weight that agents place on these in computing decision values and thus affect choices. They argue that this "provides a neurobiological foundation for the effectiveness of some marketing and behavioral public policies" (Fehr and Rangel, 2011). Nudges, in their view, are prime examples illustrating the role of attention in economic choice. Bordalo et al. (2013) develop a model of context-dependent choice that reflects this view of decision making, namely that consumers' choices are shaped by the most salient aspects of their choice context. In this model, consumers' attention is drawn to the attribute of a good that stands out most compared to that attribute's average level. Bordalo et al. (2015) further extend this model by adding limited memory, thus combining a model of limited recall based on Norm Theory (Kahneman and Miller, 1986) with a model of context dependent choice in which consumers' attention is drawn to salient features of choices.

The third is the impact of choice and competition in healthcare provision. This study advances the current literature by introducing exogenous variation in the salience of waiting time for certain clinical services and by using patient-level data to study the impact on choice. In the UK context, several papers have recently studied the impact of the 2006 reforms extending patient choice in the English NHS. Estimating reduced form models of hospital mortality outcomes, Gaynor et al. (2013) and Cooper et al. (2011) find that competition and choice improved patient outcomes at the hospital level. Gaynor et al. (2016) build on this work by explicitly modeling the choice process

both pre- and post-reform.

The rest of the chapter is structured as follows: Section 2 develops a conceptual framework while Section 3 describes the institutional context. Section 4 presents the intervention as well as the experimental design and empirical strategy. Section 5 lays out some descriptive statistics and results are presented in Section 6, while Section 7 concludes the chapter.

2.2 Conceptual Framework

In the referral process, GPs choose which services to shortlist or book from a list of services. Most searches on the ERS will by default generate a list of 99 options to choose from, with about 16 displayed on the first results page. These services differ along a number of dimensions, including distance to a patient's home address (by default this is how choices are ordered), quality, and waiting time. Referrers have beliefs about these factors based on information available to them. Following Gaynor et al. (2016) I assume that GPs try to choose the best service for their patients (they do not have any financial incentives to do otherwise), subject to the information that they have and the cost associated with finding the best service.

Note that if GPs were fully rational and well-informed, the intervention in my experiment should have no effect on the services they choose. Furthermore, providing the same information repeatedly, as the intervention does through the continued application of the 'limited capacity' alert to the same service, should not continue to affect referrer choices. However, in practice, evaluating the different options costs time and there is also a mental cost involved in remembering the different pieces of information about the various services. When the costs of time and attention are high, we might expect people to rely on shortcuts and perhaps habitually refer to a small set of services that they remember. I hypothesize that there are two mechanisms through which my intervention may affect referrers' decisions.

First, the learning channel: GPs are provided with information about one of the factors that affects the benefits of a particular service (i.e. waiting time). They update their beliefs and refer less to the place they have learned has high waiting times. However, even if they update their beliefs, they might forget over time. Second, the attention channel: Perhaps GPs already have updated their beliefs (they may have seen the intervention on a previous occasion). However, it is costly to remember this information, so reminding them of the high waiting times at the time of referral helps them bring this information back to mind.

These two channels are quite closely related. In terms of the intervention, they illustrate the difference between providing information and providing that information in a way that cannot be ignored in the moment a decision is made. I hypothesize that the attention channel is likely to play a larger role in this choice context, as the costs of remembering waiting time information are high, however, it is not possible to distinguish the two channels from each other clearly.

2.2.1 Sustained Behavioural Change

Rogers and Frey (2015) note research in classical conditioning has shown that repeatedly applying the same intervention should be unlikely to continue to have the same effect that it did initially as people are desensitized by repeated exposure to a given stimulus.

(see Rankin et al. (2009); Thompson and Spencer (1966)). Despite this, there have been behavioural interventions that have generated marginal benefits to continued treatment. In the area of health, for example, a recent study analyses how people behave when they repeatedly receive reminders for medical check-ups, by randomly assigning the incidence of being reminded and the type of reminder patients receive (Altmann and Traxler, 2014). The authors find that applying reminders repeatedly neither strengthens nor weakens their effects, addressing a concern that people may become dependent on nudges, crowding out people's active choices. Rogers and Frey (2015) discuss three features of behavioural interventions that might explain why they remain effective upon repetition. This includes interventions: 1) resisting attention habituation; 2) those producing incomplete or temporary change in behavior; and 3) interventions inducing little resistance or distrust.

In the context of the intervention I study, I believe that the first challenge - resisting attention habituation - is key to sustaining the change in referral choices. Rogers and Frey (2015) identify four features of interventions that can avoid attention habituation from occurring, citing Rankin et al. (2009) and McSweeney (2004). These are related to the nature of the intervention (ideally an intense and dynamic stimulus) and its timing (preferably long or unpredictable intervals between repeated instances of the intervention).

Note that the question I address here is distinct from that of persistence - whether the effects of the intervention are sustained when it is discontinued. Surveying previous field experiments that used financial incentives to motivate behavioural change, Brandon et al. (2017) note that only in four out of ten studies was there evidence of persistent behaviour change after removal of the incentives, consistent with habit formation, and even when effects persisted they tended to decay rapidly. Using data from 38 field experiments that used social norms-based interventions to reduce energy consumption, the authors provide further evidence that the persistence of the effect in these experiments is likely attributable to changes in physical capital in the home rather than habit formation in the household. I was not able to study persistence upon removal of the intervention due to ethical considerations (removing the successful intervention could have led to worse patient outcomes). However, as the intervention appears to operate through increased salience at the time a decision is made, I would not expect it to persist if it were to be removed.

2.3 Institutional Context

In this section I will describe the context of choice in the English NHS and the referral process. Healthcare in England is free at the point of use and mostly provided by the NHS, financed by taxes. The country is split into 209 Clinical Commissioning Groups (CCGs) who are responsible for commissioning services in their area by assessing the needs of and buying services for the population there.[1] Primary care is provided by General Practitioners (GPs) who can refer patients to hospital-based (secondary care) services for specialist treatment and diagnostics.

Currently around half of all referrals are made through the e-Referral Service (ERS) - an online platform through which services can be searched and booked.[2] GPs are

[1] https://www.nhscc.org/ccgs/

[2] https://www.england.nhs.uk/2016/03/hospital-referrals/

provided with financial incentives to use this service,[3] as part of a plan to shift all referrals from second-class mail to this faster channel. Using the ERS, GPs create a shortlist of suitable services for their patient during the appointment (and ideally in consultation with their patient). They or their practice staff then either book one of these services for the patient or send the list to the patient, who books the service of her choice out of the shortlisted options online or over the phone.

2.3.1 Patient Choice in the NHS

While patients had been able to choose their GP since the creation of the NHS in 1948, hospital choice was first opened up to patients in 2006, following some earlier pilot programs. In 2008 patient choice was extended to include any hospital listed in a national directory of services - 'free choice' of services including both NHS and independent sector providers contracted to provide services at NHS tariffs. Finally, the NHS Constitution made free choice of hospital provider a patient right in 2009.[4] (Dixon et al., 2010)

Patient choice was intended to increase hospital quality and efficiency by creating competition between hospitals (Boyce et al., 2010). For it to have this impact, the policy assumes that patients are in fact offered choices and that patients have information about hospital quality which they use in choosing where to receive treatment. An evaluation of the patient choice reforms found that, while GPs stated that they always offered their patients choice, only 45% of patient reported being offered a choice (Dixon et al., 2010). The most recent survey evidence finds that only 40% of patients who were referred to an outpatient appointment by their GP within the last 12 months recalled having been offered a choice of hospital (Populus, 2015).

Several papers have recently studied the impact of the 2006 reforms extending patient choice in the English NHS. Gaynor et al. (2013) and Cooper et al. (2011) estimate reduced form models of hospital mortality outcomes and find that competition and choice improved patient outcomes at the hospital level. Gaynor et al. (2016) build on this work by explicitly modeling the choice process both pre- and post-reform. As prices in the NHS are zero for the consumer, the authors model patient utility as a function of patient and hospital characteristics with quality of care, waiting time, and distance to hospital the key determinants of hospital choice.

2.3.2 Factors Affecting Choice

Patient choice was first piloted for patients who had been waiting for treatment for some time - initially in London (from October 2002 to April 2004) and only for cardiac surgery patients across England (from July 2002 to November 2003) (Boyce et al., 2010). The London Patient Choice Project (LPCP) focused on patients who had been waiting for treatment at a NHS hospital in London. As patients' waits neared the 6-month mark, they were offered the opportunity to receive treatment from a range of alternative providers where they would receive treatment sooner. Burge et al. (2005) evaluated this project and found that 65% of patients took up this opportunity to receive treatment sooner. Analyzing the revealed preference data the report further notes that patients seemed to try to minimise waiting and travel time, while trying

[3] https://www.england.nhs.uk/wp-content/uploads/2016/03/qualty-prem-guid-2016-17.pdf
[4] http://www.nhs.uk/Tools/Documents/The%20history%20of%20choice.htm

to receive treatment at a hospital perceived as offering high quality care. Notably, reputation seemed to be an important factor for choice, with patients placing a high negative valuation on alternative services with a worse (or simply unknown) reputation compared to the hospital they were currently assigned to.

While salience and choice architecture has not been studied in a real-world referral setting, Boyce et al. (2010) conduct a hypothetical choice online experiment of hospital choice. Unlike my experiment, this study focuses on the way information is presented to patients rather than GPs. The authors find that people lack stable preferences when it comes to choosing a hospital, suggesting that this provides an opportunity to influence patient choice by making certain aspects of hospital care more salient. Using an online survey, they presented a representative panel of the English population with a scorecard comparing five hospitals that differed on various indicators of quality and other factors. The study tested several different nudges to help participants choose the highest quality hospital. The most successful intervention was what the authors call the 'feedback' nudge, which prompted participants to reconsider their choice if they had not selected the highest quality hospital. Viewing this scorecard, participants chose the highest quality hospital 56% of the time, compared to 44% in the control group. The authors conclude that their results emphasize the importance of how information is presented in this choice context.

2.4 Field Experiment

2.4.1 Intervention

The intervention consisted of simple changes to the results screen visible to GPs searching for a service to refer their patient to through the ERS. In this referral process, GPs create a shortlist of services for patients. They or their practice staff then either book one of these services for the patient or send the list to the patient, who books the service of her choice out of the shortlisted options.

The intervention involved two changes to the ERS system: Regional commissioners were able to manually identify services with low capacity and apply a red box against those services containing the label 'limited capacity' in the lower part of the screen. They were also able to manually identify services with high capacity and make it possible for them to appear in a green box at the top of the screen. Example screenshots without and with the intervention are included in Appendix 2B.

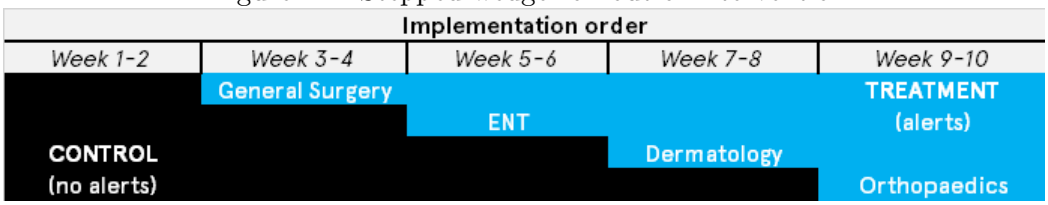
This functionality was only available to commissioners in the trial area consisting of three adjacent Clinical Commissioning Groups (Barking and Dagenham, Havering, and Redbridge). The local NHS hospital trust (Barking, Havering and Redbridge University Hospitals NHS Trust) was keen to trial the 'limited capacity' labels in order to help address the problem of very low capacity in 4 clinical specialties. For each of these clinical specialties, commissioners identified alternative local services with available capacity.

2.4.2 Methods

Experimental Design

The evaluation was conducted as a stepped wedge design. The intervention was sequentially rolled out to 4 clinical specialties at BHRUT, as depicted in Figure 2.1 below, i.e. the limited capacity and available capacity messages were added to a new specialty every fortnight. The order in which the clinical specialties crossed from the control to the treatment group was randomized in advance of the trial.

Figure 2.1: Stepped wedge roll-out of intervention



When the alerts were switched on, the intervention was visible to any GPs within the Barking & Dagenham, Havering and Redbridge CCG areas searching for services belonging to the specialties included in the trial. Crucially, this intervention did not remove or add any referral choices for GPs. If a service with 'limited capacity' was selected, GPs were shown a pop-up box with a limited capacity warning (see figure A.2.3 in Appendix), but could proceed with the selection if they so chose. This meant that where a referral to a low capacity service was clinically appropriate, the patient would not receive a barrier to referral and care.

The data used in this trial were all routinely collected by the Health and Social Care Information Centre (HSCIC), which has now become NHS Digital. The time of the referrals is defined as the time when a new referral request is created (equivalent to when the GP/practice staff press a button in the ERS which starts the process of making a referral shortlist). A referral request can be linked to multiple later actions (i.e. entry onto a waiting list (called ASI or Appointment Slot Issues list), booking, cancelling, rebooking) related to a referral. I look at the first action related to a unique referral request (which could be either a booking or entry onto a waitlist) and record the destination (i.e. which service is booked or for which service the patient is placed on a waitlist).

Ethics approval was obtained for this trial from the Ethics Committee of the University of Edinburgh School of Economics.

Empirical Strategy

I receive data at the individual referral level (when it was made, who the referrer is, and the specialty and service referred to). I estimate a logit model on the sample of referrals during the trial period (i.e. during the implementation of the intervention and two weeks before) where the dependent variable captures the probability of a referral going to one of the treated services. T_{it} is a dummy variable indicating treatment and the model includes a vector of dummies, Γ, for the clinical specialties included in the trial as well as date-specific dummies to capture any day-by-day time trends in referrals.

$$Y_{it} \overset{Indep.}{\sim} Bernoulli(p_{it}); logit(p_{it}) = \alpha + \beta_1 T_{it} + \Psi_{it}\Gamma \tag{2.1}$$

I conducted power calculations for the stepped wedge trial following Hussey and Hughes (2007). The study was powered to detect a three percentage point change in the referral rate (5% significance level, 80% power).

2.5 Descriptive Statistics

The intervention was implemented from 20th July to 14th September 2016 and the analysis also includes data from the fortnight before the intervention was rolled out (i.e. when all four specialties were in the control group).

The following analysis is based on data for electronic referrals in the trial area for four clinical specialties included in the trial (General Surgery; Ear, Nose & Throat (ENT); Dermatology; and Orthopaedics). These figures include patients who booked into available appointment slots at these services and those who went onto waiting lists (referred to as 'ASI' or Appointment Slot Issue lists).

During the trial period there were 8,273 unique referrals in the area for the four specialties included in the trial. Table A.2.1 in the Appendix shows how these referrals are distributed amongst the different specialties included in the trial and the proportions of referrals that result in bookings and waitlist entries. Orthopaedics accounts for most referrals (51 percent), while the fewest bookings are for General Surgery (10 percent).

Table 2.1 shows the numbers of referrals assigned to the treatment and control groups. The specialties are listed in the order in which the intervention was rolled out. As should be expected, there are relatively more referrals in the treatment group relative to the control group for the specialties that received the intervention earlier.

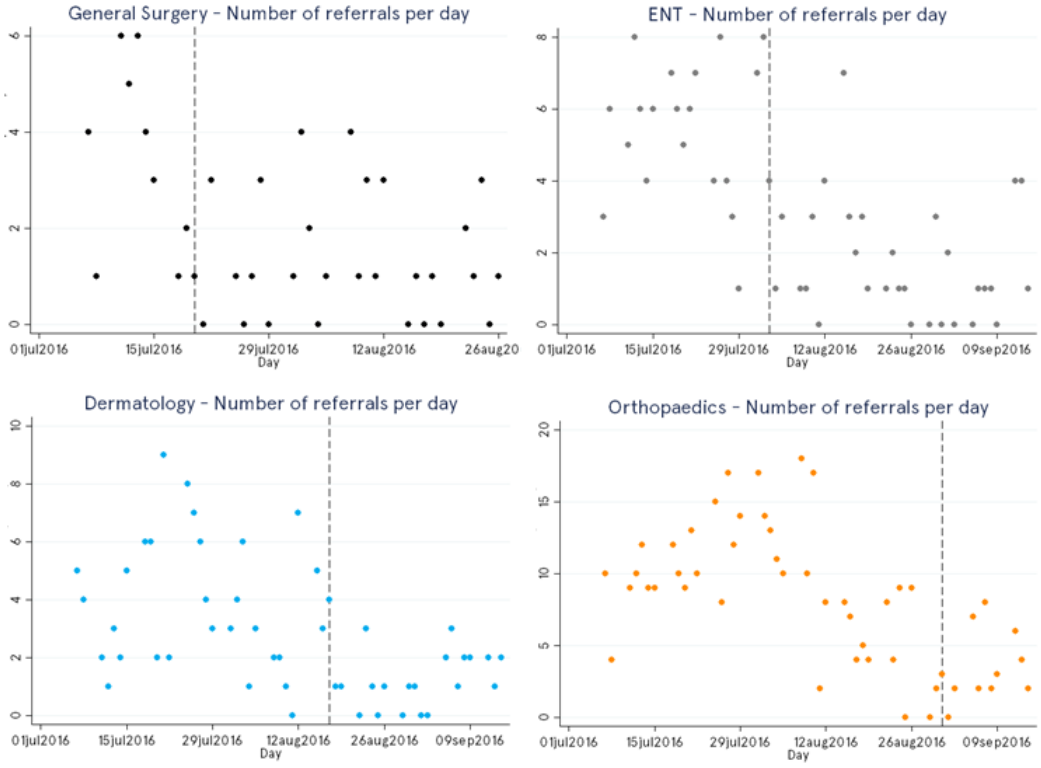
Table 2.1: Referrals during trial period - assignment to treatment and control

Specialty	Control	Treatment	Total
(1) General Surgery	201	634	835
(2) ENT	695	757	1,452
(3) Dermatology	1,138	601	1,739
(4) Orthopaedics	3,411	836	4,247
Total	5445	2828	8,273

The aim of the intervention was to reduce referrals to treated services. Looking at the outcome data, referrals to treated services at BHRUT represent 10.9 percent of referrals. Figure 2.2 shows the sum of referrals to treated services at BHRUT during

the trial period, with each vertical dashed line indicating the introduction of the intervention to an additional clinical specialty (starting with General Surgery on July 20th, followed in 14 day intervals by ENT, Dermatology and finally Orthopaedics). Looking at this panel of graphs one can see that for each specialty, the number of daily referrals appears to be lower after the introduction of the intervention.

Figure 2.2: Sum of daily referrals per specialty to treated services at BHRUT during trial period



2.6 Results

The intervention led to a statistically significant reduction in referrals to services with the 'limited capacity' alert when treated, as shown in Table 2.2. To better describe the magnitude of these results, I calculated the actual percentage of referrals to treated services when the intervention was in place (5.1 percent) compared with the share of referrals that the logistic model predicts would have gone to these services had the intervention not been introduced (7.9 percent) - a reduction of 35 percent. In terms of absolute numbers of referrals, this result means that when the limited capacity flags were in place, 145 referrals went to places with a limited capacity alert. If the alerts had not been in place, there would have been 222 referrals to these services.

Column 2 of Table 2.2 shows results from the same model as presented in column 1, but with a sample that only includes referrals to acute providers, which might be considered closer substitutes for the services provided by BHRUT. Further robustness checks are presented in the remainder of Table 2.2. One potential concern is that seasonal or time trends could play a role in inflating the observed treatment effect. This is why the main specification of the model controlled for time trends using day-specific fixed effects. As a further robustness check, I replace day fixed effects with week fixed effects, which makes only a very small difference to coefficients and standards errors. The results are also robust to adding GP-specific fixed effects, as shown in column 4.

Table 2.2: Impact of the Intervention on Propensity to Refer to Treated Services when Treated - Logit

	(1) Main	(2) Acute services only	(3) Week fixed effects	(4) Week and GP fixed effects
Treatment	-0.469***	-0.552***	-0.476***	-0.467**
	(0.160)	(0.163)	(0.183)	(0.191)
ENT	0.504***	-0.413***	0.515***	0.573***
	(0.120)	(0.122)	(0.124)	(0.145)
Orthopaedics	0.267***	-0.473***	0.269**	0.397***
	(0.0928)	(0.0942)	(0.107)	(0.111)
General Surgery	0.343*	-0.771***	0.359*	0.304
	(0.184)	(0.186)	(0.190)	(0.194)
Constant	-2.024***	-0.927***	-2.336***	-1.910***
	(0.199)	(0.209)	(0.122)	(0.431)
Day fixed effects	yes	yes	no	no
Week fixed effects	no	no	yes	yes
GP fixed effects	no	no	no	yes
Observations	8,244	5,544	8,273	7,695

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In order to explore whether the treatment effects are driven by any particular specialties and to start exploring potential spillovers across specialties, as an additional robustness check, Table 2.3 shows the results for the main model, but sequentially excluding each one of the specialties from the analysis. The main result remains robust to this check - it does not appear to be the case that the effect of the treatment is driven by any particular specialty.

Table 2.3: Robustness check: sequential exclusion of different specialties - Logit

	(1) Excl. Derma	(2) Excl. ENT	(3) Excl. Ortho	(4) Excl. Gen Surgery
Treatment	-0.473***	-0.460**	-0.408**	-0.513***
	(0.172)	(0.227)	(0.190)	(0.186)
Dermatology		0	0	0
		(.)	(.)	(.)
ENT	0		0.474***	0.517***
	(.)		(0.108)	(0.117)
Orthopaedics	-0.241**	0.268***		0.268***
	(0.0940)	(0.0916)		(0.0896)
General Surgery	-0.166	0.360*	0.286	
	(0.171)	(0.203)	(0.195)	
Constant	-1.595***	-1.869***	-1.910***	-2.158***
	(0.228)	(0.201)	(0.345)	(0.178)
Day fixed effects	yes	yes	yes	yes
Observations	6,509	6,579	3,945	7,408

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I also evaluated the impact of the intervention on referrals to services that were highlighted in the green box signifying 'available capacity'. I was not able to detect a statistically significant change in the rate of referrals to these services when the 'available capacity' alerts were in place. The results of the logistic regression model are shown in Appendix 2A.

The results show that with the alerts in place 15.3 percent of referrals went to services with an 'available capacity' alert. If the alerts had not been in place, 14.8 percent of referrals would have gone to these services. However, there were some problems with the implementation of these 'available capacity' alerts. During the trial period these services did not consistently have slots available and therefore did not always meet the criteria necessary for them to appear in the green box. The results presented here consider services as having been 'treated' if the alerts were present on at least two out of three spot checks conducted.

2.6.1 Sustained Effects of the Intervention

The effects of the intervention are sustained over time - I do not find any evidence of the effect of the intervention decaying. To analyze this I created a variable that indicates for each referral how long the intervention had been in place for the specialty. There is no statistically significant impact of the number of days the intervention has been in place on the likelihood to refer to a treated service.

Table 2.4: Impact of Number of Days Intervention is in Place on Propensity to Refer to Treated Services when Treated - Logit

	(1)
Treatment	-0.431**
	(0.169)
Treatment x Num. days since first treated	-0.00560
	(0.00837)
ENT	0.517***
	(0.119)
Orthopaedics	0.258***
	(0.0937)
General Surgery	0.392**
	(0.194)
Constant	-2.028***
	(0.197)
Day fixed effects	yes
Observations	8,244

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

One factor that might be important for the longer term effectiveness of the intervention, as outlined by Rogers and Frey (2015), is the extent to which 'attention habituation' can be avoided. Given that the pop-up feature of the 'limited capacity' alerts (only if a service with limited capacity is in fact selected) makes them relatively harder to ignore, it seems likely that this helps to prevent users from getting used to and ultimately starting to ignore the intervention. 'Alert fatigue' (see Ancker et al. (2017)) is a potential concern about alerts, however, the sustained effects of the intervention indicate that this is not a problem in the context of this study. When scaling this intervention, however, it is worth bearing in mind the potential of alert fatigue setting in as the alerts become more frequent.

2.6.2 Prior Referral Behaviour

Another question of interest is whether the intervention has a differential effect depending on how frequently a doctor referred patients to the treated services before the intervention. I create a variable that captures the share of referrals each GP made to the treated services prior to the intervention (using data from March 2016 to the start

of the intervention). Prior propensity to refer to the treated services does not appear to have any impact on the effect of the intervention - behaviour seems to be affected similarly regardless of how frequently GPs referred to the treated services prior to the intervention.

Table 2.5: Impact of Prior Propensity to Refer to Treated Services on Propensity to Refer to Treated Services when Treated

	(1)	(2)	(3)	(4)
Treatment	-0.523***	-0.382	-0.560**	-0.507**
	(0.167)	(0.294)	(0.251)	(0.256)
Percent prior referrals	0.0709***	0.0725***		
	(0.00430)	(0.00468)		
Treat x pct. prior referrals		-0.00652		
		(0.0110)		
Above median prior referrals			0.938***	1.404*
			(0.0971)	(0.735)
Treat x above median prior referrals			0.126	0.0277
			(0.250)	(0.257)
ENT	0.553***	0.555***	0.553***	0.571***
	(0.130)	(0.131)	(0.125)	(0.139)
Orthopaedics	0.390***	0.392***	0.327***	0.402***
	(0.0978)	(0.0981)	(0.0949)	(0.105)
General Surgery	0.335*	0.336*	0.322*	0.311*
	(0.182)	(0.182)	(0.183)	(0.186)
Constant	-3.369***	-3.401***	-2.597***	-2.952***
	(0.226)	(0.231)	(0.215)	(0.644)
Day fixed effects	yes	yes	yes	yes
GP fixed effects	no	no	no	yes
Observations	8,244	8,244	8,244	7,669

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To facilitate interpretation of the results, I also create a dummy variable indicating whether a GP was previously an above-median referrer to the treated services. Clearly having previously referred frequently to the treated services is highly predictive of referring to these services in the future (see column 3 in table 2.5, but again there is no statistically significant interaction effect between this variable and the treatment. As a further robustness check, I include GP-specific fixed effects to this specification of the model, but this again does not meaningfully change the observed results.

2.7 Discussion and Conclusion

In this natural field experiment I demonstrate that a small intervention - placing a simple alert against services with limited capacity - leads to a 35 percent reduction in the share of referrals to treated services. Not only does this small intervention change behaviour, I also find that it does so regardless of prior referral behaviour and that the change in behaviour is sustained over time.

I argue that this is due to the alert increasing the salience of waiting time, which leads to more weight being placed on this factor in the choice process. An alternative explanation could be that agents simply lack information. If this was the case, providing them with information on waiting times once should be enough to bring about a change in their referral behaviour. A separate randomized controlled trial (in a different location) did precisely this - providing GPs with information about limited capacity in a letter. The trial did not find a statistically significant change in referrals. In this case, waiting time information was made available, but not necessarily at the time of referral, strengthening the case for salience. Furthermore, I find that referrals are reduced only for the clinical specialties that have 'limited capacity' alerts in place - there are no spillovers to non-treated specialties, also supporting the hypothesis that GPs only use the information about waiting times when it is made salient to them.

While I did not detect a statistically significant difference in referrals to services flagged as having available capacity, this may be due to problems with the implementation of these alerts and more research is needed to establish whether patients can be nudged towards services with available capacity by making these salient or whether people react more strongly to making "bad" high waiting times salient, consistent with a model of loss averse actors.

This trial was the first to rigorously test the impact of information provision and choice architecture on referral behavior in a real-world setting. Scaling up the intervention it will also be interesting to research whether the effects of the intervention change when there are multiple 'limited capacity' alerts rather than just one. Making small changes to the salience of factors involved in choice can be a powerful tool to support decision making in complex choice situations.

Chapter 3

Mind, Behaviour and Health - a Randomised Experiment

3.1 Introduction

The growing prevalence of non-communicable diseases (NCDs), such as cardiovascular diseases, cancer, chronic respiratory diseases and diabetes in both developed and developing countries[1] has triggered significant research interest from economists and behavioural economists. These diseases are often referred to as "lifestyle diseases", that is, diseases which are at least in part caused by people's lifestyle. The standard economic approach to lifestyle choices is to portray them as outcomes of a decision process involving a trade-off between immediate benefits (such as eating tasty foods) and long-term consequences (in terms of health in particular). This trade-off involves risks and is intertemporal in nature - the decisions taken today have consequences in the future, which are assumed to be discounted. In this type of model, risk and time preferences play a key role in predicting the extent to which people engage in healthy and unhealthy behaviours. More recently, behavioural economists have shown that people have a range of "behavioural biases" such as biases toward immediate gratification. These biases can explain inconsistencies in people's choices over time, for example choosing to go on a diet and then not sticking to it (DellaVigna and Malmendier, 2006; Gul and Pesendorfer, 2004). These theories have received empirical support (Downs et al., 2009) and the insights are now increasingly used in policy design (Marteau et al., 2012).

Policies inspired by insights from behavioural economics take behavioural traits and biases as given and often propose tools that seek to exploit these behavioural biases to the advantage of individuals. For example, by allowing workers to pre-commit to later increases in their savings rate for retirement, Thaler and Benartzi (2004) exploit workers' present bias to help those who would like to save more but lack the willpower to do so. There has, however, been relatively little interest in whether one could actually try to correct for these biases or even try to alter the decision-making processes that underpin these behaviours in the first place. As I outlined in the introduction to this thesis, the recent literature in Psychology and Economics proposes models of "dual systems" that govern decision-making (see Hofmann et al. (2008), Thaler and Shefrin

[1] According to World Health Organisation (2014), about 38 million (68%) of deaths worldwide in 2012 were associated with NCDs, 40 percent of which were considered preventable.

(1981), Loewenstein and O'Donoghue (2004); and Fudenberg and Levine (2006)). One is an impulsive system that operates effortlessly and automatically, whereas the other is a reflective system that makes reasoned judgements and forms action plans to pursue longer-term goals, overriding the automatic responses that are based on impulse or habit (Strack and Deutsch, 2004). Which of the two systems wins and ultimately determines behaviour is an important question. Psychologists (e.g., Hofmann et al. (2008)) have argued that a key determinant of the "bargaining power" between the two is self-control. The key question is whether it is conceivable to train the mind to become more self-controlled, shifting the balance of power between the two systems underlying decision-making and thereby affecting behaviours related to risk and time preferences.

In this paper, we make a first attempt at investigating the possibility of training the mind to alter fundamental cognitive processes underlying decision-making. The idea that the mind can be trained is not new and is in fact at the core of old and traditional activities such as meditation, martial arts and yoga, to cite a few examples (Diamond and Lee, 2011). To achieve this, we select one of the currently most popular techniques based on what is referred to as "mindfulness training". Mindfulness, a practice that combines meditation, breathing and yoga, has recently enjoyed a rise in popularity in many countries (Huffington Post, 2014). It has been described as a process of bringing a certain quality of attention to moment-by-moment experience (Kabat-Zinn, 1990) and consists of routine exercises such as bringing the mind's attention to the present (for example, by focusing attention on one's breathing or on what one is eating). These techniques are seen as ideal training to improve self-control, perhaps because most of the exercises focus on training the ability to inhibit one's impulses (Flook et al., 2010; Friese et al., 2012; Teper and Inzlicht, 2013; Teper et al., 2013). The direct objective of these techniques is often reducing stress, and there are a number of experimental studies documenting their effectiveness in reducing chronic stress (Caldwell et al., 2012; Morledge et al., 2013; Tang et al., 2007).

Because self-control and stress are believed to play a key role in decision-making in general and health-related behaviours in particular, mindfulness techniques appear to be a promising avenue for affecting fundamental cognitive processes underlying decision-making and, in turn, health-related behaviours. The question of whether it is possible to affect decision-making processes is important and could in principle open a new domain for policy interventions, although the welfare implications of shifting the balance between systems governing decision processes are not clear. In any case, it is an open question whether these processes are malleable or not. We attempt to shed light on this by conducting a randomised controlled experiment where we treat a sub-group of participants with an online course in mindfulness called "Be Mindful".[2] The course is designed as a complete training for mindfulness and is currently one of the most popular online tools for learning mindfulness skills. It is run by the UK Mental Health Foundation.

We invited 139 students from the University of Edinburgh to participate in a six-week study (with a five-month follow-up) on "lifestyle". Students with no pre-existing medical conditions were recruited and were invited to an initial session at the Behavioural Laboratory at the University of Edinburgh. They were allocated randomly

[2]See https://www.bemindfulonline.com/ for a detailed description of mindfulness-based stress reduction course.

either to a mindfulness-based stress reduction (MBSR) programme or to a control intervention consisting of a series of documentaries called "BBC Ancient Worlds". We chose this intervention because it requires a similar degree of time commitment, but involves very different activities. While mindfulness consists of exercises that should help individuals take charge of their thought processes, a TV documentary is more likely to be distracting. Both programmes were to be followed outside the laboratory and lasted for four consecutive weeks, starting in the week immediately after the initial session. Participants were asked to return to the laboratory for five consecutive weeks after the initial session (including one week after the interventions ended) and provide feedback on the previous week (about both their engagement with the intervention and their well-being and health-related behaviours). We also conducted an additional post-intervention session five months later to document their longer-term behaviour and see whether there was evidence of long-term behavioural changes.

Using student subjects to answer these research questions provides more than logistical advantages over other subjects. There is strong evidence that students suffer from chronic stress (Galbraith and Brown, 2011; Regehr et al., 2013) and are particularly prone to engage in unhealthy behaviours such as smoking, drinking and eating unhealthy food (Dallman et al., 2003; Kandiah et al., 2006).

We evaluate the effects of the mindfulness programme on a range of outcome variables. The first outcome of interest is to what extent participants engaged with each of the programmes. Because both programmes require some form of commitment, one might expect, for example, that impulsiveness and present bias could correlate with the ability to complete the programme. We find that indicators of stress, behavioural preferences, and health behaviours are not systematically related to the likelihood of completing the mindfulness course.

We then proceed to evaluate the effects of the MBSR programme on three sets of outcome variables. Since very little is known about whether and how mindfulness affects behaviour, we find it important to present a comprehensive picture of the effects on a range of variables directly relevant to health-related behaviours. We evaluate effects on measures of chronic stress, as well as on the response to a stressful situation (measured by cortisol and self-reports), because these outcomes are primary targets of the MBSR programme. We then study impacts on risk and inter-temporal attitudes, which are believed to play a key role in decisions related to health. Finally, we evaluate impacts on health-related behaviours. As these are very difficult to measure accurately, we include a range of measures, some self-reported, others incentivised. Our aim is to identify a common pattern in these variables and mechanisms through which mindfulness might impact them.

After completion of the mindfulness course, the program participants report significantly lower levels of stress than the control group, as measured by the Perceived Stress Scale (PSS). However, their physiological responses to an acutely stressful situation, as measured by cortisol levels, do not differ significantly. We also find suggestive evidence that participants in the treatment group are more patient, less likely to suffer from present bias and also less likely to engage in "stress-eating", although these effects are not all statistically significant due to large standard errors. We do not find much evidence of effects on other health-related behaviours such as sleep, smoking, alcohol consumption or physical exercise. Overall, our results suggest that mindfulness is effective at reducing stress and there is indicative evidence it may affect time preferences

and some health-related behaviours, but few of our results are statistically significant.

Of course, the intervention we looked at was relatively short in duration (four weeks), and we have post-intervention information only for participants who continued participating in the study. We also targeted a group of adults, for whom such interventions may be less effective than interventions earlier in life. Nevertheless, we believe this is an important research agenda that deserves attention by behavioural economists, and more research is needed to understand the extent to which it is possible to train the mind to overcome behavioural biases.

The rest of this chapter is structured as follows. Section 2 outlines related literature. Section 3 lays out the experimental design and procedure and describes the participant sample and the recruitment process. Section 4 describes the outcome measures of interest collected during the experiment. In Section 5, we present descriptive statistics on background variables, randomization checks and weekly surveys. Section 6 describes the empirical strategy and presents treatment effects on the main outcome variables of interest, and Section 7 concludes.

3.2 Related Literature

This paper contributes to a growing literature in economics on decision-making. One way of modelling decision-making that has received considerable attention in recent years is based on the idea that two separate and independent mental systems may be involved in decision-making (Fudenberg and Levine, 2006; Loewenstein and O'Donoghue, 2004; Thaler and Shefrin, 1981). The first system is an impulsive system that can trigger rapid decisions; the second system is more reflective and more cognitively involved. One can think of them as two participants in the decision-making process that have different (and often conflicting) preferences. Of course, a key question is which system ultimately determines behaviour. Psychologists (e.g., Hofmann et al. (2008)) attribute a large role to self-control in determining which of these two systems has the most bargaining power. In this context, mindfulness could be seen as a form of training to shift the balance of power between the two systems.

A number of studies in psychology explore the effects of mindfulness on executive function (Flook et al., 2010; Friese et al., 2012; Teper and Inzlicht, 2013; Teper et al., 2013). Because mindfulness consists of exercises that require individuals to focus on something specific for a period of time (one's own breathing, for instance), it could be a good form of training to increase self-control and self-regulation. Some studies have shown that engaging in short-term meditation practice improves executive function, as measured by performance on the Stroop task (Wenk-Sormaz, 2005). Research by Moore and Malinowski (2009) extends this finding by showing that meditators exhibit less Stroop interference than control participants. Using the Attention Network Test (Fan et al., 2002), a related work by Jha et al. (2007) documents that experienced meditators excel at conflict monitoring. Tang et al. (2007) provide additional evidence on this effect by showing that just five days of brief meditation training improved conflict monitoring on the same test. Finally, related research investigating attentional control has demonstrated that participants who completed a 10-day intensive meditation retreat showed significant improvements in attentional switching on the Internal Switching Task (Chambers et al., 2008). Semple (2010) added to this by showing that meditation practice improved sustained attention on the Continuous Performance Test (Rosvold et al., 1956). All of the above measures capture aspects of executive

functioning (Barkley, 1997), thus providing robust evidence on the connection between meditation and executive function. Further, mindfulness shares commonalities with activities such as yoga or martial arts that have been found to improve children's executive function (Diamond and Lee, 2011).

In addition to improving executive function, mindfulness is believed to reduce chronic stress. A growing body of research finds that mindfulness, especially mindfulness-based stress reduction (MBSR) and mindfulness-based cognitive therapy (MBCT), is an effective treatment for health problems such as recurrent depression (Ma and Teasdale, 2004; Teasdale et al., 2000) and anxiety (Hofmann et al., 2010). A recent systematic review of meditation programs, including 47 randomized clinical trials with active controls, found moderate evidence that mindfulness meditation programs reduce anxiety, depression and pain, as well as low evidence of stress reduction (Goyal et al., 2014).

While MBSR has been shown to be an effective treatment for various mental and physical disorders, fewer studies have investigated its possible benefits for "healthy" subjects. A review study by Chiesa and Serretti (2009), which undertook a meta-analysis of mostly less-rigorous studies published prior to 2008, documents that MBSR may be able to reduce stress levels in healthy subjects. However, the review emphasizes the need for further research to demonstrate a robust relationship between MBSR and stress. Several studies have since found evidence of persistent reductions in perceived stress (i.e. maintained at one- to three-month follow-ups) following participation in a mindfulness intervention (Carmody and Baer, 2008; Carmody et al., 2009; Epel et al., 2009).

Krusche et al. (2013) study the impact of the online mindfulness course we use in the present study and find significant reductions in perceived stress, anxiety and depression at course completion, as well as a further decline at a one-month follow-up. The authors report effects that are comparable to those found in studies using face-to-face mindfulness courses and other types of treatment for stress, such as cognitive behavioural therapy. The amount of (self-reported) meditation practice affected outcomes when the authors controlled for baseline levels of stress, anxiety, and depression. This study, however, has two key limitations: there was no control group, and the sample consisted of self-referred individuals who were willing to pay for and take part in the course, implying a potential sample selection bias.

By contrast, a recent systematic review (Goyal et al., 2014) concluded that there is little evidence of the effects of mindfulness on health-related behaviours such as eating habits and sleep, further highlighting that stronger study designs are needed in order to determine the impact of meditation programs on stress-related behaviour.

Because of its potential effect on stress, our work also relates to the literature linking stress and decision-making. Stress triggers a physiological response in the body and neuroscientists typically distinguish between short-run (acute) and long-run (chronic) stress. The effects of acute and chronic exposure to cortisol, the primary stress hormone, can be very different, and are in many cases opposite. Most studies focus on the effects of *acute* stress on decision-making.[3] There are a few studies that look at the effects of

[3]The evidence on causal effects of acute stress and risk-taking is mixed. A number of studies find that fearful emotions increase risk aversion (Guiso et al., 2013; Kugler et al., 2012), while other studies find that elevated cortisol is associated with more risk-taking (Putman et al., 2010; van den Bos et al., 2009). Other studies have found that, in stressful situations, humans are likely to fall back on automatized reactions to risk (Porcelli and Delgado, 2009). Regarding time preferences, a recent

chronic stress, indicated by the presence of elevated cortisol levels over longer periods of time. Chronic stress has been found to impair behavioural flexibility and attentional control (Liston et al., 2009; McEwen and Morrison, 2013; Radley et al., 2004). A recent experiment (double-blind and placebo controlled) raised cortisol levels of volunteers over a period of eight days to mimic the biological effects of chronic stress, replicating levels previously observed in real financial traders (Kandasamy et al., 2014). The study found that raised cortisol levels led volunteers to become more risk-averse and that men, relative to women, increasingly over-weighed small probabilities. This suggests that the physiological stress response in humans affects risk-taking behavior.

Similarly, animal studies have found that chronic stress biases rats' behavioural strategies toward habit, making them insensitive to changes in outcome value and compromising their ability to perform actions on the basis of consequences (Dias-Ferreira et al., 2009).

To summarise, our paper builds on earlier studies that: (i) model human behaviour as the result of dual processes - one impulsive and automatic, the other reflective and deliberate; and (ii) suggest that mindfulness techniques appear effective at improving executive function and at reducing stress, both of which are believed to play a key role in decision-making and in health-related behaviours. We specifically contribute to the literature by using a randomized controlled experiment to identify the impact of a mindfulness training program on behavioural traits and anomalies that play a key role in decision-making.

3.3 Experimental Design

3.3.1 Sample

We recruited 139 participants[4] primarily through the database of the Experimental Laboratory of the School of Economics at the University of Edinburgh – called BLUE (Behavioural Laboratory at the University of Edinburgh), as well as through posters and leaflets on campus. The advertisement and recruitment emails are attached in Appendix 3A.

Participants were required to be students at the University of Edinburgh, at least 18 years old, and they could not have any pre-existing medical condition. The experiment thus targeted a healthy student population. The study was approved by the School of Economics Ethics Committee at the University of Edinburgh. The slogan used in the advertisements was "Feeling a bit stressed?", targeting students with relatively high levels of anxiety at the start of the study. This was done in order to maximise the chances of inducing an exogenous difference in chronic stress between the treatment and control groups. However, it is likely that such a slogan would attract the attention of many students, as a recent survey by the National Union of Students (Kerr, 2013) found that 92 percent of respondents reported feelings of mental distress, including feeling down, stressed and demotivated during their time in higher education. Thus, it is likely that most students at the university "feel a bit stressed".

study by Cornelisse et al. (2013) finds that temporarily elevated cortisol induces people to prefer more immediate rewards to delayed rewards.

[4]We originally intended to have 144 participants (18 participants in 8 groups). We have a smaller sample size because 5 participants did not show up. Additionally, information on BMI is missing for one participant, which is why regressions controlling for BMI only include 138 participants.

It is important to point out, however, that, unlike in previous studies, the participants in our experiment did not self-select into the treatment and did not pay for it, reducing the risk of associated biases. The prospective participants did not know beforehand what the interventions would be.

3.3.2 Experimental Interventions

Treatment: Mindfulness Based Stress Reduction Programme

The Stress Management Programme consisted of the "Be Mindful Online" mindfulness course. Combining elements of MBSR and Mindfulness Based Cognitive Therapy (MBCT), the course was developed by leading UK mindfulness instructors and is run by the Mental Health Foundation and Wellmind Media. Participants are given an individual login to the course website (http://www.bemindfulonline.co.uk), which provides instructional videos to guide formal meditation. The impact of the course on stress and anxiety has been evaluated by Krusche et al. (2013).

The course is designed to be taken over four weeks, with a total of 10 interactive online sessions lasting 30 minutes each. The course starts with a three-minute introduction video. This is followed by a questionnaire (including the 10-item version of the Perceived Stress Scale (PSS) of Cohen et al. (1983)). It also contains the Patient Health Questionnaire (PHQ-9) and the Generalised Anxiety Disorder Assessment (GAD-7). This is followed by an orientation video, which also prompts participants to write down their intentions. During the course, participants are instructed in both formal (including sitting meditation and body scan) and informal (incorporating mindfulness into daily activities) meditation techniques, through videos, assignments, and reminder emails. Participants are asked to practise exercises for both kinds of technique each week between online sessions. Upon completing the course, participants are asked to complete the same questionnaire as in the introduction session of the course.

As participants were asked to follow the programme on their own, we could not enforce compliance. However, the online platform includes a web-based administration system to track participants' activity. In addition, weekly laboratory sessions were held to maintain engagement with the participants and gather self-reported information on their experience of the course (part of the weekly questionnaire, which also included questions about participants' feelings and health-related behaviours during the previous week). Thus, we are able to study in detail the extent to which participants engage with the programme.

Control Intervention: Historical Documentary Series

The control group was asked to watch the documentary series "BBC Ancient Worlds", which was provided to them via email link each week after their visit to the laboratory. This activity was chosen because it would require a similar amount of the participants' time as the MBSR protocol, in order to avoid making the treatment group busier and reducing the time available for health-related activities such as going to gym, etc. Participants in the control group were also asked to come to the laboratory once a week to fill in a questionnaire and provide feedback on the previous week's documentary, allowing us to track their degree of engagement with the programme.

It is possible that watching the BBC ancient world series itself might have a stress-

reducing effect. In order to explore whether such effects occurred, we asked the participants to evaluate how useful they found the documentary series for relaxation purposes as part of the weekly feedback. On average, the responses were neutral, indicating slightly lower relaxing effects than reported by the treatment participants for the MBSR intervention (see Appendix 3B for details). Thus, based on these statistics, we do not see evidence indicating that the control intervention would have had a stress-reducing effect.

3.3.3 Experimental Procedure

The experimental sessions started in October 2014 and were held at the same time and day every week for each participant, with a total of eight groups each week, spread over three different times on three days. In order to minimise the chance that students would find out about the other intervention, randomisation was conducted at the group level. Table 3.1 presents a summary of the experimental procedure. Sessions 1, 6 and 7 (pre- and two post-intervention sessions) were longer than the sessions that took place during the intervention.

Table 3.1: Experimental procedure

Session	Date	Content
1	Week of 20/10/2014 Pre-intervention	1. Lifestyle and stress survey 2. Saliva sample I 3. Stressful task 4. Decision making tasks 5. Saliva sample II 6. Further survey questions 7. Picture rating task 8. Saliva sample III
2	Week of 27/10/2014	feedback and short survey
3	Week of 3/11/2014	feedback and short survey
4	Week of 10/11/2014	feedback and short survey
5	Week of 17/11/2014	feedback and short survey
6	Week of 24/11/2014 Post-intervention	same as in Session 1
7	Week of 16/3/2015 5-month follow-up	1. Lifestyle and stress survey 2. Stressful task 3. Decision making tasks 4. Further survey questions 5. Picture rating task

The structure of Sessions 1 and 6 was as follows. Participants were publicly informed about the structure of the session. They then started the computerized survey, beginning with questions relating to their lifestyle and self-reported stress (including the PSS). When all participants had completed this section, the first sample of saliva was collected simultaneously from all participants in the group. This was followed by the stressful task.[5] The task was designed to be new to participants in each session in order

[5]The stressful task involved a combination of testing cognitive ability, time pressure, monetary

to avoid participants getting used to it, which could reduce its effectiveness as a stressor. After completing the task and providing feedback on its difficulty and stressfulness, participants proceeded with survey questions on decision-making and decision-making tasks. The second saliva sample was collected precisely 15 minutes after the end of the stressful task, at a time when a peak in cortisol concentrations in response to the stressful event should be expected. Decision-making tasks aimed at eliciting risk and time preferences followed, after which participants answered further background questions (including basic demographic questions in Session 1). The third cortisol sample was taken 23-24 minutes after the second one, by which time cortisol levels can be expected to have recovered. In order to provide participants with a neutral activity during the remaining time before the final cortisol sample could be taken, participants were asked to view a series of 30 pictures of humans and 30 pictures of wildlife, rating these respectively on attractiveness and how much they liked the pictures. This task was chosen to fill the time between the two saliva collections in a way that would allow for recovery from the stressful task. Finally, participants were called individually to receive their payments for the session.

Session 7 followed the same procedure as Sessions 1 and 6, but without collection of saliva samples. For Sessions 2-5, participants were asked to complete a short survey asking for feedback about their engagement with the intervention, as well as questions on their health-related behaviours during the previous week.

3.4 Hypotheses and Outcome Variables

We will now describe the outcome variables we are interested in, as well as our hypotheses regarding how these variables could be affected by mindfulness training. These include: (1) measures of engagement and compliance with the programmes, (2) measures of chronic stress and response to a stressful situation, (3) measures of behaviour related to risk and time preferences, as well as a self-reported measure of impulsiveness, (4) measures of health-related behaviours.

3.4.1 Attrition and Compliance

Our first key variables of interest are compliance with the treatments and attrition. The interventions received by both the treatment and the control group require commitment from the participants - in each case the interventions involve watching a video at home and showing up to the laboratory every week. We chose the control intervention so that the degree of commitment required would be similar and, therefore, we do not expect compliance and attrition to systematically differ across treatment and control groups. But one could expect, for example, that certain psychological characteristics such as impulsiveness, impatience and present bias may be correlated with the probability of dropping out. Because we collected a large set of variables at baseline, we are able to test this hypothesis directly.

Our first hypothesis is as follows:

Hypothesis 1 - Attrition rates will be similar across interventions and positively correlated with psychological characteristics such as impulsiveness, impatience and present bias.

reward/loss, and social pressure. Section 3.4.2 presents the task in detail.

We construct several measures to determine the degree of engagement of participants with the programmes. First, we record participants' attendance at each session. Second, we employ three different strategies to measure compliance with the programme. One is based on self-reports of engagement in various leisure activities, which are presented in a list format. Meditation is one of the listed activities and participants are asked to report how frequently they have engaged in each activity during the previous week. Another measure is based on summaries participants are asked to write about the contents of the latest lesson (MBSR intervention) or episode (control intervention) in each weekly session. We create an indicator to reflect accuracy of the report (equal to 1 if what they wrote is correct and 0 otherwise). The last measure is based on records of online activity that we obtained from the organisation running the online MBSR course. We have detailed information about the activity and progress of each participant. We use this information to construct a variable indicating how far the participants have progressed with the course.

3.4.2 Chronic Stress and Short-Run Response to a Stressful Situation

Because the mindfulness training aims at both decreasing overall anxiety levels and improving the ability to cope with stressful situations, we are interested in measuring both chronic stress levels and the short-run response to a stressful situation (similar to what a student is likely to encounter in her or his daily life).

Measures of Chronic Stress

Self-reported measures of stress are included in the survey questions completed by participants prior to beginning the stressful cognitive task. These measurements are based on the Perceived Stress Scale (PSS), using the 10-item version of the PSS (Cohen et al., 1983). We extend the PSS with two questions that measure academic stress, which can be particularly relevant among university students. The PSS is a widely used stress measure, capturing the extent to which an individual perceives events in the previous month as overwhelming or uncontrollable. Several studies of mindfulness interventions have reported reductions in PSS scores (see Krusche et al. (2013)). In our analysis, we use as an outcome variable the sum of the scores of the 10-item PSS version.

We also collected information on stressful events to which students may have been exposed. Sources of stress are measured with a substantially shortened version of the Adolescent Perceived Events Scale (APES, based on Compas (1987)), including a selection of questions most relevant to a student population from the 90-item APES. We use a variable indicating the sum of stressful events the participant faced in the previous month, and test whether her response (in terms of PSS score) differed across treatments. Because mindfulness is supposed to improve coping skills, the hypothesis is that participants in the MBSR treatment should respond less to stressful events.

Following most studies in the literature, we also collect self-reported measures of well-being,[6] asking respondents the following standard questions: "Overall, how satisfied are you with your life nowadays?" (in weekly surveys: the previous week), which we will refer to as "life satisfaction", and "Overall, how happy are you these days?", which

[6]The well-being questions were taken from the UK Labour Force Survey. See http://www.ons.gov.uk/ons/about-ons/get-involved/taking-part-in-a-survey/information-for-households/a-to-z-of-household-and-individual-surveys/labour-force-survey/index.html.

we will refer to as "happiness". We also ask how anxious they feel these days ("anxiety these days") and how anxious they feel right now ("anxiety now"). Participants were asked these questions every week.

Short-run Response to a Stressful Situation

The second outcome of interest in relation to stress is the ability to cope with a stressful situation. Participants were asked to perform a task aimed at inducing stress through a combination of testing cognitive ability/knowledge, time pressure, monetary rewards/losses, and social pressure/shame.[7] Because stress responses decline with habituation to a particular stressful situation (Grissom and Bhatnagar, 2009), different stressful tasks were chosen for the pre- and post-intervention sessions.

In the pre-intervention session (Session 1), the task consisted of a computerized cognitive ability and knowledge test, combining numerical, spatial, and verbal reasoning questions with general knowledge questions. Students were informed that the average student would be expected to be able to answer all questions. Each question was presented on a separate page with a 20 second countdown timer ticking in the top right-hand corner of the page. Students were informed of the requirement of answering 70% of questions correctly in order to participate in a lottery to win one of the two £50 prizes.

In the post-intervention session (Session 6), the task consisted of a cognitive ability and knowledge test that was performed publicly in the laboratory. All participants were asked to stand up in the lab and questions were read aloud by the experimenter, as well as being displayed on a large screen. Immediately after reading a question, the experimenters called upon a randomly selected participant to choose the correct answer to the multiple-choice question. If the given answer was incorrect, participants were informed of this and asked to try another answer. This was repeated until the correct answer had been given. The task consisted of 36 questions. Participants were each endowed with £12 at the beginning of the task, losing £1 for every minute expired on the test. This design was chosen to add social pressure to the task, similar to the Trier Social Stress Test of Kirschbaum et al. (1993), but with the additional pressure of joint incentive payment.

Finally, in the five-month follow-up session (Session 7), participants were asked to take a computerized Stroop test (Jensen and Rohwer, 1966; Stroop, 1935). Participants were sequentially shown names of four different colours (red, blue, yellow, and green) on the screen, written either in congruent or incongruent colour. They were asked to indicate the colour in which the word was written, by clicking on one of four buttons labelled with the colour names. Upon selecting an answer, the next colour name would immediately appear on the screen. This was repeated 96 times. Participants received one penalty point for each second spent on the task, and one penalty point for each mistake made. They were informed that the two participants with the fewest penalty points would earn a bonus of £50 each.

In each session, directly after completing the task, participants were asked to rate how stressful, difficult, and enjoyable they found the task. This gives us a self-reported measure of the acute stress response. We also asked them to predict their relative

[7] See Dickerson and Kemeny (2004) for a synthesis of laboratory research on acute stressors.

performance on the task, before and after having completed it.

In addition, we measured participants' stress response using saliva measurements of cortisol levels, following a standard protocol.[8] Increased cortisol levels can be measured in saliva between 10 to 20 minutes after exposure to a stressor. If there are no further stressors, cortisol levels should return to their initial level within a short period (between 20 to 40 minutes). This is called the "recovery period". If a person experiences stress for a sustained period of time, she could experience what is called "adrenal fatigue", which leads to low levels of cortisol, a weak response to stressors and a longer recovery period (Nicolson, 2008).

Saliva samples were collected three times during the experimental session using Salivette collection devices. The timing of the saliva measurements is outlined in Section 3.3.3. The saliva samples were analysed by a professional laboratory (Salimetrics). These samples were collected for the initial session and for the post-intervention session, but not for the follow-up session.

Summarising the expected effects on chronic stress and stress response, our second hypothesis is as follows:

Hypothesis 2 - Participants in the MBSR programme will be better able to cope with stressful situations. As a consequence, chronic stress should decrease and they should be less affected by and recover faster from stressful events.

3.4.3 Risk and Time Preferences

Because risk and time preferences potentially play an important role in health-related behaviours, we are interested in evaluating how mindfulness affects risk and inter-temporal attitudes directly. We use standard experimental techniques to elicit measures of risk and time preferences.

Risk Attitudes

We use the "Bomb Risk Elicitation Task" (BRET), an intuitive procedure aimed at measuring risk attitudes (Crosetto and Filippin, 2013). Subjects decide how many out of 100 boxes to collect, but are informed that one of the boxes contains a bomb. Earnings increase linearly with the number of boxes collected, but participants receive nothing if the boxes they collect include the one that contains the bomb. Essentially, the task presents 100 lotteries which are described fully in terms of outcomes and probabilities by a single parameter (number of boxes collected). In our experiment earnings per box are £0.05, i.e. participant earnings are equal to the number of boxes collected divided by 20 (unless the bomb is collected). The major advantage of the BRET, compared with other risk elicitation tasks, is that it requires minimal numeracy skills. The task allows estimation of both risk aversion and risk-seeking, and is not affected by the degree of loss aversion.

Crosetto and Filippin (2013) present both a static and a dynamic version of the BRET (note that in both versions of the BRET the location of the bomb is only revealed after participants have made their final decision on the total number of boxes

[8] http://salimetrics.com/collection-system/adult-oral-swab.

they would like to collect). We implement a static version, with participants using a slider to choose how many boxes to collect. In contrast to the dynamic version, in which boxes are collected as time passes and subjects need to decide when to stop collecting boxes, our setup does not introduce any role for time preferences in the decision of how many boxes to collect. Subjects can also revise their decision upward and downward until they are satisfied with their choice. The number of boxes collected is used as the measure of risk aversion. The more risk averse the subject is, the fewer boxes she will collect.

In addition, we construct a non-linear measure of risk aversion, using the approximation of Crosetto and Filippin (2013). Assuming a classic power utility function, the coefficient of relative risk aversion (RRA) can be approximated as $1 - \frac{n}{100 - n}$, where n is the number of boxes collected.

How should we expect mindfulness to affect risk-taking behaviour? There is a theory that risk-taking is linked to executive function. For example, there is evidence that risk-taking observed during adolescence may be due to insufficient prefrontal executive function compared to a more rapidly developing subcortical motivation system (Romer et al., 2011). Thus, if mindfulness training improves executive function, then we would expect it to decrease risk-taking.

Impulsiveness and Time Preferences

We measure impulsiveness and time preferences using both self-reported and incentivised measures. In order to measure self-reported impulsiveness, we use the Barratt Impulsiveness Scale (Patton et al., 1995). This is a widely used measure of impulsiveness, including 30 questions assessing various impulsiveness traits (such as self-control, perseverance, and attention). Each item is reported on a four-point scale, with the total score ranging from 30 (low impulsivity) to 120 (high impulsivity).

We also elicit time preferences using an incentivised experiment. Frederick et al. (2002) review various standard methods used to elicit time preferences. This typically involves asking subjects to choose between various monetary amounts in two different time periods. We are interested both in eliciting subjects' discount rates and in testing whether their preferences are time-consistent. A simple way of determining time consistency is to offer individuals the choice between smaller amounts of money in the present and larger amounts in the future (i.e. today versus in one week), and then also offer them the identical choice between these rewards shifted further into the future (i.e. four months versus four months and one week). We follow the literature in asking subjects to make such choices for various different monetary rewards. If a subject chooses the smaller reward in the first scenario, but the larger one in the second (so-called static preference reversal), this reveals the subject's present bias. Tables A.3.2 and A.3.3 in Appendix 3D display the choice scenarios for Sessions 1, 6 and 7. Participants were informed in each session that one of their decisions would be randomly selected and implemented at the end of the session. While in Session 1 the monetary rewards were small (ranging from £2 to £4) and everyone received the selected payments, in Sessions 6 and 7 the rewards were higher (ranging from £30 to £35), but only two randomly selected participants in each session received the payments associated with their decision.

Opting for future payment creates additional uncertainty and requires subjects to trust the experimenter to pay in the future, introducing variables other than time preference. To keep transaction costs to a minimum, we chose to either provide future payments during pre-scheduled lab-sessions, or administer payments through a voucher card, which could be loaded remotely, without the subject having to come to the laboratory. This procedure, combined with the fact that the experimenters are known to use the BLUE lab regularly, should serve to minimize potential trust issues in our participants.

We construct two summary measures of time preferences using these incentivised experiments. First, we count the number of times the participant preferred to receive the money on the day of the session rather than later. We call this variable *impatience*. Second, we construct an indicator of whether the participant exhibits time-inconsistent preferences (*present bias*), preferring to receive a smaller amount of money today over a larger sum at a later date, while preferring the greater and later payment when offered a similar choice between payments on two later dates. We call this binary variable *present bias*.[9]

Because mindfulness has been shown to increase executive function, we hypothesize that greater self-control could lead the treatment group to become less present-biased than the control group. Note that, since the core exercises associated with mindfulness involve focusing the mind on the *present,* it is not necessarily obvious that this will be the case. However, there is little evidence pointing in the direction of this opposite effect. Our experiment is, however, the first to consider the effect of mindfulness on a standard measure of present bias.

Our third hypothesis regarding risk and time attitudes can be summarised as follows:

Hypothesis 3 - The MBSR programme will reduce risk-taking, increase patience and reduce present bias.

3.4.4 Health-Related Behaviours

The final set of outcomes of interest are health-related behaviours. Since these are difficult to measure accurately, we propose a range of measures, including both self-reported and incentivised ones.

Self-reported Measures

We collect self-reported information on smoking, eating, alcohol consumption and sleeping habits of our subjects, as well as on their physical activities and overall health. Most of these questions are included in the weekly survey. The survey also includes questions related to "emotional" or "comfort eating", based on the Compulsive Eating Scale (Kagan and Squires, 1983).

We collect a number of measures related to eating and healthy eating in particular. First, we construct a summary measure of unhealthy food consumption, counting the number of unhealthy items participants report having consumed the previous day (from

[9]There are three cases of inconsistent choices (i.e. people switching more than once between earlier and later dates), which we exclude from our analysis.

a list we provided; see Appendix 3C for details). Second, we focus on two measures of eating behaviour based on survey questions directly related to emotional eating. The first question asks how often participants feel "out of control" when eating; the other asks participants how often they eat too much because they are "upset, nervous or stressed".

Next, we have measures (based on self-reports) of the frequency of smoking and alcohol consumption, as well as the average number of hours slept. In Session 1, the respondents were asked generally about their smoking and drinking habits, while, in the other sessions, the questions referred to the previous week. The detailed weekly questions can be found in Appendix 3C.

Incentivised Measure

We also collected a measure of preferences for "healthy foods", using a revealed preference approach. Participants were asked to make a real choice between a high-calorie and a low-calorie option. Each option is a combination of a snack and a drink. Participants were first asked to choose sequentially among three high-calorie snacks, three high-calorie drinks, three low-calorie snacks and three low-calorie drinks. We then constructed a low-calorie option by combining their preferred low-calorie snack with their preferred low-calorie drink, and a high-calorie option by combining their preferred high-calorie drink and high-calorie snack. Participants were then endowed with £4 and asked to pick between their preferred high- and low-calorie options at different prices.[10] The options are listed in Table A.3.4 of Appendix 3D. The price of the chosen item would be deducted from the £4 endowment.

We construct a measure of preference for the low-calorie option, which corresponds to the number of times participants choose that option rather than the high-calorie option.

Because of the expected effects of mindfulness on stress, risk and inter-temporal attitudes, the hypothesis regarding health-related-behaviours follows naturally:

Hypothesis 4 - Participants in the MBSR programme will engage more in health-promoting behaviours (such as healthy eating and sleep) and less in health-harming behaviours (such as smoking, unhealthy eating and drinking alcohol).

3.5 Baseline Characteristics

In addition to the outcome variables described above, we also collected background on socio-economic characteristics in the initial session. We use these baseline characteristics to check for balance in randomisation and, later on, for evaluating the implications of attrition.

Table 3.2 presents summary statistics for our sample of participants at baseline to evaluate balance across treatment and control samples. In each panel, we report summary statistics (for the pooled sample in Column (1), the treatment sample in Column (2), and the control sample in Column (3)). We test whether the difference is statistically significant in Column (4).

[10]We also separately asked participants to make decisions involving receiving the snack and drink immediately, but paying later. Unfortunately, these measures cannot be used in the analysis due to an

Table 3.2: Baseline characteristics

Baseline characteristics							
	[1]		[2]		[3]		[4]
	Total		**Treatment**		**Control**		**Diff**
Variables	Mean	SD	Mean	SD	Mean	SD	Mean
Panel A: Individual Characteristics							
Age	24.36	3.61	23.76	1.92	24.92	4.60	-1.16*
Female	0.65	0.48	0.69	0.47	0.61	0.49	0.08
White	0.65	0.48	0.66	0.48	0.64	0.48	0.02
Weight (kg)	63.81	10.16	64.09	10.57	63.56	9.83	0.53
Body mass index (BMI)	21.83	2.59	22.25	2.73	21.44	2.41	0.81
Undergraduate	0.87	0.34	0.90	0.31	0.85	0.36	0.05
Panel B: Stress and Wellbeing							
Perceived stress score (scale: 0-40)	17.78	6.00	18.49	5.81	17.11	6.14	1.38
Anxious these days (scale: 1-11)	6.76	2.42	7.10	2.43	6.43	2.39	0.67
Anxious now (scale: 1-11)	5.50	2.43	6.01	2.45	5.03	2.33	0.98
Life satisfaction nowadays (scale: 1-11)	8.02	1.47	8.01	1.32	8.03	1.60	-0.02
Happiness these days (scale: 1-11)	7.86	1.62	7.78	1.60	7.93	1.65	-0.15
Happiness now (scale: 1-11)	7.40	1.61	7.46	1.44	7.35	1.77	0.11
Things worthwhile (scale: 1-11)	8.22	1.61	8.00	1.70	8.42	1.51	-0.42
Panel C: Behavioural measures							
Present bias (0/1)	0.08	0.27	0.07	0.26	0.08	0.28	-0.01
BIS total score (30 to 120)	64.34	9.47	65.01	10.25	63.71	8.70	1.3
# boxes collected (BRET)	45.65	20.19	48.01	22.36	43.44	17.81	4.57
Impatience (0 to 10)	0.48	1.10	0.42	1.03	0.54	1.16	-0.12
Panel D: Health related behaviours							
Unhealthy food items eaten yesterday	3.94	3.33	3.72	3.06	4.14	3.57	-0.42
Avoid fat	0.48	0.50	0.51	0.50	0.46	0.50	0.05
Eat high fibre food	0.37	0.49	0.42	0.50	0.33	0.47	0.09
Eat at regular times (1-always to 4-never)	2.10	0.82	2.13	0.81	2.07	0.83	0.06
Eat high-calorie snack while studying (0-no, 1-yes)	0.53	0.50	0.52	0.50	0.53	0.50	-0.01
Eat more than usual while preparing for exam (0-no, 1-yes)	0.45	0.50	0.46	0.50	0.43	0.50	0.03
Out of control with food (0-never to 4-always)	1.22	0.93	1.34	0.93	1.10	0.92	0.24
Eat because upset, nervous (0-never to 4-always)	1.39	1.12	1.49	1.15	1.29	1.09	0.20
Eat because bored, lonely (0-never to 4-always)	1.40	1.10	1.34	1.14	1.46	1.07	-0.12
Eat much too fast (0-never to 4-always)	0.89	1.03	1.00	1.03	0.79	1.03	0.21
Average hours of sleep/day	7.62	0.99	7.60	1.03	7.63	0.96	-0.03
Suffer from a health problem leading to a visit to doctor in past 4 weeks	0.15	0.36	0.16	0.37	0.14	0.35	0.02
Frequency of alcohol consumption (1-almost every day to 5-never)	3.40	0.84	3.48	0.73	3.32	0.93	0.16
Smoking (1-none to 4-(10-20) cigarettes per day)	1.29	0.66	1.30	0.65	1.28	0.68	0.02
Observations	139		67		72		

Notes: $***p < 0.01, **p < 0.05, *p < 0.1$.

Panel A presents basic individual characteristics that will be used in the analysis as control variables. Panels B, C and D of Table 3.2 present summary statistics for the main outcome variables at baseline. We start with self-reports of chronic stress, as well as subjective and emotional well-being. Anxiety appears to be a common problem for our sample of student subjects. Providing a more in-depth measure of stress, we also report participants' Perceived Stress Scale (PSS) score. This is based on 10 questions about the frequency of certain thoughts and feelings associated with stress, each answered on a scale from "Never" to "Very Often" (coded as 0-4, with 0 representing "Never" and 4 representing "Very Often"). Thus the highest possible PSS score would be 40. In our baseline sample, the average PSS score is 17.78. This is comparable to PSS scores in similar samples in previous studies. For example, based on samples of university students in the US, Von Ah et al. (2004) report a mean value of 19.56 and Roberti et al. (2006) report a mean of 18.3 on the ten-item PSS. Our average score is lower than the mean scores of 23.04 and 22.4 reported respectively by Krusche et al. (2013) and Morledge et al. (2013) for samples of individuals who self-selected into an online mindfulness course.

Panel C reports the behavioural measures at baseline, both those based on self-reports (impulsivity and patience), and those based on incentivised revealed-preference measures (present bias and risk aversion). Only a small proportion of participants is present-biased (8%) at baseline, which is surprisingly low. The average value of impulsivity is in the lower half of the impulsivity scale and is comparable to other recent studies involving students in the UK (see Caswell et al. (2016)). In terms of patience, the majority of the students prefer all of the later options to the earlier ones, thus the indicator of impatience at Session 1 is on average very low. Note, that impatience becomes more common in Sessions 6 and 7, when the monetary rewards are higher. Finally, the mean value of our measure of risk aversion, corresponding to the number of boxes collected in the BRET, is 46, which is very close to the mean observed in Crosetto and Filippin (2013).

Descriptive statistics on health-related behaviours are presented in Panel D of Table 3.2.[11] Taken together, the results suggest the prevalence of a high degree of anxiety and health-compromising lifestyles among the student population participating in our experiment.

3.6 Evaluation of the MBSR Intervention

3.6.1 Empirical Strategy

We estimate the reduced form effect of participating in the MBSR intervention on the outcome variables described above using the following differences-in-differences specifications. Specification (3.1) is used for outcome measures taken only at the baseline and Sessions 6 and 7, while specification (3.2) is used for outcome measures that are measured at each session.

error in programming.

[11] While the table includes all the variables, in the rest of the main text we focus on a representative set of indicators of health-related behaviours. Results relating to the remaining variables are presented in Appendix 3H.

$$
\begin{aligned}
Y_{it} &= \alpha + \beta\ MBSR_i + \gamma_1\ MBSR_i \times Session6_t + \gamma_2\ MBSR_i \times Session7_t + \delta_t week_t \\
&\quad + \phi X_i + \eta_i + \epsilon_{it} \qquad (3.1)
\end{aligned}
$$

$$
\begin{aligned}
Y_{it} &= \alpha + \beta\ MBSR_i + \gamma_1\ MBSR_i \times Session2_t + \gamma_2\ MBSR_i \times Session3_t \\
&\quad + \gamma_3\ MBSR_i \times Session4_t + \gamma_4\ MBSR_i \times Session5_t + \gamma_5\ MBSR_i \times Session6_t \\
&\quad + \gamma_6\ MBSR_i \times Session7_t + \delta_t week_t + \phi X_i + \eta_{it} + \epsilon_{it} \qquad (3.2)
\end{aligned}
$$

where Y_{it} is an outcome variable measured for individual i in week t. $MBSR$ is a dummy variable equal to 1 for individuals in the MBSR Treatment. The $Session$ variables are dummy variables that equal 1 if the outcome is measured in that particular session, where $Session7$ corresponds to the five-month follow up session. X_i is a vector of individual characteristics such as gender, age, ethnicity, a dummy for being an undergraduate student and Body Mass Index at baseline (i.e. week 1). η_{it} is an individual specific random effect and ϵ_{it} is a white noise error term. We check robustness of our results to the exclusion of the control variables (X_i). We also perform the Hausman test, which tests the null hypothesis of orthogonality (no correlation between the regressors and the individual random effects η_{it}). The test results do not reject the null, implying that our parameter estimates are consistent when estimated using the random effects specification.

Note that attrition can potentially play an important role in the analysis of all outcome variables. Similarly, we cannot be sure that all students have fully complied with the protocol to which they were assigned. Therefore our estimates will always be intention-to-treat estimates. We first discuss attrition and compliance, and then move to the other outcome variables.

3.6.2 Attrition and Compliance

We begin by testing *Hypothesis 1.* Both interventions require some degree of commitment from the participants. Our data allow us to study the determinants of continued participation in the study and, in particular, engagement with the mindfulness protocol. One would expect that certain behavioural characteristics such as impulsiveness, impatience and present bias may be correlated with the likelihood of attrition. Because we have collected a large set of variables at baseline, we are able to test this hypothesis directly.

We start with the information on attendance. The number of subjects in both the treatment and control groups declined over time due to attrition. In Session 6, 17 of the 67 original subjects in the treatment group (representing 25%) and 11 of the original 72 subjects in the control group (representing 15.3%) did not attend the experimental session. The non-attendance rate in Session 7 was 41.8% in the treatment group and 29.2% in the control group. The differences between the attrition rates in the control and treatment groups are not statistically significant ($p = 0.139$ at Session 6 and $p = 0.12$ at Session 7)[12], however, concern about bias in the estimation results due to attrition still seems justified.

First, we analyse the determinants of attrition using attrition probits (Fitzgerald et al., 1998). Attrition probits consist of estimates of binary-choice models for the determinants of attrition in later periods as a function of base year characteristics. We estimate separate attrition probit models for the treatment and control groups. We include a rich set of baseline characteristics in the models, but have to exclude some variables to avoid strong multicollinearity (anxiety now, happiness now) and perfect prediction (present bias). We will come back to the latter, since it is a variable we thought could be correlated with engagement. The dependent variable is a binary indicator of being present at Session 6 or 7.

The results presented in Table 3.3 show that, although there are some significant coefficients in the attrition probit models, there is no systematic relation between the baseline characteristics and attrition. The personal characteristics that are significantly related to attrition are those characteristics for which we control in our estimations. We also see that "anxiety these days" significantly reduces the probability of remaining in the sample within the treatment group. If the MBSR program is more effective among the subjects who report anxiety, then this selectivity can lead to underestimation of the beneficial effect of the program on anxiety. Six individuals in the control group, coded as present-biased in Session 1, have to be excluded due to perfect prediction of non-attrition by present bias. To gauge the effects of present bias on attrition, we tested for a simple mean difference in present bias as measured in Session 1 between the original sample and the sample present in Sessions 6 and 7. We found no significant differences.

As a second test of attrition, we look at whether the treatment and control samples that are present in Sessions 6 and 7 are still comparable in terms of their baseline characteristics. This check can reveal whether there is asymmetric attrition between the treatment and the control groups (on observable characteristics). We test for equality of the same set of baseline characteristics that we used in the attrition probit models. The results are presented in Table 3.4. There are statistically significant differences

[12]Results from the simple attrition probit without controls are shown in Table A.3.5 in the Appendix

Table 3.3: Attrition probits (marginal effects on non-attrition)

	Treatment Session 6		Control Session 6		Treatment Session 7		Control Session 7	
	Marginal effect	SE	Marginal effect	SE	Marginal effect	SE	Marginal effect	SE
Personal characteristics								
Age	-0.039*	0.022	0.007*	0.005	-0.062*	0.034	0.048***	0.015
Female	0.082	0.117	-0.049	0.043	0.518***	0.161	-0.130	0.102
White	-0.058	0.109	0.409***	0.143	-0.024	0.172	0.479***	0.157
BMI	0.022	0.016	-0.011	0.010	0.058**	0.026	-0.002	0.020
Stress and subjective well-being								
PSS	0.008	0.011	-0.001	0.005	0.010	0.018	-0.015	0.011
Anxious these days	-0.060*	0.033	0.022*	0.011	-0.128***	0.048	0.020	0.030
Anxious now	-0.018	0.027	-0.034**	0.015	0.016	0.038	-0.008	0.029
Life satisfaction nowadays	0.070	0.054	-0.038**	0.024	-0.033	0.073	-0.062	0.044
Happiness these days	0.006	0.044	0.019	0.020	0.033	0.058	-0.032	0.043
Behavioural measures								
Risk aversion (BRET)	0.003	0.002	0.002*	0.001	0.003	0.004	0.004	0.003
Impulsivity (BIS)	-0.003	0.005	-0.003	0.003	0.000	0.007	0.006	0.006
Impatience	-0.049	0.046	0.076**	0.049	0.018	0.084	0.027	0.074
Present bias	-0.264	0.403			-0.213	0.385	0.161	0.084
Health-related behaviours								
Unhealthy food items eaten yesterday	0.018	0.016	-0.001	0.007	0.044*	0.024	0.006	0.014
Out of control with food	0.086	0.069	0.002	0.034	-0.041	0.101	0.136*	0.074
Eat because upset, nervous	-0.035	0.062	0.033	0.027	-0.013	0.097	-0.031	0.060
Average hours of sleep/day	0.023	0.044	-0.056***	0.028	0.068	0.072	-0.053	0.049
Smoking	-0.146**	0.069	-0.096***	0.044	-0.017	0.124	-0.274***	0.082
Frequency of alcohol consumption	0.023	0.060	-0.001	0.028	0.257***	0.110	-0.006	0.064
Low calorie option chosen	0.017	0.012	-0.001	0.007	-0.014	0.020	0.015	0.014
No. of individuals	67		65		67		71	

, **, * indicate significance levels at 10%, 5% and 1% respectively*

in age, gender and BMI between the treated and control individuals at the baseline, but these are relatively small. These are also characteristics that we control for in the empirical specifications. More importantly, we do not see significant differences in terms of risk attitudes, patience or impulsiveness, our key outcome variables of interest. One variable for which we observe significant differences is stress-related eating, which is significantly more prevalent within the treatment group. We do not see evidence of significant differences for the other behavioural measures and health-related behaviours. Thus, Hypothesis 1 is not supported by the data. We do not see that engagement with the protocols is correlated with psychological traits or behavioural measures such as impulsiveness and impatience. This supports that mindfulness can possibly be an efficient instrument to target subjects with various behavioural characteristics. These findings also reduce concerns about the analysis of behavioural measures and health-related behaviours suffering from bias due to attrition. We provide a further check of the importance of attrition in Appendix 3F, where we re-estimate the results on PSS and anxiety measures using the non-attriting sub-sample.

Table 3.4: Comparison of Baseline Means of the Non-attrited Subsamples of Treatment and Control Groups

	Present at Session 6 treatment-control		Present at Session 7 treatment-control	
	Diff.	SE	Diff.	SE
Personal characteristics				
Age	-1.478**	0.732	-1.991**	0.874
Female	0.097	0.09	0.252**	0.097
White	-0.032	0.091	-0.045	0.101
BMI	1.269**	0.498	-1.356**	0.579
Stress and subjective well-being				
PSS	0.992	1.165	1.148	1.35
Anxious these days	0.192	0.456	-0.062	0.533
Anxious now	0.724	0.458	0.297	0.525
Life satisfaction nowadays	0.233	0.273	0.299	0.312
Happiness these days	-0.034	0.319	0.056	0.370
Behavioural measures				
Impulsivity (BIS)	1.412	1.881	1.090	2.033
Risk aversion (BRET)	3.081	3.831	2.428	4.004
Impatience	-0.343	0.213	-0.268	0.241
Present bias	-0.058	0.049	-0.047	0.057
Health-related behaviours				
Unhealthy food items	-0.258	0.673	-0.208	0.762
Out of control with food	0.358**	0.180	0.351*	0.203
Eat because upset, nervous	0.221	0.217	0.407*	0.236
Average hours of sleep/day	0.036	0.191	0.069	0.211
Smoking	0.007	0.108	0.106	0.123
Frequency of alcohol consumption	0.196	0.164	0.237	0.186
Low calorie option chosen	1.025	0.654	0.520	0.719

, **, * indicate significance levels at 10%, 5% and 1% respectively*

The next variables of interest are the degree of engagement and compliance with the interventions. We have designed three strategies to measure these. First, we asked participants to report every week to what extent they engaged in various activities to relax, such as meeting with friends, going to the theatre, etc. (see Appendix 3C for full

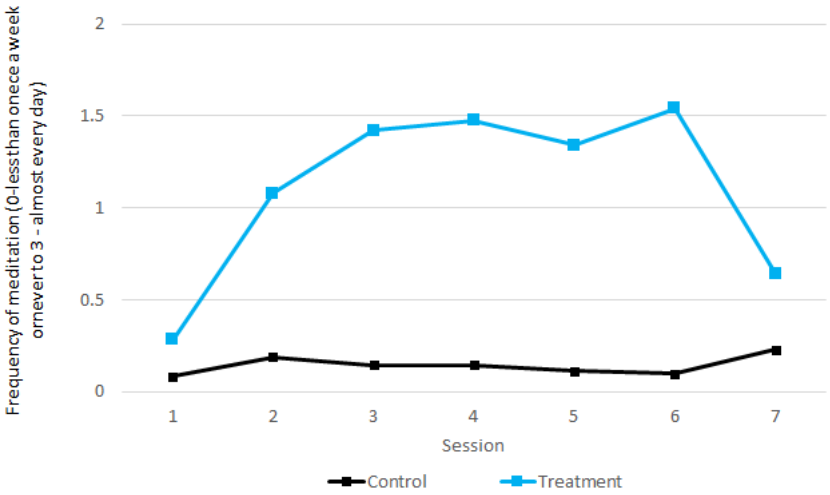


Figure 3.1: Average frequency of meditation (0-less than once a week or never to 3-almost every day), by session and treatment

questionnaire). Meditating is one of the activities they were asked about. Figure 3.1 shows the average report on the extent to which participants meditate, with 0 being never or less than once a week and 3 being almost every day. We report the difference-in-difference analysis in Appendix 3E (Table A.3.6). We find a significantly positive treatment effect on the frequency of mediation during Sessions 2-6. The effect remains positive but is no longer statistically significant at the follow-up session five months later.

Second, we asked participants to describe the contents of the weekly mindfulness lesson (in at least 100 words). We also asked the participants in the control intervention to describe the documentary episode they were asked to watch. We constructed a dummy variable indicating whether the text was indeed an accurate description of the lesson/episode.[13] Participants complied to a large extent with the interventions. Based on this binary indicator of compliance, more than 90% of the subjects in the control group followed the control intervention every week. Compliance also exceeded 90% in the treatment group every week, except for the first week, when 82% of the subjects followed the MBSR program (this statistic is based on the survey during Session 2). Of course it is worth noting that the number of participants falls over time, and attriting participants are unlikely to still be engaged with the intervention.

Finally, the last strategy to check for engagement and compliance of the treatment group with the mindfulness intervention involves using data from the website, which tracked participants' activities during the online sessions. The website tracks when participants logged in and completed the various stages of the intervention. By Session 6, 36% of the non-attrited individuals in the treatment group had fully completed the online mindfulness course (while 72% had reached at least Week 4 of the course by this point). By Session 7, the completion rate increased to 59%. The estimated marginal effects on the probability of completing the course (based on probit models)

[13]We also coded the response as a zero if the description was generic and did not demonstrate that they engaged with the intervention or if they mention not having done the activity at all.

are reported in Table 3.5. According to these results, indicators of stress, behavioural preferences and health behaviours are not systematically related to the probability of completing the mindfulness course.

Table 3.5: Probability of completing the course

	Marginal effect	SE	Marginal effect	SE
Personal characteristics				
Age	0.086*	0.046	0.046	0.040
Female	0.244	0.148	0.315	0.209
White	-0.051	0.140	0.047	0.177
BMI	0.035	0.026	0.027	0.030
Stress and subjective well-being				
PSS	0.007	0.014	0.014	0.014
Anxious these days	-0.068*	0.039	-0.104**	0.043
Anxious now	0.036	0.035	0.070*	0.041
Life satisfaction nowadays	0.005	0.060	-0.003	0.075
Happiness these days	0.007	0.053	0.016	0.062
Behavioural measures				
Impulsivity (BIS)	-0.001	0.006	0.008	0.007
Risk aversion (BRET)	0.000	0.003	-0.001	0.003
Impatience	-0.034	0.090	-0.074	0.082
Present bias	-0.073	0.346	-0.112	0.406
Health-related behaviours				
Unhealthy food items			0.010	0.022
Out of control with food			-0.147	0.124
Eat because upset, nervous			0.048	0.095
Average hours of sleep/day			-0.133**	0.067
Smoking			0.086	0.117
Frequency of alcohol consumption			0.237**	0.121
Low calorie option chosen			-0.016	0.019
No. of individuals	65		65	

, **, * indicate significance levels at 10%, 5% and 1% respectively*

3.6.3 Effects on Chronic Stress and Stress Response

We now turn to *Hypothesis 2,* which relates to the impact of the MBSR intervention on chronic stress and on the response to a stressful situation. These outcomes are the primary targets of the mindfulness programme.

We use three different sources to construct a measure of chronic stress. The first is based on the total score of the Perceived Stress Scale (measured in the initial session and in the two post-intervention sessions). The second and third are based on responses to weekly questions about how anxious the participants feel "now" and "these days", both on a scale from $1 - 11$. Table 3.6 reports the treatment effects of the intervention on these three measures from a difference-in-differences estimator (MBSR & Session 6, and MBSR & Session 7 show the post-intervention estimates of the treatment effect).

Table 3.6: The Impact of MBSR on Perceived Stress Score (PSS) and Anxiety Measures

	[1]		[2]		[3]	
	PSS		Anxiety Now		Anxiety These Days	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
MBSR	1.463	0.981	1.256***	0.407	0.878**	0.408
Session 2	.	.	0.531	0.346	0.273	0.336
Session 3	.	.	0.918***	0.327	0.302	0.275
Session 4	.	.	0.227	0.371	-0.095	0.304
Session 5	.	.	0.344	0.351	-0.141	0.306
Session 6	0.999*	0.511	0.279	0.344	-0.131	0.245
Session 7	2.205***	0.855	0.747*	0.387	0.344	0.355
MBSR & Session 2	.	.	-0.857*	0.505	-0.454	0.450
MBSR & Session 3	.	.	-1.163**	0.520	-0.882*	0.458
MBSR & Session 4	.	.	-0.402	0.529	-0.387	0.445
MBSR & Session 5	.	.	-0.360	0.511	-0.296	0.495
MBSR & Session 6	-1.809*	0.926	-0.068	0.542	-0.069	0.454
MBSR & Session 7	-2.464*	1.320	-1.095*	0.650	-0.765	0.582
Intercept	17.363**	7.063	6.982***	2.009	8.575***	2.116
Individual random effects	Yes		Yes		Yes	
Control variables	Yes		Yes		Yes	
No. of individuals	138		138		138	

Notes: Robust standard errors; *** $p < 0.01$,** $p < 0.05$, * $p < 0.1$.

Results show that the MBSR intervention leads to a decrease in participants' PSS scores both in the week immediately following completion of the course (session 6, p=0.051) and at the 5 month follow-up (session 7, p=0.062), though this is statistically significant only at the 10% level. Compared to the baseline PSS scores, this amounts to a decrease of around 10 percent. The estimated treatment effect resulted from both decreasing levels of stress among the treatment group and increasing levels of stress among the control group. The effect is comparable to the effect found by Morledge et al. (2013) after 8 and 12 weeks of an internet-based mindfulness program, but smaller than the effect found by Krusche et al. (2013). Krusche et al. (2013) estimate that the online mindfulness course reduces the average PSS score by around 8 points and by a further 1.5 points a month later; however, these estimates are based on a sample of self-selected individuals, without the inclusion of a control group in their analysis. The MBSR intervention also appears to reduce reported anxiety throughout session 2 to 7, but these estimates are mostly not statistically significant. The results indicate that the treatment is more effective in reducing the current level of anxiety (anxiety "now") than the general level of anxiety (anxiety "these days"). We do not find any significant treatment effects on other measures of subjective well-being, including measures of life satisfaction, happiness and "considering things worthwhile".[14]

[14]We conducted a series of specification checks to investigate further the the impact of the MBSR intervention on PSS score and anxiety. First, to check for the importance of attrition, we re-estimated the treatment effects using the sub-sample of individuals who were present at Session 6 or 7. Although the precision of the estimated treatment effects declines, the main conclusions remain robust. These results are reported in Table A.3.7 in Appendix 3F. Next, we estimated the effect of MBSR on the sum of the two indicators of academic stress (worries about grades in the current semester and in the future). We find no significant treatment effects. Finally, while we see that stressful events (measured by the

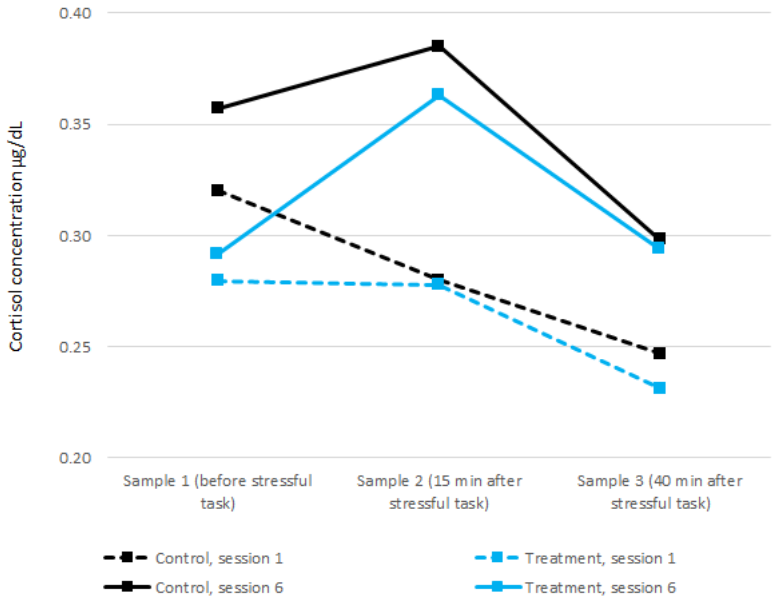


Figure 3.2: Salivary cortisol concentration averages by session and by treatment

Next, we examine how the intervention affected the response to a stressful situation. Appendix 3G (Table A.3.8) summarises how stressful, not enjoyable and difficult the participants found each task. Based on these indicators, while the stressfulness of all three tasks was rated around 6-7 on average on a 10-point scale, the computerised ability and knowledge test was considered on average less enjoyable and more difficult than the post-intervention tasks.

Considering the salivary cortisol measurements, we do not find evidence that the MBSR intervention significantly affected the physiological measures of stress levels and stress responses. The average levels of the three cortisol measurements by session and by treatment are displayed in Figure 3.2. These cortisol levels are within the normal ranges of cortisol concentration.

Summarising, the evidence based on self-reported measures is supportive of *Hypothesis 2,* but the evidence based on physiological measurements is inconclusive.

3.6.4 Effects on Risk Attitudes, Time Preferences and Impulsiveness

We now turn to investigating the effects of the intervention on risk attitudes and time preferences of the program participants (*Hypothesis 3*). The impulsivity and risk measures are identical across Sessions 1, 6 and 7. The time preferences measure differs slightly across the three sessions. The exact choice scenarios are described in the Appendix.

Table 3.7 presents difference-in-differences regression results on the impact of our MBSR intervention on risk preferences. Results in Column [1] show that the number

modified Adolescent Perceived Events Scale, APES) increase the PSS score, we do not see evidence that the PSS score of the treatment group responds less to such stressful events.

of boxes collected decreases in the treatment group relative to the control group (which is indicative of a decrease in risk-taking), but only in Session 7, and the effect is not statistically significant at conventional levels. Using the approximated coefficient of relative risk aversion (Column [2]), the estimated treatment effects are more robust and statistically significant in Session 7. As the coefficient "MBSR" indicates, subjects in the treatment group were initially significantly less risk-averse than those in the control group, as measured by the coefficient of relative risk aversion. The gap between the two groups appears to be eliminated by the MBSR treatment. This effect is in line with *Hypothesis 3*.

Table 3.7: The Impact of MBSR on Risk Aversion

	[1]		[2]	
	Risk aversion (BRET)		Risk aversion (BRET)	
	Number of boxes		RRA coefficient	
	Coeff.	SE	Coeff.	SE
MBSR	5.04	3.453	-0.733**	0.362
Session 6	-4.081	2.693	0.138	0.111
Session 7	2.844	2.87	-0.047	0.096
MBSR & Session 6	0.407	4.378	0.431	0.414
MBSR & Session 7	-4.522	3.987	0.779**	0.398
Intercept	63.203***	13.99	-1.560**	0.757
Individual RE	Yes		Yes	
Control variables	Yes		Yes	
No. of individuals	138		138	

, **, * indicate significance levels at 10%, 5% and 1% respectively*

Because the measure of time preferences was not identical across sessions, we conduct a simple difference analysis on patience and present-bias between treatment and control groups for Sessions 1, 6 and 7 separately and report the results in Tables 3.8 and 3.9 respectively. Results in Table 3.8 show that participants in the treatment group became more patient after the intervention, but the effects are not statistically significant. The point estimates and standard errors are quite large, however, so there is a possible issue of statistical power. In Table 3.9, we show that participants have a similar propensity of being present-biased in Session 1, but the treatment group appears less present-biased immediately after the intervention, although the effects are again not statistically significant. We find no significant difference in Session 7 either. However, it is useful to point out that the baseline measure of present-bias was very low (with only 8% of the participants categorised as present-biased). Overall, we take our results as somewhat indicative that patience may have increased and the propensity to be present-biased decreased, but these results are not statistically significant at conventional levels. Thus, we do not find strong support for *Hypothesis 3*.

Table 3.8: The Impact of MBSR on Impatience

	Impatience Session 1		Impatience Session 6		Impatience Session 7	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
MBSR	-0.018	0.178	-0.179	0.411	-0.12	0.486
Intercept	3.002**	1.294	5.184***	1.891	-3.571	2.501
Individual RE	No		No		No	
Control variables	Yes		Yes		Yes	
No. of individuals	136		112		80	

, **, * indicate significance levels at 10%, 5% and 1% respectively*

Table 3.9: The Impact of MBSR on Present Bias

	Present bias Session 1		Present bias Session 6		Present bias Session 7	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
MBSR	0.016	0.033	-0.113	0.101	0.026	0.121
Individual RE	No		No		No	
Control variables	Yes		Yes		Yes	
No. of individuals	136		112		80	

Notes: Probit estimates (marginal effects reported).
, **, * indicate significance levels at 10%, 5% and 1% respectively*

Finally, we look at how mindfulness affects the scores on the Barratt Impulsiveness Scale (BIS). We conduct a difference-in-differences analysis and report the results in Table 3.10. Here, we document a significant effect in Session 6, but we find that participants in the treatment group increased their average score on the BIS relative to the participants in the control group; that is, if anything, mindfulness appears to have increased impulsiveness rather than decreased it. These survey-based results go in the opposite direction to those we found for patience and present-bias using revealed preference methods.[15]

Table 3.10: The Impact of MBSR on Impulsivity

	Coeff.	SE
MBSR	0.809	1.589
Session 6	-0.11	0.735
Session 7	0.217	0.756
MBSR Session 6	2.229*	1.174
MBSR Session 7	0.983	1.298
Intercept	59.199***	9.135
Individual RE	Yes	
Control variables	Yes	
No. of individuals	138	

Notes: *, **, *** indicate significance levels at 10%, 5% and 1% respectively.

3.6.5 Effects on Health-related Behaviours

We finally investigate the impact of the MBSR intervention on health-related behaviours (*Hypothesis 4*). First, we look at our three main measures of eating behaviour: the number of unhealthy food items eaten the day before, and two measures related to emotional eating (how often participants feel out of control while eating and how often they eat too much because of being upset or nervous). The results are reported in Table 3.11.[16]

Overall, participants in the MBSR programme appear to adopt somewhat healthier eating habits. For the second measure of emotional eating, the treatment effect is quite substantial in Session 7, with an estimated coefficient of -0.417 (the indicator ranges from 0-never to 4-always, with baseline mean of 1.4), which is statistically significant at the 10% level. However, it is important to bear in mind that the attrition analysis indicates non-random attrition related to this particular indicator. Thus, the results might partly be driven by self-selection of participants with emotional eating problems.

[15]This result seems to be driven by the sub-category of questions related to "self-control" and, more specifically, four of the 30 questions that make up the Barratt Impulsiveness Scale. In Session 6, participants in the treatment group are statistically significantly more likely to describe themselves as "I am happy-go-lucky" (Barratt item 4) and significantly less likely to say "I am self-controlled" (item 8), "I am a careful thinker" (item 12), or "I am a steady thinker" (item 20). The effect on "I am happy-go-lucky" (Barratt item 4) persists to Session 7 five months later. Given that these are self-reported measures, it is of course possible that doing the mindfulness course has simply made participants think of themselves differently (and possibly more critically, in terms of self-control).

[16]We document effects on additional variables related to eating behaviour in Appendix 3H.

Next, we look at number of hours of sleep, alcohol consumption and smoking, all self-reported as well. We do not find any consistent pattern here - there is little evidence that these behaviours were affected in one direction or the other (Appendix 3H, Table A.3.9).

Finally, we consider the revealed preference measure of healthy eating based on real choices between a low-calorie and a high-calorie option. We counted the number of times participants preferred the low-calorie option over the high-calorie option, which is a direct indicator of their relative preferences. We find no evidence that participants changed their choices significantly. They all appear to value the low-calorie option less in the later sessions, but this effect is not significantly different across treatment and control groups. (Appendix 3H, Table A.3.10)

Table 3.11: The Impact of MBSR on Eating Habits

	Unhealthy food items yesterday		Out of control with food		Eat because upset, nervous	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
MBSR	-0.267	0.593	0.126	0.156	0.086	0.176
Session 2	-0.245	0.429	-0.274***	0.101	-0.386***	0.139
Session 3	-0.152	0.389	-0.275**	0.111	-0.431***	0.138
Session 4	0.142	0.496	-0.246**	0.123	-0.549***	0.135
Session 5	0.127	0.455	-0.306***	0.115	-0.530***	0.143
Session 6	-0.236	0.492	-0.16	0.108	-0.379***	0.14
Session 7	-0.226	0.55	0.015	0.131	0.076	0.145
MBSR & S2	0.355	0.588	0.167	-0.007	0.178	-0.066
MBSR & S3	0.235	0.501	0.035	0.186	0.025	0.195
MBSR & S4	-0.246	0.618	-0.223	0.185	-0.041	0.191
MBSR & S5	-0.405	0.559	-0.045	0.168	0.057	0.202
MBSR & S6	-0.488	0.647	-0.195	0.162	-0.038	0.191
MBSR & S7	-0.791	0.7	-0.094	0.207	-0.417*	0.228
Intercept	5.910*	3.201	1.458**	0.688	0.628	0.78
Ind. RE	Yes		Yes		Yes	
Control var.	Yes		Yes		Yes	
No. of ind.	138		138		138	

, **, * indicate significance levels at 10%, 5% and 1% respectively*

Summarising, we find indicative evidence for *Hypothesis 4*, that is, eating habits may have improved, particularly stress-related eating; but all other measures of health-related behaviours appear unaffected by the intervention. Considering that we look at several outcome variables of health related behaviours, we would have wanted to see a consistent pattern across all of these (as well as adjusting for multiple hypothesis testing) if we wanted to conclude that there was an effect of the intervention on health-related behaviours.

3.7 Discussion and Conclusions

Lifestyle choices are thought to be factors central to explaining the increasing prevalence of non-communicable diseases around the world. This paper explores the possibility of training the mind to alter fundamental cognitive processes underpinning decision-making, using a series of novel experiments. We conducted a randomised field experiment on 139 participants to investigate the effects of mindfulness training on stress, risk-taking and time preferences, as well as on health-related behaviours. Around half of the participants were assigned to a "Mindfulness Based Stress Reduction" programme while the other half were asked to watch a documentary series called "BBC Ancient Worlds". Both interventions ran for four consecutive weeks, and, to measure their long-term impact on behaviour, we conducted a post-intervention session five months later. Importantly, our participants did not self-select into the programmes.

Analysing the patterns of attrition and compliance, we find no evidence of behavioural characteristics predicting engagement with the mindfulness program. We find that the mindfulness intervention appears to reduce perceived stress, but the evidence based on physiological measures of stress (cortisol) is less conclusive. We show indicative evidence that participants may have become more risk averse, as well as more patient and less present-biased, but these results are not statistically significant. On the other hand, participants in the MBSR treatment score higher on the Barratt Impulsiveness Scale based on survey questions. Finally, we find that participants in the MBSR treatment are less likely to engage in stress-related eating, and the point estimates relating to other variables are also suggestive that eating habits may have improved overall, although again the results are not statistically significant. We fail to find any evidence - even suggestive - of changes in other health-related behaviours such as sleep, alcohol consumption and smoking.

Overall, we conclude that such interventions appear to be effective at targeting people with various behavioural characteristics and reducing "feelings of stress", but the effects on decision-making and health-related behaviours are not entirely clear. Looking at the set of point estimates we have, we cannot reject the hypothesis that mindfulness may have in fact increased patience and risk aversion, reduced present-bias and encouraged healthier eating habits, by a significant magnitude. The effects of mindfulness on such outcomes may be more diffuse, especially as we are considering a population of healthy subjects. We believe that further research is needed to obtain more robust evidence of the effects of such techniques on decision-making.

Bibliography

H. Allcott. Social norms and energy conservation. *Journal of Public Economics*, 95(9): 1082–1095, 2011.

S. Altmann and C. Traxler. Nudges at the dentist. *European Economic Review*, 72:19 – 38, 2014.

J. S. Ancker, A. Edwards, S. Nosal, D. Hauser, E. Mauer, and R. Kaushal. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Medical Informatics and Decision Making*, 17(1):36, 2017.

P. Arcidiacono, V. J. Hotz, and S. Kang. Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16, 2012.

G. Azmat and N. Iriberri. The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(7):435–452, 2010.

R. A. Barkley. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1):65, 1997.

M. E. Beier and P. L. Ackerman. Determinants of health knowledge: an investigation of age, gender, abilities, personality, and interests. *Journal of Personality and Social Psychology*, 84(2):439, 2003.

J. Blanes i Vidal and M. Nossol. Tournaments without prizes: Evidence from personnel records. *Management Science*, 57(10):1721–1736, 2011.

M. Bobba and V. Frisancho. Learning about oneself: The effects of performance feedback on school choice. *IZA Discussion Paper*, 10360, 2016.

P. Bordalo, N. Gennaioli, and A. Shleifer. Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843, 2013.

P. Bordalo, N. Gennaioli, and A. Shleifer. Memory, attention and choice. *Working paper*, 2015.

T. Boyce, A. Dixon, B. Fasolo, and E. Reutskaja. *Choosing a high-quality hospital: The role of nudges, scorecard design and information.* The King's Fund, 2010.

A. Brandon, P. Ferraro, J. List, R. Metcalfe, M. Price, F. Rundhammer, et al. Do the effects of social nudges persist? Theory and evidence from 38 natural field experiments. *NBER Working Paper*, 23277, 2017.

C. N. Brinch, E. Herns, and Z. Jia. Salience and social security benefits. *Journal of Labor Economics*, 35(1):265–297, 2017.

P. Burge, N. Devlin, J. Appleby, C. Rohr, and J. Grant. *London Patient Choice Project Evaluation: A model of patients' choices of hospital from stated and revealed preference choice data.* RAND Corporation, 2005.

S. V. Burks, J. P. Carpenter, L. Goette, and A. Rustichini. Overconfidence and social signalling. *The Review of Economic Studies*, 80(3):949–983, 2013.

K. Caldwell, M. Baime, and R. Wolever. Mindfulness based approaches to obesity and weight loss maintenance. *Journal of Mental Health Counseling*, 34(3):269–282, 2012.

C. Camerer and D. Lovallo. Overconfidence and excess entry: An experimental approach. *The American Economic Review*, 89(1):306–318, 1999.

J. Carmody and R. A. Baer. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. *Journal of Behavioral Medicine*, 31:23–33, 2008.

J. Carmody, R. A. Baer, E. L B Lykins, and N. Olendzki. An empirical study of the mechanisms of mindfulness in a mindfulness-based stress reduction program. *Journal of Clinical Psychology*, 65:613–626, 2009.

A. J. Caswell, M. A. Celio, M. M. J, and D. T. Impulsivity as a multifaceted construct related to excessive drinking among uk students. *Alcohol and Alcoholism*, 51(1): 77–83, 2016.

J. Cawley, J. A. Rizzo, and K. Haas. Occupation-specific absenteeism costs associated with obesity and morbid obesity. *Journal of Occupational and Environmental Medicine*, 49(12):1317–1324, 2007.

R. Chambers, B. C. Y. Lo, and N. B. Allen. The impact of intensive mindfulness training on attentional control, cognitive style, and affect. *Cognitive Therapy and Research*, 32(3):303–322, 2008.

R. Chetty, A. Looney, and K. Kroft. Salience and taxation: Theory and evidence. *The American Economic Review*, 99(4):1145–1177, 2009.

A. Chiesa and A. Serretti. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. *Journal of Alternative and Complementary Medicine*, 15:593–600, 2009.

S. Cohen, T. Kamarck, and R. Mermelstein. A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4):385–396, 1983.

B. E. Compas. Stress and life events during childhood and adolescence. *Clinical Psychology Review*, 7(3):275–302, 1987.

Z. Cooper, S. Gibbons, S. Jones, and A. McGuire. Does hospital competition save lives? Evidence from the English NHS patient choice reforms. *The Economic Journal*, 121 (554):F228–F260, 2011.

S. Cornelisse, V. Van Ast, J. Haushofer, M. Seinstra, and M. Joels. Time-dependent effect of hydrocortisone administration on intertemporal choice. *SSRN*, 2294189, 2013.

P. Crosetto and A. Filippin. The bomb risk elicitation task. *Journal of Risk and Uncertainty*, 47(1):31–65, 2013.

M. F. Dallman, N. Pecoraro, S. F. Akana, et al. Chronic stress and obesity: a new view of comfort food. *Proceedings of the National Academy of Sciences*, 100(20): 11696–11701, 2003.

S. DellaVigna and U. Malmendier. Paying not to go to the gym. *The American Economic Review*, 96(3):694–719, 2006.

A. Diamond and K. Lee. Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333(6045):959–964, 2011.

E. Dias-Ferreira, J. a. C. Sousa, I. Melo, P. Morgado, A. R. Mesquita, J. a. J. Cerqueira, R. M. Costa, and N. Sousa. Chronic Stress Causes Frontostriatal Reorganization and Affects Decision-Making. *Science*, 325(5940):621–625, July 2009.

S. S. Dickerson and M. E. Kemeny. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3):355, 2004.

A. Dixon, R. Robertson, J. Appleby, P. Burge, and N. J. Devlin. *Patient choice: how patients choose and how providers respond*. The King's Fund, 2010.

R. Dizon-Ross. Parents' beliefs and children's education: Experimental evidence from Malawi, 2016.

T. Dohmen, A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550, 2011.

J. S. Downs, G. Loewenstein, and J. Wisdom. Strategies for promoting healthier food choices. *The American Economic Review*, 99(2):159–164, 2009.

R. Dur. Salience of law enforcement: A field experiment. *Tinbergen Institute Discussion Paper*, 7, 2017.

M. R. Durantini, D. Albarracin, A. L. Mitchell, A. N. Earl, and J. C. Gillette. Conceptualizing the influence of social agents of behavior change: A meta-analysis of the effectiveness of hiv-prevention interventionists for different groups. *Psychological Bulletin*, 132(2):212, 2006.

E. Epel, J. Daubenmier, J. T. Moskowitz, S. Folkman, and E. Blackburn. Can meditation slow rate of cellular aging? Cognitive stress, mindfulness, and telomeres. In *Annals of the New York Academy of Sciences*, volume 1172, pages 34–53, 2009.

A. Falk and A. Ichino. Clean evidence on peer effects. *Journal of Labor Economics*, 24 (1):39–57, 2006.

J. Fan, B. D. McCandliss, T. Sommer, A. Raz, and M. I. Posner. Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14(3): 340–347, 2002.

E. Fehr and A. Rangel. Neuroeconomic foundations of economic choice - recent advances. *The Journal of Economic Perspectives*, 25(4):3–30, 2011.

A. Filippin and P. Crosetto. A reconsideration of gender differences in risk attitudes. *Management Science*, 62(11):3138–3160, 2016.

A. Finkelstein. E-ztax: Tax salience and tax rates. *The Quarterly Journal of Economics*, 124(3):969–1010, 2009.

J. Fitzgerald, P. Gottschalk, and R. Moffitt. The impact of attrition in the panel study of income dynamics on intergenerational analysis. *Journal of Human Resources*, 33 (2):123–134, 1998.

L. Flook, S. L. Smalley, M. J. Kitil, B. M. Galla, S. Kaiser-Greenland, J. Locke, E. Ishijima, and C. Kasari. Effects of mindful awareness practices on executive functions in elementary school children. *Journal of Applied School Psychology*, 26(1):70–95, 2010.

M. H. Forouzanfar, L. Alexander, H. R. Anderson, V. F. Bachman, S. Biryukov, M. Brauer, R. Burnett, D. Casey, M. M. Coates, A. Cohen, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013. *The Lancet*, 386 (10010):2287–2323, 2015.

S. Frederick, G. Loewenstein, and T. ODonoghue. Time discounting and time preference: a critical review. *Journal of Economic Literature*, 40(2):351–401, 2002.

M. Friese, C. Messner, and Y. Schaffner. Mindfulness meditation counteracts self-control depletion. *Consciousness and Cognition*, 21(2):1016–1022, 2012.

D. Fudenberg and D. K. Levine. A dual-self model of impulse control. *The American Economic Review*, 96(5):1449–1476, 2006.

N. D. Galbraith and K. E. Brown. Assessing intervention effectiveness for reducing stress in student nurses: quantitative systematic review. *Journal of Advanced Nursing*, 67(4):709–721, 2011.

D. M. Gates, P. Succop, B. J. Brehm, G. L. Gillespie, and B. D. Sommers. Obesity and presenteeism: the impact of body mass index on workplace productivity. *Journal of Occupational and Environmental Medicine*, 50(1):39–45, 2008.

M. Gaynor, R. Moreno-Serra, and C. Propper. Death by market power: reform, competition, and patient outcomes in the National Health Service. *American Economic Journal: Economic Policy*, 5(4):134–166, 2013.

M. Gaynor, C. Propper, and S. Seiler. Free to choose? Reform, choice, and consideration sets in the English National Health Service. *The American Economic Review*, 106(11):3521–3557, 2016.

B. Gilbert and J. G. Zivin. Dynamic salience with intermittent billing: Evidence from smart electricity meters. *Journal of Economic Behavior & Organization*, 107, Part A:176 – 190, 2014.

M. Goyal, S. Singh, E. M. S. Sibinga, N. F. Gould, A. Rowland-Seymour, R. Sharma, Z. Berger, D. Sleicher, D. D. Maron, H. M. Shihab, P. D. Ranasinghe, S. Linn, S. Saha, E. B. Bass, and J. a. Haythornthwaite. Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. *JAMA Internal Medicine*, 174(3):357–68, Mar. 2014.

N. Grissom and S. Bhatnagar. Habituation to repeated stress: get used to it. *Neurobiology of Learning and Memory*, 92(2):215–224, 2009.

L. K. Grossman, L. N. Rich, and C. Johnson. Decreasing nonurgent emergency department utilization by medicaid children. *Pediatrics*, 102(1):20–24, 1998.

L. Guiso, P. Sapienza, and L. Zingales. Time varying risk aversion. *NBER Working Paper*, 19284, 2013.

F. Gul and W. Pesendorfer. Self-control and the theory of consumption. *Econometrica*, pages 119–158, 2004.

R. J. Herrnstein, G. F. Loewenstein, D. Prelec, and W. Vaughan. Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making*, 6(3):149–185, 1993.

S. G. Hofmann, A. T. Sawyer, A. A. Witt, and D. Oh. The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78:169–183, 2010.

W. Hofmann, M. Friese, and R. W. Wiers. Impulsive versus reflective influences on health behavior: A theoretical framework and empirical review. *Health Psychology Review*, 2(2):111–137, 2008.

Huffington Post. Why 2014 will be the year of mindful living. `http://www.huffingtonpost.com/2014/01/02/will-2014-be-the-year-of-_0_n_4523975.html`, 2014. Accessed: 2016-04-12.

M. A. Hussey and J. P. Hughes. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*, 28(2):182–191, 2007.

A. R. Jensen and W. D. Rohwer. The Stroop color-word test: A review. *Acta Psychologica*, 25:36–93, 1966.

A. P. Jha, J. Krompinger, and M. J. Baime. Mindfulness training modifies subsystems of attention. *Cognitive, Affective, & Behavioral Neuroscience*, 7(2):109–119, 2007.

L. K. John, G. Loewenstein, A. B. Troxel, L. Norton, J. E. Fassbender, and K. G. Volpp. Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of General Internal Medicine*, 26(6):621–626, 2011.

J. Kabat-Zinn. Full catastrophe living: The program of the stress reduction clinic at the University of Massachusetts Medical Center. *New York: Delta*, 1990.

D. M. Kagan and R. L. Squires. Dieting, compulsive eating, and feelings of failure among adolescents. *International Journal of Eating Disorders*, 3(1):15–26, 1983.

D. Kahneman and D. T. Miller. Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2):136, 1986.

N. Kandasamy, B. Hardy, L. Page, M. Schaffner, J. Graggaber, A. S. Powlson, P. C. Fletcher, M. Gurnell, and J. Coates. Cortisol shifts financial risk preferences. *Proceedings of the National Academy of Sciences*, 111(9):3608–3613, 2014.

E. Kandel and E. P. Lazear. Peer pressure and partnerships. *Journal of Political Economy*, pages 801–817, 1992.

J. Kandiah, Y. Melissa, J. Jones, et al. Stress influences appetite and comfort food preferences in college women. *Nutrition Research*, 26(3):118–123, 2006.

H. Kerr. National Union of Students: Mental distress survey overview. `http://www.nus.org.uk/Global/Campaigns/20130517%20Mental%20Distress%20Survey%20%20Overview.pdf`, 2013. Accessed: 2016-04-12.

C. Kirschbaum, K.-M. Pirke, and D. H. Hellhammer. The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, 28(1-2):76–81, 1993.

A. Krusche, E. Cyhlarova, and J. M. G. Williams. Mindfulness online: an evaluation of the feasibility of a web-based mindfulness course for stress, anxiety and depression. *BMJ Open*, 3(11):e003498, Jan. 2013.

T. Kugler, T. Connolly, and L. D. Ordóñez. Emotion, decision, and risk: Betting on gambles versus betting on people. *Journal of Behavioral Decision Making*, 25(2): 123–134, 2012.

A. J. Larner. NHS Direct telephone helpline: frequency of use over time and by age and gender in an outpatient population. *Telemedicine and e-Health*, 15(2):199–201, 2009.

C. Liston, B. S. McEwen, and B. J. Casey. Psychosocial stress reversibly disrupts prefrontal processing and attentional control. *Proceedings of the National Academy of Sciences of the United States of America*, 106:912–917, 2009.

G. Loewenstein and T. O'Donoghue. Animal spirits: Affective and deliberative processes in economic behavior. *Available at SSRN 539843*, 2004.

G. Loewenstein, D. A. Asch, J. Y. Friedman, L. A. Melichar, and K. G. Volpp. Can behavioural economics make us healthier? *BMJ: British Medical Journal (Online)*, 344, 2012.

R. Luengo-Fernandez, J. Leal, A. Gray, S. Petersen, and M. Rayner. Cost of cardiovascular diseases in the united kingdom. *Heart*, 92(10):1384–1389, 2006.

S. H. Ma and J. D. Teasdale. Mindfulness-based cognitive therapy for depression: replication and exploration of differential relapse prevention effects. *Journal of Consulting and Clinical Psychology*, 72(1):31, 2004.

U. Malmendier and G. Tate. Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1):20–43, 2008.

T. M. Marteau, G. J. Hollands, and P. C. Fletcher. Changing human behavior to prevent disease: the importance of targeting automatic processes. *Science*, 337(6101): 1492–1495, 2012.

B. McEwen and J. Morrison. The Brain on Stress: Vulnerability and Plasticity of the Prefrontal Cortex over the Life Course, 2013.

F. K. McSweeney. Dynamic changes in reinforcer effectiveness: Satiation and habituation have different implications for theory and practice. *The Behavior Analyst*, 27 (2):171, 2004.

K. L. Milkman, D. Chugh, and M. H. Bazerman. How can decision making be improved? *Perspectives on Psychological Science*, 4(4):379–383, 2009.

A. Moore and P. Malinowski. Meditation, mindfulness and cognitive flexibility. *Consciousness and Cognition*, 18(1):176–186, 2009.

T. J. Morledge, D. Allexandre, E. Fox, A. Z. Fu, M. K. Higashi, D. T. Kruzikas, S. V. Pham, and P. R. Reese. Feasibility of an online mindfulness program for stress management - a randomized, controlled trial. *Annals of Behavioral Medicine*, 46(2): 137–148, 2013.

National Audit Office. Emergency admissions to hospital: managing the demand., 2013.

NHS England. Transforming urgent and emergency care services in England: Urgent and Emergency Care review. End of phase 1 report., 2013.

N. A. Nicolson. Measurement of cortisol. *Handbook of Physiological Research Methods in Health Psychology*, pages 37–74, 2008.

S. Nikolova, M. Harrison, and M. Sutton. The impact of waiting time on health gains from surgery: Evidence from a national patient-reported outcome dataset. *Health Economics*, 2015.

W. E. Northington, J. H. Brice, and B. Zou. Use of an emergency department by nonurgent patients. *The American Journal of Emergency Medicine*, 23(2):131–137, 2005.

J. H. Patton, M. S. Stanford, et al. Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, 51(6):768–774, 1995.

Populus. Populus summary: NHS England & Monitor outpatient appointment referrals, 2015.

A. J. Porcelli and M. R. Delgado. Acute stress modulates risk taking in financial decision making. *Psychological Science*, 20(3):278–283, Mar. 2009.

P. Putman, N. Antypa, P. Crysovergi, and W. A. J. Van Der Does. Exogenous cortisol acutely influences motivated decision making in healthy young men. *Psychopharmacology*, 208:257–263, 2010.

J. J. Radley, H. M. Sisti, J. Hao, A. B. Rocher, T. McCall, P. R. Hof, B. S. McEwen, and J. H. Morrison. Chronic behavioral stress induces apical dendritic reorganization in pyramidal neurons of the medial prefrontal cortex. *Neuroscience*, 125:1–6, 2004.

C. H. Rankin, T. Abrams, R. J. Barry, S. Bhatnagar, D. F. Clayton, J. Colombo, G. Coppola, M. A. Geyer, D. L. Glanzman, S. Marsland, et al. Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, 92(2):135–138, 2009.

T. S. Rector, P. J. Venus, and A. J. Laine. Impact of mailing information about nonurgent care on emergency department visits by Medicaid beneficiaries enrolled in managed care. *The American Journal of Managed Care*, 5(12):1505–1512, 1999.

C. Regehr, D. Glancy, and A. Pitts. Interventions to reduce stress in university students: A review and meta-analysis. *Journal of Affective Disorders*, 148(1):1–11, 2013.

J. W. Roberti, L. N. Harrington, and E. A. Storch. Further psychometric support for the 10-item version of the Perceived Stress Scale. *Journal of College Counseling*, 9 (2):135–147, 2006.

T. Rogers and E. Frey. Changing behavior beyond the here and now. *The Wiley Blackwell Handbook of Judgment and Decision Making*, pages 723–748, 2015.

D. Romer, L. M. Betancourt, N. L. Brodsky, et al. Does adolescent risk taking imply weak executive function? A prospective study of relations between working memory performance, impulsivity, and risk taking in early adolescence. *Developmental Science*, 14(5):1119–1133, 2011.

H. E. Rosvold, A. F. Mirsky, I. Sarason, E. D. Bransome Jr, and L. H. Beck. A continuous performance test of brain damage. *Journal of Consulting Psychology*, 20 (5):343, 1956.

P. Scarborough, P. Bhatnagar, K. K. Wickramasinghe, S. Allender, C. Foster, and M. Rayner. The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006–07 NHS costs. *Journal of Public Health*, 33(4):527–535, 2011.

R. J. Semple. Does mindfulness meditation enhance attention? A randomized controlled trial. *Mindfulness*, 1(2):121–130, 2010.

K. E. Stanovich and R. F. West. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23(5):645–665, 2000.

R. Stinebrickner and T. Stinebrickner. A major in science? Initial beliefs and final outcomes for college major and dropout. *The Review of Economic Studies*, 81(1): 426–472, 2014.

T. Stinebrickner and R. Stinebrickner. Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4):707–748, 2012.

F. Strack and R. Deutsch. Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8(3):220–247, 2004.

J. R. Stroop. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6):643, 1935.

Y.-Y. Tang, Y. Ma, J. Wang, Y. Fan, S. Feng, Q. Lu, Q. Yu, D. Sui, M. K. Rothbart, M. Fan, et al. Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences*, 104(43):17152–17156, 2007.

S. E. Taylor and S. C. Thompson. Stalking the elusive "vividness" effect. *Psychological Review*, 89(2):155, 1982.

J. D. Teasdale, Z. V. Segal, J. M. G. Williams, V. A. Ridgeway, J. M. Soulsby, and M. A. Lau. Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, 68(4):615, 2000.

R. Teper and M. Inzlicht. Meditation, mindfulness and executive control: the importance of emotional acceptance and brain-based performance monitoring. *Social Cognitive and Affective Neuroscience*, 8(1):85–92, 2013.

R. Teper, Z. V. Segal, and M. Inzlicht. Inside the mindful mind: How mindfulness enhances emotion regulation through improvements in executive control. *Current Directions in Psychological Science*, 22(6):449–454, 2013.

R. H. Thaler and S. Benartzi. Save more tomorrow ^{TM}: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1):S164–S187, 2004.

R. H. Thaler and H. Shefrin. An economic theory of self-control. *Journal of Political Economy*, 89(2):392–406, 1981.

R. H. Thaler and C. R. Sunstein. *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press, 2008.

The Behavioural Insights Team. EAST: Four simple ways to apply behavioural insights., 2014.

The University of Sheffield. Evaluation of NHS 111 pilot sites: Final report. Technical report, The University of Sheffield, 2012.

R. F. Thompson and W. A. Spencer. Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, 73(1):16, 1966.

L. Uscher-Pines, J. Pines, A. Kellermann, E. Gillen, and A. Mehrotra. Deciding to visit the emergency department for non-urgent conditions: a systematic review of the literature. *The American Journal of Managed Care*, 19(1):47, 2013.

R. van den Bos, M. Harteveld, and H. Stoop. Stress and decision-making in humans: Performance is related to cortisol reactivity, albeit differently in men and women. *Psychoneuroendocrinology*, 34(10):1449–1458, Nov. 2009.

E. M. VanEpps, J. S. Downs, and G. Loewenstein. Advance ordering for healthier eating? Field experiments on the relationship between the meal order–consumption time delay and meal content. *Journal of Marketing Research*, 53(3):369–380, 2016.

K. G. Volpp, A. B. Troxel, M. V. Pauly, H. A. Glick, A. Puig, D. A. Asch, R. Galvin, J. Zhu, F. Wan, J. DeGuzman, et al. A randomized, controlled trial of financial incentives for smoking cessation. *New England Journal of Medicine*, 360(7):699–709, 2009.

D. Von Ah, S. Ebert, A. Ngamvitroj, N. Park, and D.-H. Kang. Predictors of health behaviours in college students. *Journal of Advanced Nursing*, 48(5):463–474, 2004.

Y. Wang, K. Hunt, I. Nazareth, N. Freemantle, and I. Petersen. Do men consult less than women? an analysis of routinely collected uk general practice data. *BMJ Open*, 3(8):e003320, 2013.

H. Wenk-Sormaz. Meditation can reduce habitual responding. *Alternative Therapies in Health and Medicine*, 11(2):42, 2005.

M. Whitehead, R. Jones, R. Howell, R. Lilley, and J. Pykett. Nudging all over the world: Assessing the global impact of the behavioral sciences on public policy. Technical report, Economic and Social Research Council, UK, 2014.

World Health Organisation. World Health Organisation: Global status report on non-communicable diseases. Technical report, World Health Organisation, 2014.

Appendices

Appendix for Chapter 1

Appendix 1A: Tables and Graphs

Table A.1.1: OtherModels

	(1)	(2)	(3)	(4)
VARIABLES	LPM	LPM	Probit	Logit
Treatment	0.00122	0.00152	0.0151	0.0298
	(0.00430)	(0.00429)	(0.0401)	(0.0860)
Age		-0.000372***	-0.00360***	-0.00786***
		(0.000107)	(0.00112)	(0.00243)
Female		0.00181	0.0157	0.0349
		(0.00430)	(0.0401)	(0.0861)
2.hospital		-0.000253	-0.00848	-0.00420
		(0.00777)	(0.0724)	(0.159)
3.hospital		-0.0103	-0.116	-0.263
		(0.00685)	(0.0757)	(0.170)
1.area		0.0181***	0.169***	0.371***
		(0.00687)	(0.0619)	(0.135)
2.area		0.0207***	0.198***	0.419***
		(0.00710)	(0.0685)	(0.148)
3.area		0.0110	0.125	0.288
		(0.00734)	(0.0819)	(0.184)
Constant	0.0522***	0.0579***	-1.577***	-2.802***
	(0.00294)	(0.00747)	(0.0738)	(0.160)
Controls	No	Yes	Yes	Yes
Observations	10,889	10,889	10,889	10,889
R-squared	0.000	0.004		

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.1.2: Gender interaction

VARIABLES	(1) LPM
Treatment	0.405
	(0.620)
Female	0.408
	(0.586)
Treatment X Female	-0.481
	(0.860)
Age	-0.0372***
	(0.0107)
Hospital and Area Dummies	Yes
Constant	5.673***
	(0.773)
Observations	10,889
R-squared	0.004

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.1.3: Cox Proportional Hazard Model

VARIABLES	(1) Cox
Treatment	0.0340
	(0.0840)
Age	-0.00778***
	(0.00238)
Female	0.0405
	(0.0841)
2.hospital	-0.0127
	(0.156)
3.hospital	-0.256
	(0.167)
1.area	0.352***
	(0.132)
2.area	0.408***
	(0.146)
3.area	0.269
	(0.181)
Observations	10,883

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.1.4: Cox Proportional Hazard Model - Age Splits (Women Only)

	(1) Women up to 30	(2) Women 31-60	(3) Women over 60	(4) Women low tertile	(5) Women mid tertile	(6) Women top tertile
Treatment	0.0841	-0.127	0.0996	0.0382	-0.180	0.172
	(0.171)	(0.179)	(0.324)	(0.183)	(0.197)	(0.227)
Hospital & Area Dummies	yes	yes	yes	yes	yes	yes
Observations	2,178	2,437	1,105	1,934	1,827	1,959

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A.1.1: Survival Analysis - Women

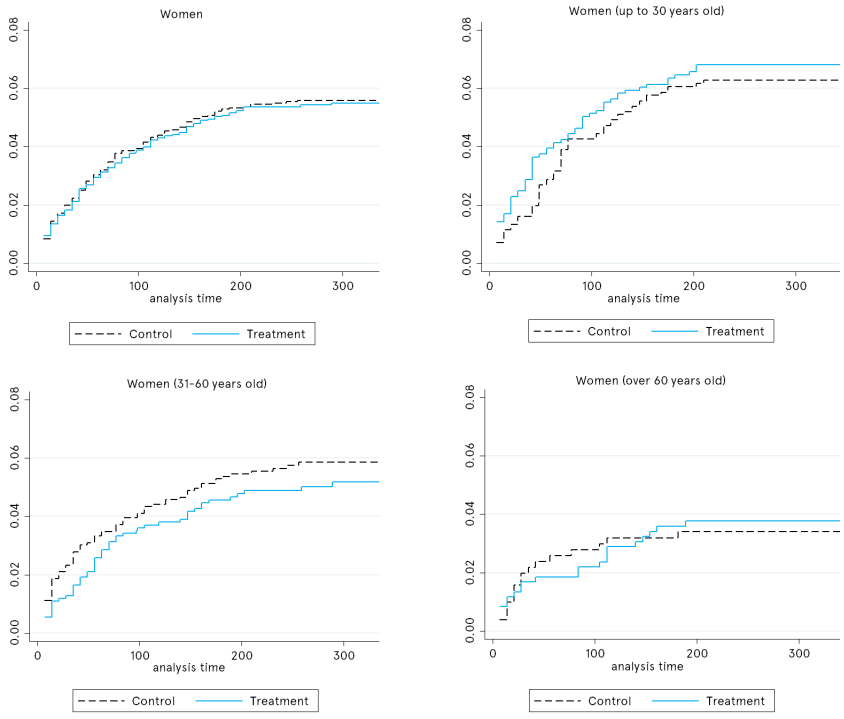
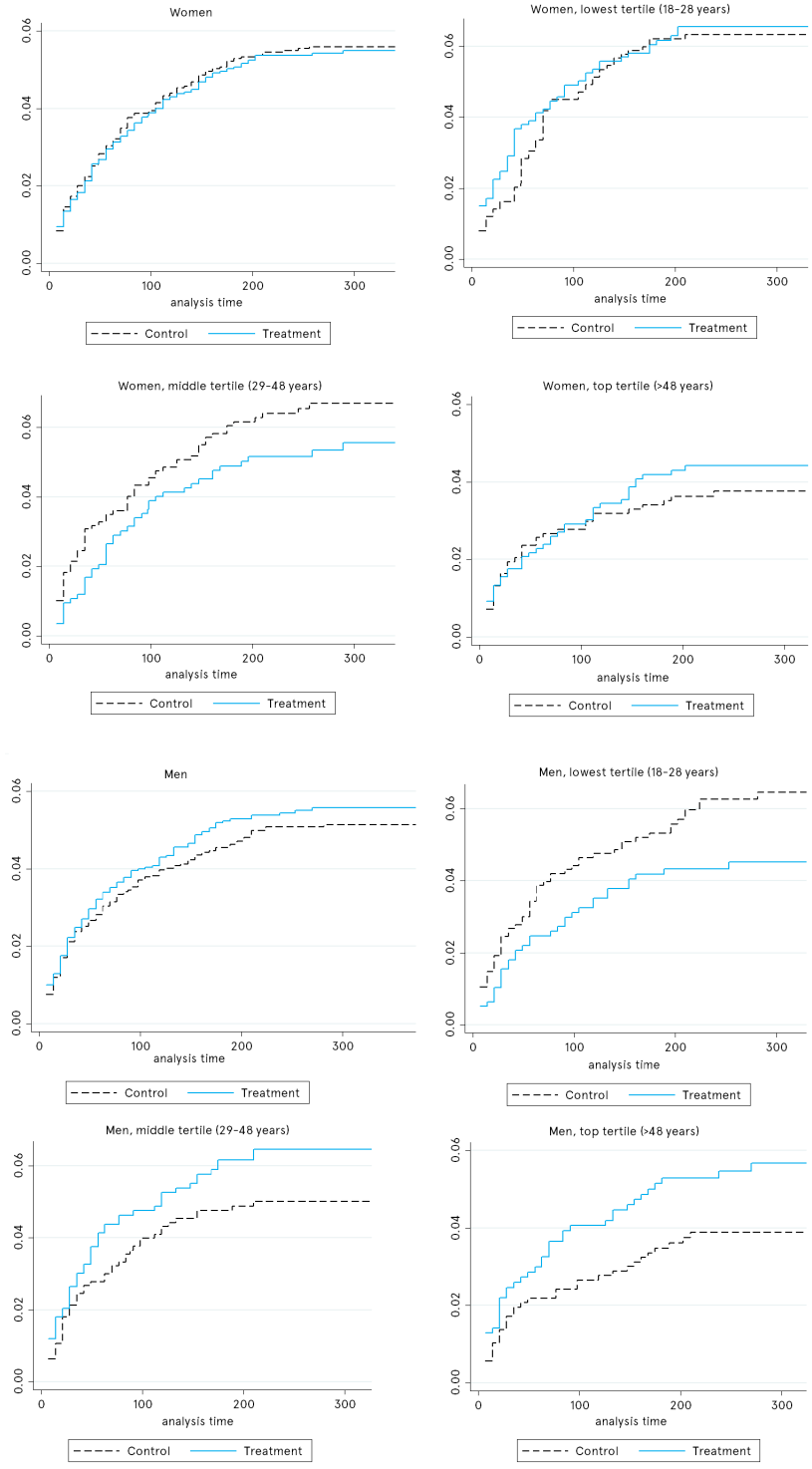


Figure A.1.2: Survival Analysis - Age Tertiles (women and men)



Appendix 1B: Intervention Letter and Leaflet

EKHUFT Hospital
Road
Kent Postcode
Call hospital phone number

Firstname Lastname
Road
Kent
Postcode

Dear Firstname Lastname,

Our records show you recently attended the Emergency Department at EKHUFT Hospital. We just wanted to tell you that there are **easy, local ways to get medical advice**, without having to wait in A&E.

HEALTH help NOW.	**If you need to get advice and find the right local health service for you** • Health Help Now is a free website and app to find the most appropriate local health service for common symptoms, and to get medical advice for people of all ages. Visit www.healthhelpnow-nhs.net
CALL 111	**If it's not an emergency, but you need medical help fast** • NHS 111 is available 24 hours a day. Calls are free from landlines and mobile phones. • 111 will give you advice and directions to any local health service that you need.
Go to a Minor Injuries Unit	**If you need to see someone today for minor illnesses and injuries** • A Nurse Practitioner can give advice and treatment for minor illnesses and injuries in adults. You do not need an appointment. • Please see the enclosed leaflet to find the centre closest to you.
Call your GP	**If you need to see a doctor for an illness or injury that won't go away** • If you need to see a GP urgently when your GP practice is closed, call 111.

Finally, if you know it is an emergency, then go to A&E or **call 999**. This information could help you in the future, so please keep this letter as a reminder of the services available to you. We have to let your GP know if you attend A&E, so we have also written to them.

Yours sincerely,

Julie Pearce

Chief Nurse, Director of Quality & Deputy Chief Executive
East Kent Hospitals University NHS Foundation Trust
@Julie_Pearce1

HEALTH help NOW.

NHS

GP weekend walk-in service

For urgent matters only

- **Ivy Court Surgery**, Tenterden TN30 6RB 01580 763666 - Saturday, 9am to noon (for patients registered at either Ivy Court, Woodchurch or Hamstreet)
- **Hamstreet Surgery,** Hamstreet TN26 2NJ 01233 730190 - Sunday, 9am to noon (for patients registered at either Ivy Court, Woodchurch or Hamstreet)
- **Faversham Medical Practice,** Bank Street, Faversham, ME13 8QR – Saturday and Sunday, 9am to 1pm (pilot project until March 2015 for patients registered at any Faversham GP practice)
- **Queen Victoria Memorial Hospital,** King Edward Avenue, Herne Bay, CT6 6EB – Saturday and Sunday, 9am to 1pm (pilot project until March 2015 for patients registered at any Herne Bay GP practice)
- **Estuary View Medical Centre**, Boorman Way, Whitstable, CT5 3SE – Saturday and Sunday, 9am to 1pm (pilot project until March 2015 for patients registered at any Whitstable GP practice)

Minor injuries units

For fractured arms or lower legs, sprains, cuts and minor burns.

- **Estuary View,** Whitstable, CT5 3SE 01227 284309 Open 8am-8pm, seven days a week. X-ray 8am-8pm, seven days a week
- **Emergency Care Centre**, Kent and Canterbury Hospital, CT1 3NG 01227 864244 Open 24 hours, seven days a week. X-ray 24 hours, seven days a week
- **Faversham Cottage Hospital,** ME13 8PS 01795 562069 Open 8am-8pm, seven days a week. No X-ray
- **Buckland Hospital**, Dover CT17 0HD 01304 222 612 Mon-Fri 9am–7pm, 10am-6pm weekends and bank holidays. X-ray 9am-4.30pm Mon-Sun
- **Victoria Hospital,** Deal CT14 9UA 01304 865400 Open 8am–8pm every day. X-ray 9am-5pm Mon-Fri
- **Royal Victoria Hospital,** Folkestone CT19 5BN 01303 854 494 8am–8pm every day. X-ray 9am-5pm Mon-Fri

Small wounds, sprains, minor burns only. Unregistered or Registered patients. Monday - Friday 8am - 6.30pm

- **Charing Surgery** TN27 0AW 01233 714490
- **Hamstreet Surgery** TN26 2NJ 01233 730190
- **Ivy Court Surgery** Tenterden TN30 6RB 01580 763666
- **Kingsnorth Medical Practice** TN23 3ED 01233 610140
- **Willesborough Health Centre** TN24 0HZ 01233 621 626
- **Woodchurch Surgery** TN26 3SF 01233 860236
- **Wye Surgery** TN25 5AY 01233 220699
- **Orchard House Surgery** (8.30am–6pm, Mon-Fri) TN29 9AE 01797 320307
- **Aylesham Medical Practice** CT3 3BB 01304 840415

Whatever the time, wherever you are, find the right treatment for you.

www.healthhelpnow-nhs.net

Appendix for Chapter 2

Appendix 2A: Tables

Table A.2.1: Descriptive statistics - referrals during the trial period

Specialty	Bookings	Waitlist entries	Total	Percent
Dermatology	1,393	346	1,739	21.02
ENT	1,190	262	1,452	17.55
Orthopaedics	3,636	611	4,247	51.34
General Surgery	714	121	835	10.09
Total	6,933	1,340	8,273	100.00

Table A.2.2: Descriptive statistics - referrals during the trial period (broader sample used for analysis of 'available capacity' alerts)

Specialty	Bookings	Waitlist entries	Total	Percent
Dermatology	1,393	346	1,739	16.84
ENT	1,190	262	1,452	14.06
Orthopaedics	3,636	611	4,247	41.12
General Surgery	714	121	835	8.08
GI and Liver	978	415	1,393	13.49
Rheumatology	472	190	662	6.41
Total	8,383	1,945	10,328	100.00

Table A.2.3: Impact of the Intervention on Propensity to Refer to Services with Available Capacity

	(1) Logistic regression
Treatment	0.0483 (0.31)
ENT	-1.027*** (-10.97)
GI and Liver	-3.297*** (-15.35)
Orthopaedics	-2.310*** (-20.07)
Rheumatology	-5.404*** (-7.89)
Day fixed effects	yes
Constant	-1.025*** (-4.38)
Observations	9452

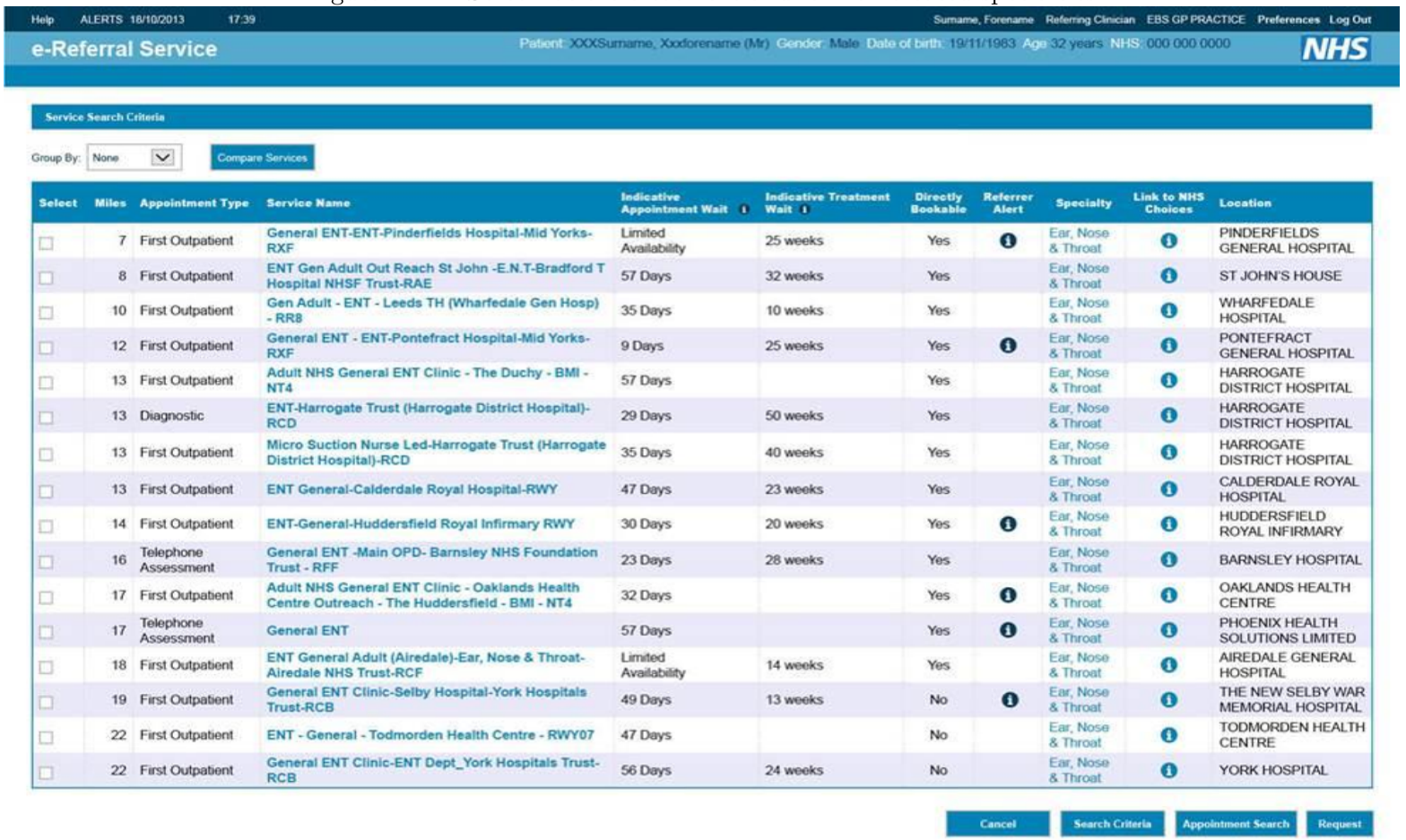
t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The overall sample size is slightly larger than in the analysis of the 'limited capacity' alerts, as I also include referrals to two additional specialties because some of the 'available capacity' services identified as suitable alternatives for the 'limited capacity' services were assigned to a different clinical specialty (Liver & GI for General Surgery and and Rheumatology for Orthopaedics).

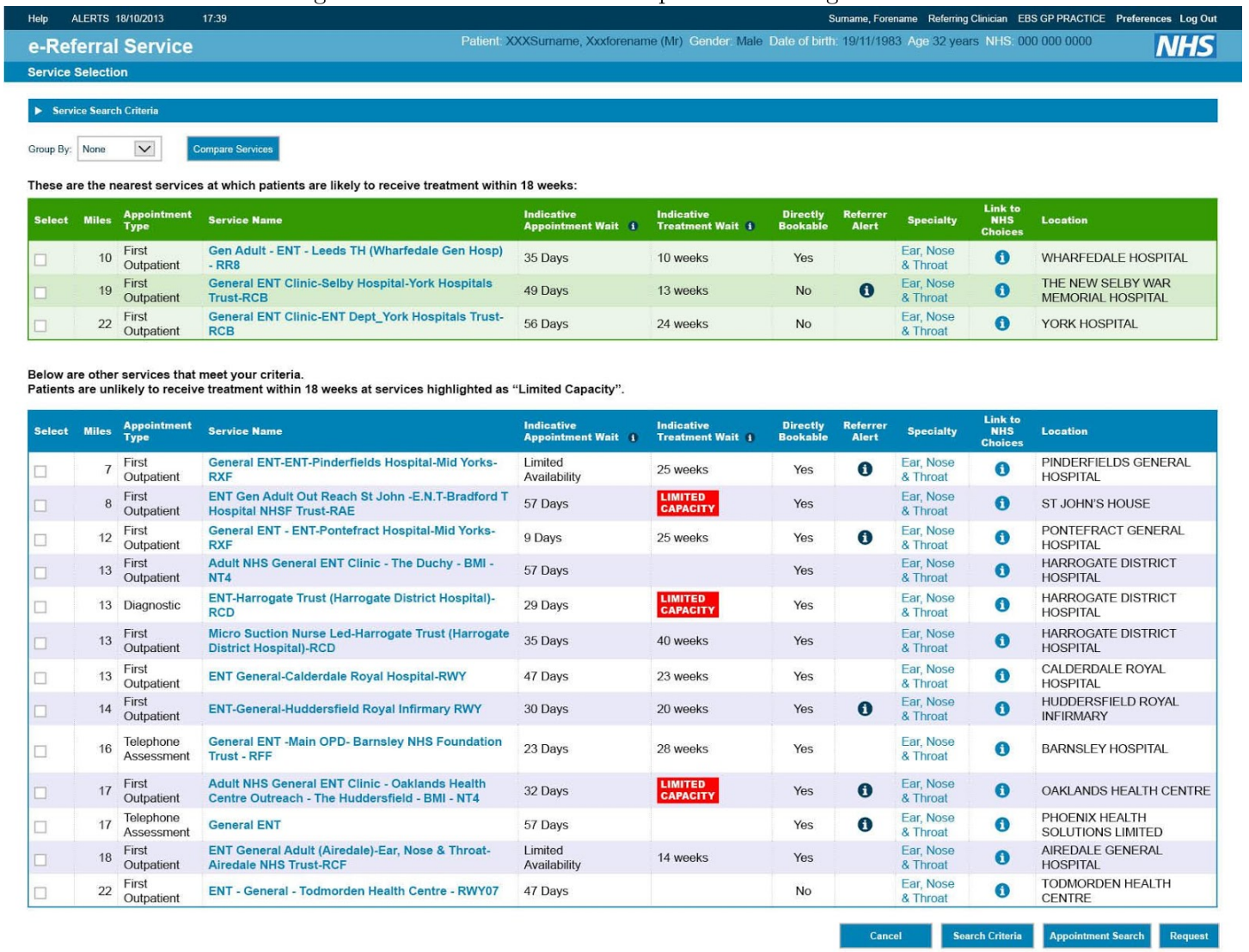
Appendix 2B: Intervention Screens

Figure A.2.1: Screen before treatment and for non-treated specialties



Select	Miles	Appointment Type	Service Name	Indicative Appointment Wait	Indicative Treatment Wait	Directly Bookable	Referrer Alert	Specialty	Link to NHS Choices	Location
☐	7	First Outpatient	General ENT-ENT-Pinderfields Hospital-Mid Yorks-RXF	Limited Availability	25 weeks	Yes	ℹ	Ear, Nose & Throat	ℹ	PINDERFIELDS GENERAL HOSPITAL
☐	8	First Outpatient	ENT Gen Adult Out Reach St John -E.N.T-Bradford T Hospital NHSF Trust-RAE	57 Days	32 weeks	Yes		Ear, Nose & Throat	ℹ	ST JOHN'S HOUSE
☐	10	First Outpatient	Gen Adult - ENT - Leeds TH (Wharfedale Gen Hosp) - RR8	35 Days	10 weeks	Yes		Ear, Nose & Throat	ℹ	WHARFEDALE HOSPITAL
☐	12	First Outpatient	General ENT - ENT-Pontefract Hospital-Mid Yorks-RXF	9 Days	25 weeks	Yes	ℹ	Ear, Nose & Throat	ℹ	PONTEFRACT GENERAL HOSPITAL
☐	13	First Outpatient	Adult NHS General ENT Clinic - The Duchy - BMI - NT4	57 Days		Yes		Ear, Nose & Throat	ℹ	HARROGATE DISTRICT HOSPITAL
☐	13	Diagnostic	ENT-Harrogate Trust (Harrogate District Hospital)-RCD	29 Days	50 weeks	Yes		Ear, Nose & Throat	ℹ	HARROGATE DISTRICT HOSPITAL
☐	13	First Outpatient	Micro Suction Nurse Led-Harrogate Trust (Harrogate District Hospital)-RCD	35 Days	40 weeks	Yes		Ear, Nose & Throat	ℹ	HARROGATE DISTRICT HOSPITAL
☐	13	First Outpatient	ENT General-Calderdale Royal Hospital-RWY	47 Days	23 weeks	Yes		Ear, Nose & Throat	ℹ	CALDERDALE ROYAL HOSPITAL
☐	14	First Outpatient	ENT-General-Huddersfield Royal Infirmary RWY	30 Days	20 weeks	Yes	ℹ	Ear, Nose & Throat	ℹ	HUDDERSFIELD ROYAL INFIRMARY
☐	16	Telephone Assessment	General ENT -Main OPD- Barnsley NHS Foundation Trust - RFF	23 Days	28 weeks	Yes		Ear, Nose & Throat	ℹ	BARNSLEY HOSPITAL
☐	17	First Outpatient	Adult NHS General ENT Clinic - Oaklands Health Centre Outreach - The Huddersfield - BMI - NT4	32 Days		Yes	ℹ	Ear, Nose & Throat	ℹ	OAKLANDS HEALTH CENTRE
☐	17	Telephone Assessment	General ENT	57 Days		Yes	ℹ	Ear, Nose & Throat	ℹ	PHOENIX HEALTH SOLUTIONS LIMITED
☐	18	First Outpatient	ENT General Adult (Airedale)-Ear, Nose & Throat-Airedale NHS Trust-RCF	Limited Availability	14 weeks	Yes		Ear, Nose & Throat	ℹ	AIREDALE GENERAL HOSPITAL
☐	19	First Outpatient	General ENT Clinic-Selby Hospital-York Hospitals Trust-RCB	49 Days	13 weeks	No	ℹ	Ear, Nose & Throat	ℹ	THE NEW SELBY WAR MEMORIAL HOSPITAL
☐	22	First Outpatient	ENT - General - Todmorden Health Centre - RWY07	47 Days		No		Ear, Nose & Throat	ℹ	TODMORDEN HEALTH CENTRE
☐	22	First Outpatient	General ENT Clinic-ENT Dept_York Hospitals Trust-RCB	56 Days	24 weeks	No		Ear, Nose & Throat	ℹ	YORK HOSPITAL

Figure A.2.2: Screen for treated specialties during treatment



Help ALERTS 18/10/2013 17:39 Surname, Forename Referring Clinician EBS GP PRACTICE Preferences Log Out

e-Referral Service

Patient: XXXSurname, Xxxforename (Mr) Gender: Male Date of birth: 19/11/1983 Age 32 years NHS: 000 000 0000

NHS

Service Selection

▶ Service Search Criteria

Group By: None Compare Services

These are the nearest services at which patients are likely to receive treatment within 18 weeks:

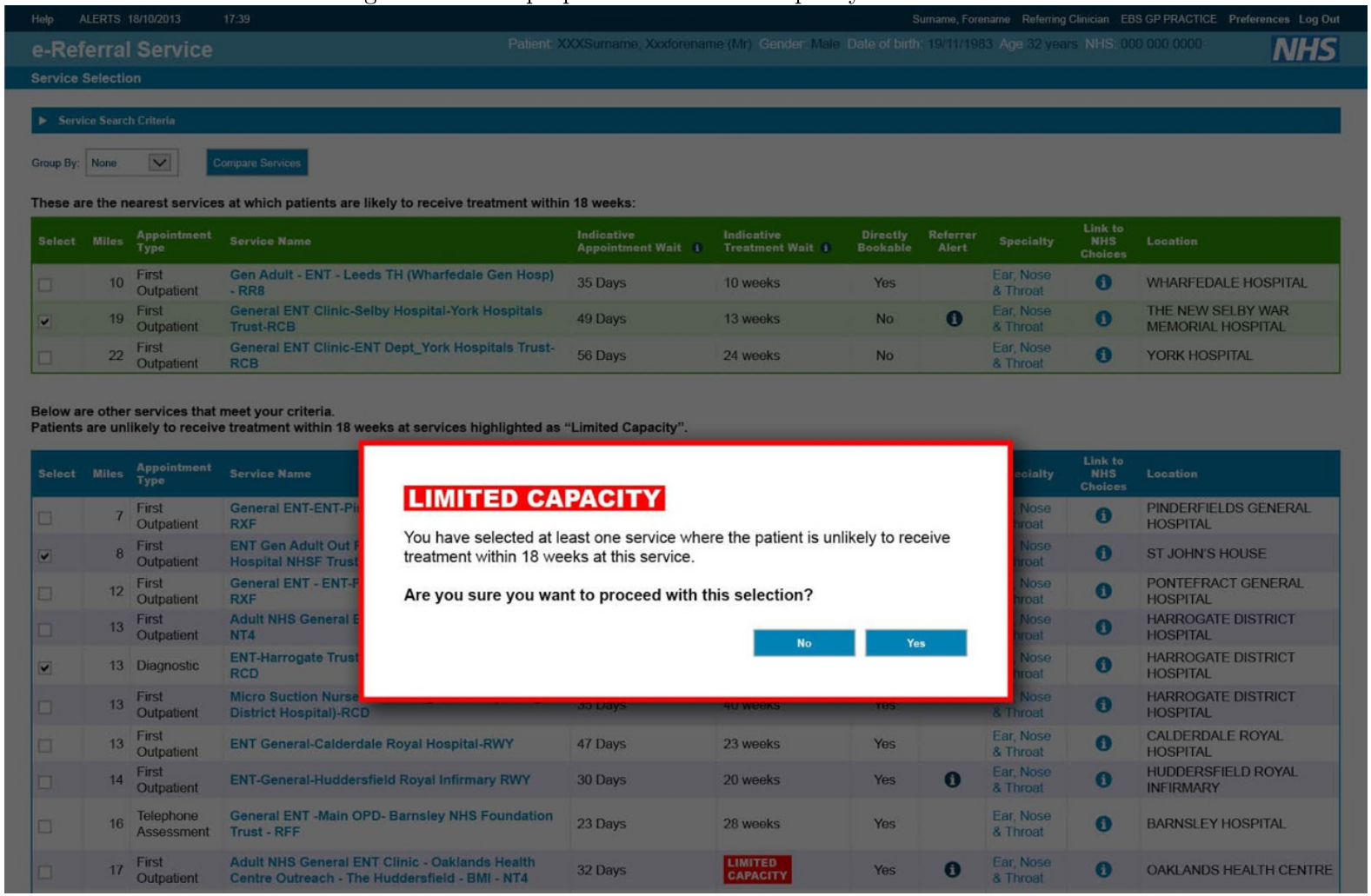
Select	Miles	Appointment Type	Service Name	Indicative Appointment Wait	Indicative Treatment Wait	Directly Bookable	Referrer Alert	Specialty	Link to NHS Choices	Location
☐	10	First Outpatient	Gen Adult - ENT - Leeds TH (Wharfedale Gen Hosp) - RR8	35 Days	10 weeks	Yes		Ear, Nose & Throat	ℹ	WHARFEDALE HOSPITAL
☐	19	First Outpatient	General ENT Clinic-Selby Hospital-York Hospitals Trust-RCB	49 Days	13 weeks	No	ℹ	Ear, Nose & Throat	ℹ	THE NEW SELBY WAR MEMORIAL HOSPITAL
☐	22	First Outpatient	General ENT Clinic-ENT Dept_York Hospitals Trust-RCB	56 Days	24 weeks	No		Ear, Nose & Throat	ℹ	YORK HOSPITAL

Below are other services that meet your criteria.
Patients are unlikely to receive treatment within 18 weeks at services highlighted as "Limited Capacity".

Select	Miles	Appointment Type	Service Name	Indicative Appointment Wait	Indicative Treatment Wait	Directly Bookable	Referrer Alert	Specialty	Link to NHS Choices	Location
☐	7	First Outpatient	General ENT-ENT-Pinderfields Hospital-Mid Yorks-RXF	Limited Availability	25 weeks	Yes	ℹ	Ear, Nose & Throat	ℹ	PINDERFIELDS GENERAL HOSPITAL
☐	8	First Outpatient	ENT Gen Adult Out Reach St John -E.N.T-Bradford T Hospital NHSF Trust-RAE	57 Days	LIMITED CAPACITY	Yes		Ear, Nose & Throat	ℹ	ST JOHN'S HOUSE
☐	12	First Outpatient	General ENT - ENT-Pontefract Hospital-Mid Yorks-RXF	9 Days	25 weeks	Yes	ℹ	Ear, Nose & Throat	ℹ	PONTEFRACT GENERAL HOSPITAL
☐	13	First Outpatient	Adult NHS General ENT Clinic - The Duchy - BMI - NT4	57 Days		Yes		Ear, Nose & Throat	ℹ	HARROGATE DISTRICT HOSPITAL
☐	13	Diagnostic	ENT-Harrogate Trust (Harrogate District Hospital)-RCD	29 Days	LIMITED CAPACITY	Yes		Ear, Nose & Throat	ℹ	HARROGATE DISTRICT HOSPITAL
☐	13	First Outpatient	Micro Suction Nurse Led-Harrogate Trust (Harrogate District Hospital)-RCD	35 Days	40 weeks	Yes		Ear, Nose & Throat	ℹ	HARROGATE DISTRICT HOSPITAL
☐	13	First Outpatient	ENT General-Calderdale Royal Hospital-RWY	47 Days	23 weeks	Yes		Ear, Nose & Throat	ℹ	CALDERDALE ROYAL HOSPITAL
☐	14	First Outpatient	ENT-General-Huddersfield Royal Infirmary RWY	30 Days	20 weeks	Yes	ℹ	Ear, Nose & Throat	ℹ	HUDDERSFIELD ROYAL INFIRMARY
☐	16	Telephone Assessment	General ENT -Main OPD- Barnsley NHS Foundation Trust - RFF	23 Days	28 weeks	Yes		Ear, Nose & Throat	ℹ	BARNSLEY HOSPITAL
☐	17	First Outpatient	Adult NHS General ENT Clinic - Oaklands Health Centre Outreach - The Huddersfield - BMI - NT4	32 Days	LIMITED CAPACITY	Yes	ℹ	Ear, Nose & Throat	ℹ	OAKLANDS HEALTH CENTRE
☐	17	Telephone Assessment	General ENT	57 Days		Yes	ℹ	Ear, Nose & Throat	ℹ	PHOENIX HEALTH SOLUTIONS LIMITED
☐	18	First Outpatient	ENT General Adult (Airedale)-Ear, Nose & Throat-Airedale NHS Trust-RCF	Limited Availability	14 weeks	Yes		Ear, Nose & Throat	ℹ	AIREDALE GENERAL HOSPITAL
☐	22	First Outpatient	ENT - General - Todmorden Health Centre - RWY07	47 Days		No		Ear, Nose & Throat	ℹ	TODMORDEN HEALTH CENTRE

Cancel Search Criteria Appointment Search Request

Figure A.2.3: Pop-up shown if limited capacity service selected



Appendix for Chapter 3

Appendix A: Recruitment Poster and Email

BEHAVIOURAL LABORATORY AT THE UNIVERSITY OF EDINBURGH

Participate in a LIFESTYLE study!

The School of Economics at the University of Edinburgh is currently seeking

150 HEALTHY individuals

Do you sometimes feel stressed? Want to participate in a scientific study and earn a bit of money?

Are you available for a couple of hours a week (at a time of your choice) between mid-October and mid-November?

GOAL OF THE STUDY

The goal of this study is to look at stress and lifestyle among university students. If you decide to participate, you will be given a specific protocol that you will be kindly requested to follow for four consecutive weeks. These protocols are non-invasive and will include requests to undertake certain activities during the week, for a period of four weeks starting immediately after the initial session.

Note: you must be **at least 18 years old,** a **student** at the University of Edinburgh, and have **NO pre-existing medical conditions**

INTERESTED? WANT MORE INFORMATION?

E-mail: blue@ed.ac.uk

FEELING A BIT STRESSED? WANT TO EARN SOME MONEY? HAVE A COUPLE OF HOURS PER WEEK TO SPARE IN THE COMING TWO MONTHS?

Participate to our study on "Stress and Lifestyle among University Students".

The Behavioural Laboratory at the University of Edinburgh is currently seeking 150 HEALTHY individuals for a scientific study on stress and lifestyle among university students.

- **You will be asked to come every week to our laboratory at a specific timeslot (the same day and same time every week) for a period of 6 weeks (starting in the week of October 22d) and another time 4 months later (in March 2015).**
- **You will be asked to follow a specific protocol in between (more information below).**

Please read on before signing up.

DESCRIPTION OF THE STUDY

The goal of this study is to look at stress and lifestyle among university students. If you decide to participate, you will be given a **specific protocol** that you will be kindly requested to follow for **four consecutive weeks**. These protocols are non-invasive and will include requests to undertake certain activities during the week, for a period of four weeks starting immediately after the initial session. **It is very important for our study that you agree to follow the protocol's instructions**. These activities should not take more than 2 hours a week and **we will pay all costs involved**.

The study will take place over the course of 6 weeks and an additional follow-up in six months. You will be asked to come to our experimental laboratory (situated at the) every week (6 times in total including the 6 months follow up). You will be asked to come every week **on the same day and same time slot (this is VERY important for our analysis, so please do pick your timeslot carefully and make sure you can come every week)**.

Each time, you will be asked to answer basic survey questions (including basic background information), questions about your lifestyle and health, feedback on the protocol you have been asked to follow and you will be asked to take decisions that will involve monetary rewards (you can only earn positive amounts). **You should expect to receive between £3 and £10 in each session (the exact amount will depend on your decisions).** The sessions in weeks 1 and 6 (and in March 2015) will last about an hour and a half. The other sessions will take less than half an hour each.

Note that in sessions 1 and 6, we will collect saliva samples using a standard scientific protocol. The protocol is non invasive and completely safe. The goal is to measure cortisol concentration (as an indicator of stress levels). We kindly ask you not to drink or eat anything one hour prior to the session.

Eligibility criteria:

- You must be older than 18 years old, student at the University of Edinburgh, with NO medical condition

Possible time slot options (you can only choose one option and are asked to stick to the days and times once you have picked that option)

	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6
Week 1 Initial session (1 hour 30 min) Pay: £3-10	Tue 21/10 10 am	Tue 21/10 12.30 pm	Tue 21/10 3 pm	Wed 22/10 10 am	Wed 22/10 12.30 pm	Wed 22/10 3.30 pm
Week 2 (30 min) Pay: £7.50	Tue 28/10 10 am	Tue 28/10 12.30 pm	Tue 28/10 3 pm	Wed 29/10 10 am	Wed 29/10 12.30 pm	Wed 29/10 3.30 pm
Week 3 (30 min) Pay: £7.50	Tue 4/11 10 am	Tue 4/11 12.30 pm	Tue 4/11 3 pm	Wed 5/11 10 am	Wed 5/11 12.30 pm	Wed 5/11 3.30 pm
Week 4 (30 min) Pay: £7.50	Tue 11/11 10 am	Tue 11/11 12.30 pm	Tue 11/11 3 pm	Wed 12/11 10 am	Wed 12/11 12.30 pm	Wed 12/11 3.30 pm
Week 5 (30 min) Pay: £7.50	Tue 18/11 10 am	Tue 18/11 12.30 pm	Tue 18/11 3 pm	Wed 19/11 10 am	Wed 19/11 12.30 pm	Wed 19/11 3.30 pm
Week 6 (1 hour 30 min) Pay: £3-£10	Tue 25/11 10 am	Tue 25/11 12.30 pm	Tue 25/11 3 pm	Wed 26/11 10 am	Wed 26/11 12.30 pm	Wed 26/11 3.30 pm
4 months later (1 hour 30 min) Pay: £3-£10	Tue 18/3 10 am	Tue 18/3 12.30 pm	Tue 18/3 3 pm	Wed 19/3 10 am	Wed 19/3 12.30 pm	Wed 19/3 3.30 pm

Location of the sessions: Behavioural Laboratory at the University of Edinburgh, School of Economics, 31 Buccleuch Place, 4th floor

Ethical issues and Informed consent

Note that the study is conducted with ethical approval of the School of Economics at the University of Edinburgh. All the data will be anonymised and treated with confidentiality, in accordance with the ethical guidelines.

You will be asked to sign an informed consent form at the beginning of the initial session. Although we emphasize the importance of participating to all sessions, you will be free to withdraw from the study at any point in time.

Interested? Please e-mail blue@ed.ac.auk and indicate which option you would prefer.

Appendix B: Evaluation of the Treatment and Control Intervention for Relaxation Purposes

How useful was the program (control: BBC Ancient Worlds; treatment: Mindfulness) for relaxation purposes? 1-very useful, 2-somewhat useful, 3-not useful at all.

Session		Control, previous week	Control, overall	Treatment, previous week	Treatment, overall
2	mean (sd)	2.00 (0.54)		2.05 (0.57)	
	median	2		2	
3	mean (sd)	2.10 (0.65)		1.92 (0.65)	
	median	2		2	
4	mean (sd)	1.97 (0.54)		1.97 (0.60)	
	median	2		2	
5	mean (sd)	2.03 (0.57)		1.86 (0.52)	
	median	2		2	
6	mean (sd)	2.10 (0.61)	2.13 (0.55)	1.98 (0.51)	2.02 (0.51)
	median	2	2	2	2

Table A.3.1: Evaluation of the Treatment and Control Intervention for Relaxation Purposes

Appendix C: Weekly Survey

Weekly survey

Q1 Please enter your desk number.

Q2 This survey aims to collect information on various aspects of your lifestyle and activities in the previous week. The survey takes around 30 minutes to complete. All data collected in this survey will be held anonymously and securely, and will be used for research purposes only. Please click next to continue.

Q3 Please enter your participant ID. This is the code consisting of letters and numbers on the card you were given at registration. Please don't put any spaces between the letters and numbers.

Q4 We would like to ask you some questions about your feelings on aspects of your life. There are no right or wrong answers. For each of these questions we'd like you to give an answer on a scale of 0 to 10, where 0 is 'not at all' and 10 is 'completely'.

	0 (1)	1 (2)	2 (3)	3 (4)	4 (5)	5 (6)	6 (7)	7 (8)	8 (9)	9 (10)	10 (11)
Overall, how satisfied were you with your life in the previous week? (1)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Overall, to what extent do you feel that the things you did the previous week were worthwhile? (2)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Overall, how happy are you these days? (3)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Overall, how happy are you right now? (5)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Overall, how anxious do	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

you feel these days? (6) Overall, how anxious do you feel right now? (4)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

Q97 How many hours did you spend studying the previous week? Do NOT include hours spent in classes, but DO include hours spent studying alone, in the library or with classmates.

Q98 Did you have any midterm exams in the previous week?
❍ Yes (1)
❍ No (2)

Q99 Did you have to submit any assignments in the previous week?
❍ Yes (1)
❍ No (2)

Q100 Did anything unusually upsetting or stressful happen to you in the previous week?
❍ Yes (1)
❍ No (2)

Answer If Did anything unusually upsetting or stressful happen to you in the previous week? Yes Is Selected

Q101 Please provide some details.

Q5 The next questions are about your health behaviours during the PREVIOUS WEEK.

Q102 Please read all the following statements carefully and tick the box next to the one that best describes you. During the previous week:
❍ I did not smoke any cigarette, not even a puff (1)
❍ I smoked cigarettes, but fewer than one per day (2)
❍ I smoked between 1 and 10 cigarettes per day (3)
❍ I smoked between 10 and 20 cigarettes per day (4)
❍ I smoked more than 20 cigarettes per day (5)

Q8 How often did you eat breakfast in the previous week?
❍ Almost every day (1)
❍ Most days a week (2)
❍ About once a week (3)
❍ Never (5)

Q9 How often did you eat lunch in the previous week?

- ❍ Almost every day (1)
- ❍ Most days a week (2)
- ❍ About once a week (3)
- ❍ Never (5)

Q10 How often did you eat dinner in the previous week?

- ❍ Almost every day (1)
- ❍ Most days a week (2)
- ❍ About once a week (3)
- ❍ Never (5)

Q11 Did you eat at regular times of the day during the previous week?

- ❍ Almost always (1)
- ❍ Most days (2)
- ❍ Sometimes (3)
- ❍ Never (4)

Q12 On average, how many meals did you eat each day during the previous week?

Q13 On average, how many between-meal snacks did you eat each day during the previous week?

Q16 These questions ask about what you ate or drank YESTERDAY. Tick all relevant boxes for each item (you can tick more than once as you could have the same type of meal for example for lunch and for dinner as well). There are no right or wrong answers. Did you eat (drink)…

	For breakfast (1)	For lunch (2)	For dinner (3)	Between main meals (4)	Not at all (5)
Processed meats like bacon, ham, sausage, or processed lunchmeats (1)	❑	❑	❑	❑	❑
Deep fried food, e.g. chips, onion rings, fried chicken, battered fish (2)	❑	❑	❑	❑	❑
Burgers, hot dog, pizza, sausage rolls (3)	❑	❑	❑	❑	❑
Potato crisps (4)	❑	❑	❑	❑	❑
Vegetable crisps (5)	❑	❑	❑	❑	❑
Popcorn, salted peanuts (6)	❑	❑	❑	❑	❑
Unsalted nuts, seeds (e.g. sunflower, pumpkin) (7)	❑	❑	❑	❑	❑
Pretzels, crackers, e.g. Ritz (8)	❑	❑	❑	❑	❑
Pies (savoury or sweet) (9)	❑	❑	❑	❑	❑
Cakes, muffins, brownies, cookies (10)	❑	❑	❑	❑	❑
Doughnut, pastry, e.g. Danish pastry, croissant, pain chocolat (11)	❑	❑	❑	❑	❑

Chocolate, candy bars, candies (12)	❑	❑	❑	❑	❑
Ice cream (13)	❑	❑	❑	❑	❑
Energy bar, high protein bar, e.g. Zone, PowerBar (14)	❑	❑	❑	❑	❑
Breakfast bar, e.g. Nutri-Grain (15)	❑	❑	❑	❑	❑
Soft drinks, e.g. Coke, Fanta, sugared sweetened fruit juices (16)	❑	❑	❑	❑	❑
Energy drinks, e.g. RedBull (17)	❑	❑	❑	❑	❑

Q17 How many servings of fruit did you eat yesterday? One serving is about one cup of chopped or sliced fruits, or one medium sized apple of banana.

Q18 Not counting potatoes, how many servings of vegetables did you eat yesterday? One serving is about one cup of chopped or sliced vegetables.

Q19 How often did you drink coffee, latte or cappuccino (not decaf) in the previous week?

- ❍ Never (1)
- ❍ About once a week (3)
- ❍ Every 2 or 3 days (4)
- ❍ Once a day (5)
- ❍ Twice a day (6)
- ❍ At least three times a day (7)

Q25 The next questions are about drinking alcohol, including beer, wine, spirits and any other alcoholic drink.

Q26 How many days over the previous week did you have an alcoholic drink?

- ❍ Almost every day (1)
- ❍ Most days a week (2)
- ❍ About once or twice (3)
- ❍ Never (5)

Q27 On the days that you did drink during the previous week, how many drinks did you have, on average? One drink is a glass of wine, or a pint of beer or cider, or 25 ml of spirits.

Q28 How often did each of the following happen to you during the previous week?

	All of the time (1)	Often (2)	Sometimes (3)	Rarely (4)	Never (5)
Felt completely out of control when it came to food. (1)	❍	❍	❍	❍	❍
Ate too much because you were upset, nervous or stressed. (2)	❍	❍	❍	❍	❍
Ate too much because you were bored or felt lonely. (3)	❍	❍	❍	❍	❍
Ate so much food so fast that you didn't know how much you ate or how it tasted. (4)	❍	❍	❍	❍	❍
Ate more than usual while preparing for an exam or working on an assignment. (5)	❍	❍	❍	❍	❍
Ate high calorie snacks while studying or working on assignments. (6)	❍	❍	❍	❍	❍

Q33 At what time did you go to sleep most days during the previous week?

- ❍ before 8pm (1)
- ❍ 8-9pm (2)
- ❍ 9-10pm (3)
- ❍ 10-11pm (4)
- ❍ 11pm-midnight (5)
- ❍ midnight-1am (6)
- ❍ 1am-2am (7)
- ❍ after 2am (8)

Q34 On average, how many hours of sleep did you get in a 24 hour period during the previous week?

Q35 How did you relax during the previous week?

	Almost every day (1)	Most days a week (2)	About once a week (3)	Never (5)
Watch movies / read books / listen to music (1)	❍	❍	❍	❍
Go to the cinema / theatre / concert (2)	❍	❍	❍	❍
Meet with friends (3)	❍	❍	❍	❍
Yoga / pilates / tai chi / chi gong or similar exercises (4)	❍	❍	❍	❍
Meditate / do breathing exercises / practice mindfulness (5)	❍	❍	❍	❍
Do sport activities (6)	❍	❍	❍	❍
Other (7)	❍	❍	❍	❍

Answer If How do you relax? Other - Almost every day Is Selected Or How do you relax? Other - About once or twice a week Is Selected Or How do you relax? Other - About once or twice a month Is Selected

Q36 What other activities did you do to relax, not listed above?

Q75 What time did you get up today?

Q76 What time did you go to sleep last night?

Q47 Please enter the code announced by the experimenters to continue.

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to AW37

Q103 Please summarize in at least 100 words the episode of the Ancient Worlds series you watched the previous week. You might add some of the following details: What were the main locations and topics? What did you learn from the documentary? Which parts did you find the most interesting or stunning?

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to AW37

Q104 Please provide a critical review of at least 50 words of the Ancient Worlds episode you watched the previous week. Feel free to add positive and negative remarks as well.

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to AW37

Q105 Did you find watching the Ancient Worlds episode useful for relaxation purposes?

- ❍ Very useful (1)
- ❍ Somewhat useful (2)
- ❍ Not useful at all (3)

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to AW37

Q106 Overall, how would you rate the Ancient Worlds episode you watched during the previous week?

- ❍ Excellent (1)
- ❍ Very good (2)
- ❍ Fair (3)
- ❍ Poor (4)

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to AW37

Q107 Would you recommend the Ancient Worlds documentary series to a friend?

- ❍ Yes (1)
- ❍ No (2)

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q108 How many days of the previous week did you practice mindfulness?

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q109 Overall, how many hours did you spend with learning and practicing mindfulness during the previous week?

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q110 Please describe in at least 100 words the mindfulness exercises you did the previous week.

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q111 What did you gain from the mindfulness course during the previous week? Please describe in at least 50 words. Feel free to add critical remarks as well.

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q112 How difficult/easy did you find practising mindfulness during the previous week?

- ❍ Very difficult (1)
- ❍ Difficult (2)
- ❍ Neither difficult nor easy (3)
- ❍ Easy (4)
- ❍ Very easy (5)

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q113 How useful did you find mindfulness for relaxation purposes during the previous week?

- ❍ Very useful (1)
- ❍ Somewhat useful (2)
- ❍ Not useful at all (3)

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q114 Overall, how would you rate the previous week's mindfulness instruction?

- ❍ Excellent (1)
- ❍ Very god (2)
- ❍ Good (3)
- ❍ Fair (4)
- ❍ Poor (5)

Answer If Please enter the code announced by the experimenters to continue. Text Response Is Equal to MI28

Q115 Would you recommend the mindfulness course to a friend?

- ❍ Yes (1)
- ❍ No (2)

Appendix D: Time Preference and Revealed Preference Measures

Table A.3.2: Time preference measure Session 1

Question	This Week (£)	Next Week (£)	Question	Next Week (£)	In 2 Weeks (£)
1	3.80	4.00	11	3.80	4.00
2	3.60	4.00	12	3.60	4.00
3	3.40	4.00	13	3.40	4.00
4	3.20	4.00	14	3.20	4.00
5	3.00	4.00	15	3.00	4.00
6	2.80	4.00	16	2.80	4.00
7	2.60	4.00	17	2.60	4.00
8	2.40	4.00	18	2.40	4.00
9	2.20	4.00	19	2.20	4.00
10	2.00	4.00	20	2.00	4.00

Table A.3.3: Time preference measure Sessions 6 and 7

Question	This Week (£)	In 2 Weeks (£)	Question	In 4 Months (£)	In 4 Months & 2 Wks (£)
1	30	31	6	30	31
2	30	32	7	30	32
3	30	33	8	30	33
4	30	34	9	30	34
5	30	35	10	30	35

Table A.3.4: Revealed preference measure

Scenario	Current Choice	Price (£)
1	Option 1: high calorie	2.60
	Option 2: low calorie	2.00
2	Option 1: high calorie	2.40
	Option 2: low calorie	2.00
3	Option 1: high calorie	2.20
	Option 2: low calorie	2.00
4	Option 1: high calorie	2.00
	Option 2: low calorie	2.00
5	Option 1: high calorie	1.80
	Option 2: low calorie	2.00
6	Option 1: high calorie	1.60
	Option 2: low calorie	2.00
7	Option 1: high calorie	1.40
	Option 2: low calorie	2.00

Appendix E: Further Results on Compliance

Table A.3.5: Simple Attrition Probit without Controls

	(1) Present Session 6	(2) Present Session 7
Treatment	-0.362	-0.341
	(0.245)	(0.219)
Constant	1.025***	0.549***
	(0.180)	(0.156)
N	139	139

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3.6: Effect of MBSR on the frequency of meditation (0-less than once a week to 3-almost every day)

	Coeff.	SE
MBSR	0.178**	0.088
Session 2	0.103	0.063
Session 3	0.060	0.059
Session 4	0.060	0.041
Session 5	0.022	0.047
Session 6	0.012	0.041
Session 7	0.165*	0.090
MBSR & Session 2	0.694***	0.138
MBSR & Session 3	1.076***	0.123
MBSR & Session 4	1.145***	0.122
MBSR & Session 5	1.032***	0.116
MBSR & Session 6	1.212***	0.114
MBSR & Session 7	0.113	0.152
Intercept	0.099	0.369
Individual random effects	Yes	
Control variables	Yes	
No. of individuals	138	

Notes: Robust standard errors; $***p < 0.01, **p < 0.05, *p < 0.1$.

Appendix F: The Impact of MBSR on Perceived Stress Score (PSS) and Anxiety Measures, Sample of Individuals Present at Session 6 or 7

Table A.3.7: The Impact of MBSR on PSS and Anxiety, Individuals Present at Session 6 or 7

	[1]		[2]		[3]	
	PSS		Anxiety Now		Anxiety These Days	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
MBSR	1.057	1.109	1.094**	0.467	0.462	0.462
Session 2	0.000	.	0.624	0.382	0.298	0.357
Session 3	0.000	.	1.066***	0.359	0.377	0.304
Session 4	0.000	.	0.213	0.387	-0.049	0.323
Session 5	0.000	.	0.457	0.371	-0.073	0.312
Session 6	1.156**	0.505	0.382	0.359	-0.084	0.282
Session 7	2.321***	0.873	0.839**	0.402	0.387	0.365
MBSR & Session 2	0.000	.	-0.844	0.593	-0.278	0.494
MBSR & Session 3	0.000	.	-0.926	0.595	-0.557	0.517
MBSR & Session 4	0.000	.	-0.153	0.576	-0.131	0.484
MBSR & Session 5	0.000	.	-0.164	0.554	0.130	0.490
MBSR & Session 6	-1.796*	0.951	0.038	0.585	0.184	0.474
MBSR & Session 7	-2.417*	1.344	-0.986	0.690	-0.510	0.604
Intercept	16.162**	7.253	6.546***	2.070	8.145***	2.205
Individual random effects	Yes		Yes		Yes	
Control variables	Yes		Yes		Yes	
No. of individuals	112		112		112	

Notes: Robust standard errors; $***p < 0.01, **p < 0.05, *p < 0.1$.

Appendix G: Summary Statistics of the Stressful Tasks

Table A.3.8: Summary Statistics of the Stressful Tasks

Session 1: Computerised cognitive ability and knowledge test					
	Treatment		Control		Diff
	Mean	SD	Mean	SD	
Task stressful (0-10)	7.373	0.204	7.611	0.166	-0.238
Task not enjoyable (0-10)	6.373	0.284	6.069	0.279	0.304
Task difficult (0-10)	7.179	0.187	6.917	0.192	0.262
Session 6: Publicly performed cognitive ability and knowledge test					
	Treatment		Control		Diff
	Mean	SD	Mean	SD	
Task stressful (0-10)	6.36	0.282	6.746	0.235	-0.386
Task not enjoyable (0-10)	4.68	0.376	5.206	0.307	-0.526
Task difficult (0-10)	5.16	0.272	5.016	0.234	0.144
Session 7: Computerised Stroop test					
	Treatment		Control		Diff
	Mean	SD	Mean	SD	
Task stressful (0-10)	6.524	0.311	6.151	0.226	0.373
Task not enjoyable (0-10)	4.762	0.381	4.623	0.285	0.139
Task difficult (0-10)	4.714	0.296	4.642	0.264	0.073

, **, * indicate significance levels at 10%, 5% and 1% respectively*

Appendix H: The Impact of MBSR on Health-related Behaviours and Eating Habits (measures not reported in main text)

Table A.3.9: The Impact of MBSR on Health-related Behaviours

	Average hours sleep/day		Smoking		Freq. alcohol consumption	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
MBSR	-0.028	0.192	0.025	0.120	0.180	0.130
Session 2	0.055	0.347	0.038	0.033	0.052	0.109
Session 3	-0.264**	0.112	-0.049	0.057	0.200*	0.108
Session 4	-0.176*	0.098	-0.049	0.061	0.479***	0.111
Session 5	-0.241**	0.121	-0.081	0.052	0.355***	0.120
Session 6	-0.583***	0.143	-0.002	0.044	-0.195**	0.098
Session 7	0.029	0.487	0.047	0.051	-0.295***	0.084
MBSR S2	1.436	1.205	0.040	0.064	0.076	0.159
MBSR S3	0.049	0.191	0.065	0.071	0.177	0.150
MBSR S4	-0.325*	0.185	0.086	0.073	-0.171	0.167
MBSR S5	-0.214	0.223	0.071	0.070	-0.179	0.182
MBSR S6	0.140	0.198	0.002	0.069	0.096	0.138
MBSR S7	-0.378	0.520	-0.044	0.090	0.072	0.127
Intercept	9.452***	1.578	1.513***	0.490	5.003***	1.009
Ind. RE	Yes		Yes		Yes	
Control var.	Yes		Yes		Yes	
No. of ind.	138		138		138	

, **, * indicate significance levels at 10%, 5% and 1% respectively*

Table A.3.10: Effects on Food Choices (# of low-calorie options chosen, 0 to 10)

	Coeff.	SE
MBSR	0.596	0.714
Session 6	-1.812***	0.420
Session 7	-1.451***	0.471
MBSR & Session 6	-1.008	0.671
MBSR & Session 7	-0.994	0.737
Intercept	3.980	2.482
Random effects	Yes	
Control variables	Yes	
No. of individuals	132	

Notes: Robust standard errors; $***p < 0.01, **p < 0.05, *p < 0.1$.

Table A.3.11: The Impact of MBSR on Eating Habits (measures not reported in main text)

	Avoid fat		Eat high fibre		Eat regular (1 always to 4 never)		Snack, studying (0 never to 4 always)		Eat more while studying (0 to 4)		Eat because bored (0 to 4)		Eat much fast (0 to 4)	
	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
MBSR	-0.04	0.087	-0.082	0.083	0.028	0.14	0.17	0.244	-0.102	0.25	-0.21	0.168	0.161	0.18
Session 2		.		.	-0.086	0.079	-0.145	0.179	-0.441**	0.183	-0.516***	0.144	-0.197*	0.111
Session 3		.		.	-0.166*	0.096	-0.269	0.176	-0.617***	0.179	-0.596***	0.133	-0.373***	0.102
Session 4		.		.	-0.078	0.088	-0.401**	0.157	-0.793***	0.158	-0.758***	0.129	-0.270**	0.118
Session 5		.		.	-0.208**	0.085	-0.325**	0.158	-0.702***	0.182	-0.752***	0.135	-0.294**	0.119
Session 6	0.079*	0.048	-0.046	0.058	-0.183*	0.098	-0.125	0.167	-0.432**	0.171	-0.616***	0.126	-0.216*	0.119
Session 7	-0.005	0.064	-0.1	0.069	-0.051	0.106	.		.		-0.336**	0.144	-0.185	0.136
MBSR S2	.		.	0.062	0.124	0.206	0.269	0.391	0.272	-0.128	0	0.179	-0.133	0.167
MBSR S3		.		.	0.04	0.14	0.101	0.25	0.214	0.257	0.232	0.182	0.015	0.176
MBSR S4		.		.	-0.078	0.136	0.234	0.252	0.385*	0.227	0.241	0.185	0.058	0.168
MBSR S5		.		.	0.143	0.138	-0.149	0.261	0.354	0.254	0.201	0.196	-0.021	0.166
MBSR S6	-0.005	0.067	0.014	0.09	0.145	0.144	-0.101	0.26	0.198	0.246	0.084	0.189	-0.142	0.191
MBSR S7	0.08	0.08	0.135	0.093	-0.056	0.177	.		.		0.222	0.222	0.086	0.202
Intercept	1.848***	0.455	1.573***	0.412	1.868**	0.9	1.269	0.935	1.485	1.065	0.222	0.962	1.714**	0.723
Ind. RE	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Control var.	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
No. of ind.	138		138		138		134		134		138		138	

, **, * indicate significance levels at 10%, 5% and 1% respectively*