A broader economic evaluative space for public health interventions: an integrated approach

Willings Botha

Thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy (PhD)

Health Economics and Health Technology Assessment

Institute of Health & Wellbeing

University of Glasgow

January 2017

Abstract

Introduction: There is an increasing need for economic evaluation of public health interventions to ensure efficient allocation of resources. Outcomes of such interventions often consists of health and non-health and do not fit in the conventional economic evaluation of quality-adjusted life year (QALY) framework. A cost-benefit analysis (CBA) could be appropriate but has concerns of assigning monetary values to health outcomes. Questions remain on how to consider the broad outcomes of a public health intervention in an economic evaluation.

Objective: This thesis aimed to develop an integrated approach for an economic evaluation of a public health intervention that combines the standard cost-utility analysis (CUA) for health outcomes with the stated preference discrete choice experiment (SPDCE) approach for non-health outcomes on a single monetary metric.

Methods: A natural experiment of the Woods In and Around Towns (WIAT) study in Scotland was used for empirical analysis. Costs were assessed using a top-down approach based on resources used. A difference-in-differences (DiD) approach was used to establish the impact. A CUA valued the health outcomes in terms of QALYs while a previously developed conceptual model of the WIAT was used to identify the SPDCE attributes and levels for the non-health outcomes. The WIAT study questionnaire was mapped to the SPDCE which generated relative willingness to pay (WTP) values from a general Scottish population. The WTP estimates were applied to the changes or improvements in the attributes and levels resulting from the intervention. A net monetary benefit (NMB) framework was then used to combine the CUA with the SPDCE on the same monetary scale, effectively deriving a CBA.

Results: The WIAT interventions were of low cost despite the base case DiD analysis showing a statistically insignificant effect for interventions. The incremental cost-effective ratios (ICERs) for the interventions revealed that they were cost-effective. The probabilistic sensitivity analysis (PSA) showed that the physical intervention was 73% likely to be cost-effective at WTP of £20,000 and £30,000. The combined physical and social interventions had 74% and 75% likelihood of being cost-effective at WTP of £20,00 and £30,000, respectively. There was a great deal of uncertainty around QALY results. Overall, the integrated approach revealed that the WIAT interventions were cost-beneficial in terms of both health and non-health outcomes.

Conclusion: This thesis has proposed and demonstrated the integrated approach that combines the conventional QALY framework with the SPDCE on a single monetary scale, hence a broader economic evaluative space particularly suitable for an economic evaluation of a public health intervention.

TABLE OF CONTENTS

List of tables

List of figures

Working papers and presentations

- Increasing access to green spaces as a means of improving population health: A discrete choice experiment to explore preferences for woodland attributes
  Willings Botha[1, 3], Emma Mcintosh[1], Richard Mitchell[2, 3], Andrew Briggs[1]

- Cost-utility analysis of the Woods In and Around Towns (WIAT) programme for mental well-being in Scotland
  Willings Botha[1, 3], Aldo Elizalde[2, 3], Richard Mitchell[2, 3], Andrew Briggs[1]

- Using the net monetary benefit framework to develop a broader economic evaluative space for public health interventions
  Willings Botha[1, 3], Richard Mitchell[2, 3], Andrew Briggs[1]

- Using discrete choice experiments (DCEs) within a cost-benefit analysis (CBA) framework to estimate benefits of Woods in and Around Towns (WIAT) project for mental well-being in Scotland
  Willings Botha[1, 3], Richard Mitchell[2, 3], Andrew Briggs[1]
  Presented at the Health Economics Study Group (HESG) UK- Sheffield, January 2013.

[1] University of Glasgow, Health Economics and Health Technology Assessment, 1 Lilybank Gardens G12 8RZ, Glasgow, UK
[2] University of Glasgow, Public Health, 1 Lilybank Gardens G12 8RZ, Glasgow, UK
[3] Centre for Research on Environment, Society, and Health (CRESH), UK

Statement of original authorship

The work contained within this thesis is my own and has not been done in collaboration, except where otherwise stated. No part of this thesis has been submitted to any other University for a degree.

Willings Botha

Acknowledgements

I would like to thank Professor Andrew Briggs and Professor Richard Mitchell for giving me the opportunity to undertake the PhD study and their continuous support and guidance throughout my study.

I also wish to acknowledge the financial support for this PhD study from the Forestry Commission Scotland (FCS).

This thesis has benefited largely from the guidance and comments of numerous people: Professor Emma McIntosh; Professor Catherine Ward-Thompson; colleagues and others too numerous to mention. To you all thank you.

Finally, I would like to thank my family and friends for providing necessary support, encouragement, and distractions to keep me sane throughout my study.

Chapter 1: Scope of the study

1.1 Introduction

Currently, conventional economic evaluation approach in healthcare focuses mainly on, and aims to maximize, health-related outcomes given finite healthcare budgets. The quality-adjusted life year (QALY) is the predominant standard unit of outcome. While this approach is well suited for interventions within healthcare, it has been found to be inadequate or unsuitable for valuing the broad health and non-health outcomes of interventions particularly related to public health (Weatherly et al., 2009; Curtis, 2014; Lawson et al., 2014). There is also an increasing recognition of the methodological challenges in how to deal with both health and non-health related outcomes of public health interventions in standard economic evaluation framework of healthcare. To date, no clear guidance on how to conduct economic evaluations of public health interventions exists (Owen et al., 2011; Payne and Thompson, 2015).

Public health is defined as the art and science of preventing disease, prolonging life and promoting health through the organized efforts of society (WHO, 2014). Three key notions can be highlighted from this definition: prevention; protection; and lastly, promoting health at population-level. Based on this definition, public health interventions can be considered to be a collective social effort to promote health and prevent diseases through population surveillance, regulation of determinants of health and the provision of key health services with an emphasis on prevention (Ruger and Ng, 2014). The focus of public health interventions seems to converge towards attaining the well-being of the wider society rather than only the health of individuals. This implies that potential outcomes of public health interventions are broader with outcomes that go beyond health. Throughout the thesis, health is broadly defined as per World Health Organisation's definition to mean a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity (Huber et al., 2011).

1.2 Overall objective of the thesis

This thesis is methodological. It explores the development of a broader economic evaluative space for public health interventions capable of valuing both health and non-health related outcomes. It uses an existing natural experiment of the Woods In and Around Towns (WIAT) study in Scotland as a case study. The WIAT study is an environmental improvement project with physical and social interventions to enhance access to natural environments in deprived communities in Scotland. The expectation of the programme is that it could result in increased physical activity and improved mental well-being of individuals (Silveirinha de Oliveira et al., 2013). A description of the WIAT case study is presented later in chapter three and detailed information about the wider WIAT study is found in Silveirinha de Oliveira et al. (2013).

This thesis takes advantage of the broad array of outcomes (health and non-health related) of the WIAT study which are examples of outcomes of some public health interventions. Traditionally, economic evaluations in healthcare measure the health-related outcomes of an intervention using the standardised EQ-5D tool which uses questions on five dimensions: mobility, self-care, usual activities, pain/discomfort and depression/anxiety (Edlin et al., 2015). The health outcomes are valued in terms of quality-adjusted life years (QALYs) which combines the health-related quality of life and the duration in a particular health state (quantity of life) (Dolan, 1997). The details on the EQ-5D tool and QALYs are presented in chapter two.

This thesis focuses on the health-related outcomes of the WIAT study as measured by the EQ-5D questionnaire and valued in QALYs. However, the QALY framework is narrow in focus and cannot value the outcomes of an intervention that go beyond health which can lead to undervaluation of the overall outcomes of an intervention with broad outcomes.

The WIAT case study provides an opportunity to explore the development of an approach that considers both health and non-health related outcomes of public health interventions.

1.2.1 Research questions

To be able to develop the broader economic evaluative space for public health interventions using the WIAT case study, the following research questions are explored:

1. What are the costs of resources involved in the delivery of the WIAT study?
2. How are the health outcomes of the WIAT study measured and valued?
3. Is the WIAT study effective in terms of improving the health-related quality of life (HRQoL)?
4. Is the WIAT study cost-effective in terms of the health outcomes?
5. How can the non-health outcomes of the WIAT study be assessed and valued?
6. How can both the health and non-health related outcomes of the WIAT study be considered in an economic evaluation on a single metric scale?

1.3 Structure of the thesis

With the exclusion of this chapter, which provides the scope of the study, this thesis is divided into two parts which build upon one another. Part one aims to provide the theoretical aspects of standard economic evaluations in healthcare; the understanding of the relationship between green spaces and well-being of individuals to facilitate the understanding of the WIAT case study; an overview of economic evaluations of public health interventions and their challenges. It goes on to discuss the approaches suggested by the literature on how to handle these challenges. Lastly, the proposed approach that deals with and combines the broad outcomes of public health interventions is presented.

Part two is the implementation of the proposed approach through an empirical analysis of the WIAT case study. It begins with costing of resources used in the delivery of the WIAT programme followed by the valuation of the health and non-health outcomes. Then, these broad outcomes are combined in the new proposed approach. Finally, a general discussion and conclusion including suggestions for future work is given.

There are nine chapters in total with the first part consisting of chapter two, three, and four. The second part includes chapter five, six, seven, eight and nine. These chapters have been organised as follows:

Chapter one presents the scope of this research; and the overall objective of the thesis. Then, the case study research questions are presented to aid the development of a broader economic evaluative space for public health interventions.

<u>Part one</u>

This part is made up of chapter two, three, and four. Chapter two presents the fundamental theoretical aspects of economic evaluations in healthcare. The chapter begins with a definition of economic evaluation and its importance in healthcare. Then, costing and the types of methods of analysis when undertaking an economic evaluation in healthcare are discussed. A discussion on the role of decision-making, perspectives in economic evaluations and important considerations related to on comparators and study population, time horizon, discount rate and uncertainty is then given. Then a discussion on the welfarist and extra-welfarist viewpoints in economic evaluations together with the viewpoint taken by the proposed approach of this thesis follows. Lastly, the chapter presents a discussion and conclusion.

Chapter three discusses the relationship between nature and the well-being of individuals. It begins by defining nature and presents its positive health effects on the well-being of individuals. The focus is on green spaces which is a subset of nature. Then, the mechanisms behind the link between green spaces and well-being of individuals together with their conceptual framework are presented. The chapter proceeds to discuss evidence on the mechanisms behind the association of green spaces and the well-being of individuals. Then, it looks at the positive health effects of green spaces particularly on mental well-being. Lastly, the chapter introduces the WIAT case study and provides a conclusion.

Chapter four looks at the economic evaluation of public health interventions. The chapter begins by providing an overview of the methodological challenges of conducting an economic evaluation of public health interventions. Then, an overview of the approaches suggested in literature to address these challenges is presented. A presentation of the proposed approach for this thesis; a discussion; and a conclusion of the chapter then follows.

Part two

This second part consists of chapter five, six, seven, eight and nine. Chapter five begins the implementation of the proposed approach for this thesis. It discusses the costing of the resources used in the implementation of the WIAT study. This includes the methods used to assess and value the cost of resources used, and the results of the costing exercise. Then, the chapter provides a discussion and a conclusion to wrap up.

The health-related outcomes of the WIAT study measured by EQ-5D questionnaire are valued in chapter six. This chapter discusses the methods used; data collection and analysis; the valuation and presents the results. This is followed by a discussion and a conclusion.

Chapter seven is a presentation of the assessment and valuation of the non-health related outcomes of the WIAT study using the stated preference discrete choice experiment (SPDCE). It discusses the details of how the SPDCE was undertaken including the identification of attributes and the assignment of levels; the construction of experimental designs for pilot and final surveys; the questionnaire development and administration; and data input and analysis. Following this, the results of the SPDCE are presented. The chapter, then, offers a discussion and a conclusion at the end.

Chapter eight presents the proposed approach for this thesis, referred to as the integrated approach. The chapter discusses the methods used to develop this approach, then the results of the approach are given. This is followed by a discussion and a conclusion.

Finally, chapter nine concludes the thesis by providing an overview of the thesis, and key results of the empirical analysis which helped to develop the broader economic evaluative space which is argued to be particularly suitable for a public health intervention. A discussion on how this study contributes to the existing body of knowledge is also given followed by a brief discussion on the implications for research, policy and decision-making that this study has, in general. The chapter concludes with a summary of the strengths and weaknesses of the study and suggestions for future work.

Part 1: Economic evaluation-theoretical aspects

The first part of this thesis that consists of chapter two, three and four introduces the fundamental theoretical aspects of the standard economic evaluations of healthcare interventions. The aim is to apply the theory to the context of the economic evaluation of the Woods In and Around Towns (WIAT) case study in part two of this thesis. Key issues pertaining to the conduct of economic evaluations of interventions with broad outcomes, especially public health interventions, are highlighted.

Chapter 2: Economic evaluations in healthcare

2.1 Introduction

This chapter begins with a definition of economic evaluation in healthcare followed by a discussion on the importance of undertaking economic evaluations of healthcare interventions. Then, it proceeds to explain how the costs and outcomes of interventions are identified, measured, and valued. Following this, a discussion on the role of decision-making and perspective in economic evaluations of healthcare is provided. The chapter then goes on to discuss important considerations in economic evaluations such as comparators and study population; time horizon; discounting; and uncertainty. Lastly, the unabated debate on the two dominant economic evaluation viewpoints of welfarism and extra-welfarism is presented. Then the position on this debate taken by the methodology being proposed in this thesis is highlighted followed by a general discussion and conclusion which wraps up the chapter.

To begin with, it is essential to explain what healthcare and intervention are. Healthcare is one of the ways of modifying the occurrence and impact of ill health through some course of action referred to as intervention (Rychetnik et al., 2002; Morris et al., 2012). Thus, an intervention is aimed at bringing about change or some identifiable outcomes. In general, an intervention in healthcare has considerable resource allocation decision-making implications. Economic evaluations, therefore, help the decision-making process by identifying the implications of allocating resources to one course of action rather than another, in the face of finite healthcare resources. The next section defines economic evaluation in healthcare.

2.2 Meaning of economic evaluation

Economic evaluation in healthcare is defined as an undertaking that involves comparing the costs and outcomes associated with two or more interventions and choosing the option that is more beneficial than the comparator (Drummond et al., 2005; Elliott and Payne, 2005; Lessard, 2007; Payne and Thompson, 2013). An economic evaluation process begins with the identification, measurement and valuation of both costs and outcomes of the intervention being compared to facilitate a choice decision between the alternative intervention (Drummond et al., 2015).

The definition of economic evaluation presented above, in the strictest sense, implies that two features have to be present: first, the information on both costs and outcomes of an intervention; and second, the comparison of an alternative intervention to allow choices to be made for an option that offers maximum benefits (Drummond et al., 2015).

2.3 The importance of economic evaluations in healthcare

Having defined economic evaluation in healthcare, an important question that can be asked relates to why it is important to undertake economic evaluations. Economic evaluations inform decisions in competing choice situations on how to commit healthcare resources to one use instead of another in order to maximize on the given outcome, subject to some resource constraint and uncertainty (Elliott and Payne, 2005; McIntosh and Luengo-Fernandez, 2006a; Miller, 2009; Luyten et al., 2016).

It is well known that healthcare resources are finite and insufficient to meet all the demand. This means that decisions should be made on how to allocate the limited resources. One approach that is used to help resource allocation decisions in healthcare is the undertaking of economic evaluations (McFarland, 2014; Mason et al., 2016). This way, economic evaluations provide a framework of comparing costs and outcomes of alternative interventions in a systematic, formal, explicit and transparent way so that only options that are more beneficial than the comparator are chosen or at least that information can inform the choice (Griffin et al., 2010; Drummond et al., 2015).

This type of economic evaluation is sometimes referred to as efficiency evaluation and is the one commonly undertaken in healthcare (Cunningham, 2001; Drummond et al., 2005). The motivation behind economic evaluation in healthcare is to maximize the benefits, mainly health-related, from a given healthcare budget (while, arguably, ignoring benefits that go beyond health); to address inequitable access to healthcare; to contain cost and manage healthcare demand; to regulate or negotiate reimbursement prices in healthcare markets; and where applicable, to minimize the value of benefits forgone for choosing a particular allocation of resources over another (the opportunity cost); or where the concept of opportunity cost does not arise, an economic evaluation can ensure that an intervention does more good than harm to individuals (Brousselle and Lessard, 2011; Morris et al., 2012).

In general, formal economic evaluations are recommended in the UK for healthcare interventions, especially those which would exclude resources from other alternative uses within or outside the healthcare system (Sculpher and Price, 2003; NICE, 2013; Richards and Hallberg, 2015). The next section begins with a discussion on costing and its approaches in economic evaluations of healthcare.

2.4 Costing in economic evaluations

While there is considerable literature on the outcome side of economic evaluations of interventions, the cost side which essentially represents resource utilisation, is often neglected despite being a key element in an economic evaluation (McIntosh et al., 2014). This section discusses costing in economic evaluations and highlights the steps involved in its approaches.

There are generally three steps that are undertaken in order to capture all costs that are needed or have produced or are consequent to the outcome of the intervention of interest: cost identification; cost measurement; and cost valuation (Drummond et al., 2005; Evers et al., 2015).

2.4.1 Cost identification

During the cost identification stage, a decision is made as to which costs to include in the economic evaluation. This is normally decided on the basis of a number of factors such as the perspective of the study, its broader impact, and the time horizon (Simoens, 2009). As will be noted in chapter five, the costs for the WIAT case study used in this study were identified to include resources used in terms of Forestry Commission Scotland's time for the delivery of the project and other costs related to physical inputs, community involvement and contractors who were involved in the implementation of the intervention work.

2.4.2 Cost measurement

After identifying the costs, the next step is to measure them. The use of diaries, questionnaires, and databases for recording cost items and activities can facilitate the measurement process. Two approaches are generally used in measuring the costs in healthcare economic evaluation: the top-down approach, also referred to as gross costing or macro costing; and the bottom-up approach, also known as activity-based costing (ABC) or micro costing approach (Chapko et al., 2009; McIntosh et al., 2010; Carey and Stefos, 2011).

With the top-down costing approach, the whole budget is looked at alongside the total cost incurred in an intervention. It comprises high level summaries of cost items, hence fails to capture the smallest details of resource use of an intervention (Oostenbrink et al., 2002; Peng Yu, 2009; Federowicz et al., 2010; Perkins et al., 2015). However, despite this drawback, it is generally preferred because it does not require the assessment of the details of each resource use to generate aggregate cost, hence, easy to use in economic evaluations (Oostenbrink et al., 2002; Olsson, 2011; Jacobs and Barnett, 2016).

On the other hand, the bottom-up costing approach establishes a detailed inventory of resources used and measures each resource item separately without resorting to high level summaries (Peng Yu, 2009; Federowicz et al., 2010; McIntosh et al., 2010; Olsson, 2011). Where activities in an intervention are varied in nature, the 'bottom-up' method captures the details of each activity's resource use through, for example, the average time used or wage rates (Oostenbrink et al., 2002). It requires identifying all underlying activities and resource use items of the intervention, then costs are traced back to these activities or resource use items that generated them (Canby Iv, 1995). The basis of this approach is that an intervention consists of activities which use resources, regardless of whether these resources are fixed, variable, direct or indirect (Dowless, 2007).

As can be noted, the bottom-up approach is more comprehensive because it tries to incorporate all details of resource use and aims for accuracy, hence, is referred to as gold standard in economic evaluations (Federowicz et al., 2010; McIntosh et al., 2010; Jacobs and Barnett, 2016). One interesting characteristic of this approach is that it can be used retrospectively to estimate resource use over a particular time (McIntosh et al., 2010). However, when this method is used retrospectively, it is recommended that the ideal time for resource use recall should be within three months because memory fades away considerably beyond this time (McIntosh et al., 2010).

Despite its comprehensiveness and attention to detail as has been discussed, its drawbacks include being costly to implement, too involving to perform, very restrictive when it comes to generalization of results, and sometimes very difficult to get required costing information from relevant sources (Oostenbrink et al., 2002).

2.4.3 Cost valuation

Having discussed the approaches of measuring costs in an economic evaluation, the final step involves valuation. The process of valuing resource use is based on the concept of opportunity cost or shadow prices (Simoens, 2009). Thus, the opportunity cost represents the cost of using resources for some purpose measured in terms of the value of the next best alternative use. On the other hand, shadow prices represent market prices for similar resources. An example of a shadow price can be a wage or salary used to value lost productivity. To calculate cost values, resources used are multiplied by their unit prices. For example, an activity measured in days multiplied by the wage rate per day or an item multiplied by its unit price.

It is essential to note that the costing processes of identifying, measuring and valuing resources used in an intervention is similar for all the methods of analysis in an economic evaluation described earlier. The next section briefly discusses outcome measurement and valuation in economic evaluations.

2.5 Outcome measurement and methods of analysis

There are many ways that outcomes can be measured and valued in economic evaluations. However, only five ways and methods of analysis are commonly used in healthcare: first, is the cost-minimization approach

(CMA); second, is the cost-consequences analysis (CCA); third, is the cost-effectiveness approach (CEA); fourth, is the cost-utility approach (CUA); and lastly, the cost-benefit approach (CBA) (Briggs and O'Brien, 2001; Cunningham, 2001; Brent, 2003; Homik and Suarez-Almazor, 2004; Lewis, 2004; Brown and Brown, 2005; Drummond et al., 2005; Tan et al., 2006; Lorgelly et al., 2010; Brousselle and Lessard, 2011).

Cost-effectiveness analysis is, in some cases, used as an umbrella term for both CEA and CUA approaches. When this is the case, CUA is considered as a special type of CEA (Edlin et al., 2015). Each economic evaluation method of analysis identifies, measures and values the outcomes of an intervention differently (Lewis, 2004; Drummond et al., 2005; Deber, 2009). This poses a challenge because the choice of a particular method of analysis to use depends on the type of outcome of interest which can be very context specific. The challenge becomes more apparent as discussed later in chapter four when an intervention results in multiple and varied outcomes, as is normally the case with public health interventions. The next section presents a detailed discussion of each of the five methods of analysis in economic evaluations of healthcare.

2.5.1 Cost-minimization analysis (CMA)

This method only measures and compares the costs related to the intervention with those of an alternative intervention with the assumption that the outcomes are similar. This means that CMA approach, essentially, culminates to only measuring costs and not the outcomes resulting in a decision rule solely based on the cost, with the lowest-cost option being preferred (Hailey et al., 2002; Salazar et al., 2007; Davalos et al., 2009).

This method has the advantage being simple to use especially that it does not require the complex task of handling outcomes as they are assumed to be similar to those of alternative interventions (Lewis, 2004; Tan et al., 2006). However, given that CMA does not consider the outcome side of an intervention, it does not qualify to be considered a full economic evaluation technique as per the definition of economic evaluation given by Drummond et al. (2005) . For this reason, it is considered as a partial economic evaluation technique (Cunningham, 2001; Lewis, 2004; Kobelt, 2013).

Furthermore, the CMA can only be used to compare interventions with the same outcomes which in practice, is difficult to demonstrate or rarely exist (Lewis, 2004; Salazar et al., 2007). This has led to the argument that the assumption of similar outcomes in CMA is very unrealistic and ineffective in informing healthcare decision-making. For this reason, CMA should not be used in economic evaluations of healthcare because it is not fit for purpose (Briggs and O'Brien, 2001; Dakin and Wordsworth, 2011).

2.5.2 Cost-consequences analysis (CCA)

Cost-consequences analysis (CCA) approach is another important technique of economic evaluation. This approach does not compare costs with outcomes of an intervention, hence cannot be considered to be an economic evaluation technique in the strict sense like CEA,CUA and CBA (Lorgelly et al., 2010). Nevertheless, it is an important tool in economic evaluations particularly in the context of public health interventions with multiple and varied outcomes. The reason being that it allows listing of multiple and different elements of an economic evaluation under the cost side and the outcomes side in a balance sheet format (Payne and Thompson, 2015). For example, different elements making up the cost of resource utilization can be listed under costs while multiple outcomes of a public health intervention such as the WIAT would include QALYs, increased physical activities, reduced stress level can be listed under the outcomes side. When this approach is used, the onus is on decision makers to choose the relevant information in the CCA table on costs and multiple outcomes to make decisions (Mauskopf et al., 1998; Coast, 2004).

The variety of outcomes presented in CCA is both advantageous and problematic at the same time. The advantage is that decision makers are presented with a comprehensive, wide range of outcomes to choose from for various decisions. However, the problem is that not all data are presented in a single metric, as can be noted from the example given above (Mauskopf et al., 1998). This presents comparability problems. Furthermore, the CCA approach places the burden on decision makers to choose their outcome of interest from a variety of outcomes. For this reason, the CCA has received some criticisms (Payne and Thompson, 2015). While this approach cannot be used to rank interventions, it is easily understood and simple to use, hence more likely to aid decision making (Coast, 2004; Lorgelly et al., 2010).

Given that the CMA and CCA are not considered as a full economic evaluation method of analysis, it, effectively, means that only three of the five methods of analysis qualify to be full economic evaluation techniques in the strictest terms, namely: CEA, CUA and CBA.

2.5.3 Cost-effectiveness analysis (CEA)

This approach compares various options, in which costs are measured in monetary units, then aggregated, and outcomes are expressed in non-monetary natural units (Culyer, 2015). For example, cost per case of a disease prevented; cost per hospitalization avoided because of an immunization programme or cost per quitter for an anti-smoking campaign (Cunningham, 2001; Hailey et al., 2002; Davalos et al., 2009; Lorgelly et al., 2010). Its objective is to identify where more benefit can be produced at the same cost or where the same benefit can be achieved at a lower cost (Edlin et al., 2015).

The CEA offers a restrictive evaluative space to a single outcome in its natural units only. This becomes problematic as it clearly excludes other aspects of outcomes of an intervention and comparisons of relative effectiveness of interventions with different outcomes with different units of measurement cannot be made (Cunningham, 2001; Lewis, 2004; McIntosh, 2006; McIntosh and Luengo-Fernandez, 2006a; Davalos et al., 2009; Lorgelly et al., 2010). Clearly, this method of analysis cannot be used for interventions with multiple and varied outcomes consisting of health and non-health.

Despite the limitations above, the CEA approach is relatively easy to carry out and has the advantage of having the outcomes of an intervention clearly measured with the unit of measurement that is more intuitive and is easily understood by individuals (Lewis, 2004; Polinder et al., 2011; Edlin et al., 2015).

The next section discusses the CUA approach which was developed to address the weaknesses of the CEA approach of not permitting comparisons of relative effectiveness of interventions with different outcomes with different units of measurement (Cunningham, 2001; Lorgelly et al., 2010). The CUA approach is of interest in this thesis as noted in chapter one. It forms part of the approach that this thesis is proposing for valuing the health outcomes aspect of interventions with broad outcomes. For this reason, a more detailed discussion is presented.

2.5.4 Cost-utility analysis (CUA)

The CUA approach compares alternative interventions in which costs are measured in monetary units and outcomes are measured in terms of a utility-based, health-related quality of life (HRQoL) measure such as a quality-adjusted life year (QALY) (Culyer, 2015). It is a form of CEA with the effectiveness of an intervention measured in terms of both quality and quantity of life.

Quality-adjusted life year (QALY)

A QALY as a HRQoL measure is a composite metric that incorporates both quality and quantity of life (mortality and morbidity) by weighting a year of life by the quality of life (utility). This utility is a measure of preference or value that an individual places on a particular health state, usually with the value of 1 representing full or perfect health and 0 representing death (Edlin et al., 2015). Health states considered worse than dead have a negative value. The QALY allows comparison among diverse health outcomes to be made with each other (Howard, 2009).

There exist other alternative utility-based HRQoL measures to QALY which include disability-adjusted life years (DALYs) and Health Years Equivalent (HYE) (Whitehead and Ali, 2010). The DALY was developed in the 1990s as an indicator of the relative impact of illness and injuries on losses of health life years. It is commonly used for international comparisons of disease burden and is recommended by the World Bank and World Health Organisation (WHO) (Edejer, 2003). Its derivation is like that of QALY.

In general, QALYs are by far the most frequently used HRQoL measure in healthcare (Johannesson, 1995; Gray et al., 2010; Kobelt, 2013). As discussed earlier, a QALY measure combines length and quality of life (utilities). These utilities are elicited using direct and indirect preference-based techniques.

Direct QALY elicitation methods

Direct methods of QALY utility elicitation techniques include time-trade off (TTO); visual analogue scale (VAS); and standard gamble (SG) (Lorgelly et al., 2010; Whitehead and Ali, 2010). The TTO method of eliciting QALY utilities presents two alternative scenarios to individuals, thus, between living for the rest of their life in an impaired state and living in full health for a shorter period and asks which scenario they would prefer. The time in full health is varied until individuals are indifferent between the two options, then individuals are asked how much time they would be willing to sacrifice to avoid the impaired health state. The second method, the VAS, is a form of rating scale on a single line with the top labelled the 'best imaginable health' and the bottom of the scale labelled the 'worst imaginable health'. Individuals are asked to indicate where on the scale they consider their health state of interest to be. This method is generally considered weak and full of biases relating to scaling because individuals are usually not keen to place health states at the extremes of the scale. Appendix 1, questionnaire item F17 of the WIAT main study shows the VAS questionnaire.

The third method is the standard gamble which adds risk in the decisions faced by individuals. Two choices are presented; the first one is that of remaining in a particular health state with certainty and the second one is a gamble of either being in full health or risking death with some probability which is varied until the individual is indifferent between certainty and the gamble. The more severe the health state, the greater the risk of death that the individual would be willing to be cured of. Some well-illustrated examples on the direct QALY elicitation methods have been provided elsewhere (see Whitehead and Ali (2010)).

<u>Indirect QALY elicitation methods</u>

The direct preference elicitation approaches can be difficult to undertake, time-consuming and sometimes considered unethical because of the inclusion of "death" in the elicitation process. For these reasons, there has been wide use of indirect elicitation methods also known as 'off-the-shelf' methods such as EQ-5D, SF-6D and the Health Utilities Index (HUI) instruments (Horsman et al., 2003; Walters and Brazier, 2005; McCrone et al., 2009; Weinstein et al., 2009; Whitehead and Ali, 2010).

The EQ-5D approach is the most commonly used instrument of all these indirect methods (Edlin et al., 2015). It uses questions on five dimensions: mobility, self-care, usual activities, pain/discomfort, and depression/anxiety. These dimensions are assessed by a question on a three-level ordinal scale of 1(no problem), 2 (some/slight problem) or 3 (major/severe problem) (Walters and Brazier, 2005; McCrone et al., 2009). Recently, a new version of the EQ-5D with a five-levels ordinal scale of 1(no problem), 2 (some/slight problem), 3 (moderate problem), 4 (major/severe problem), and 5 (extreme problem) has been developed (Oemar and Janssen, 2014). This new version maintains the three old levels with an addition of only two levels (moderate problem and extreme problem). Appendix 1, questionnaire item F12 to F16 of the WIAT main study shows the EQ-5D 3L and Appendix 2 shows the EQ-5D 5L version.

The reason for the additions of the two levels, moderate problem and extreme problem, in the EQ-5D questionnaire is that previous studies have shown that these additional levels would increase reliability and sensitivity of the results while maintaining feasibility and potentially reducing ceiling effects (Oemar and Janssen, 2014).

The EQ-5D tool also has a visual analogue scale (VAS) where respondents record their assessment of overall health on a scale from 100, which represents the best imaginable health to 0, which is the worst imaginable health (Feng et al., 2014). The VAS can be used to derive utilities through mapping onto other preference elicitation techniques such as standard gamble or time trade-off (Brazier et al., 2003; Whitehead and Ali, 2010). However, there are arguments against the use of VAS for deriving utilities for economic evaluations because its questionnaire does not involve any trading-off or sacrifices, hence, there is lack of choice in its elicitation task (Brazier et al., 2003). For this reason, the VAS can only be considered as a second-best approach of deriving utilities for use in economic evaluations compared with the commonly used health state information from the EQ-5D descriptive system of five dimensions.

Despite this concern, it has been argued that, in some cases, the VAS results can complement the results of EQ-5D descriptive system in an economic evaluation because they represent respondents’ view of their own health (Feng et al., 2014; Oemar and Janssen, 2014; Devlin, 2016). In this case, the use of VAS could be useful in addressing the question of whose “values” matter in respondents’ reported outcome measures in an economic evaluation. To date, however, there is no guidance on the extent and how the VAS can complement the EQ-5D descriptive system in an economic evaluation. The EQ-5D descriptive system and the VAS are conceptually different tools. Therefore, it can be problematic to use the results of the VAS to complement the standard EQ-5D descriptive system results.

The advantages of the VAS, however, are that it is broader in focus, simple to use, administer, and score compared with the EQ-5D five-dimension descriptive system (Torrance et al., 2001; Brazier et al., 2003; Feng et al., 2014). Furthermore, unlike the EQ-5D descriptive system, the VAS does not use predetermined value sets of utilities obtained from another representative sample, hence it is considered to be free from external bias unrelated to the concerned sample of respondents (Dolan, 1997; Parkin et al., 2010). Despite these advantages, the VAS is known to have administrative limitations. Respondents sometimes fail to indicate a precise position of their overall health on the VAS scale (Feng et al., 2014).

Given that the VAS is not conventionally used to produce health state utilities for calculating QALYs in economic evaluations, most studies do not report its analysis (Brazier et al., 2003; Parkin and Devlin, 2006; Devlin, 2016). For this reason, the rationale for focusing on the EQ-5D descriptive system for the purposes of estimating QALYs for cost effectiveness analysis becomes much clear. Questions, therefore, remain on the relevance and use of the VAS in economic evaluations.

The responses to the EQ-5D descriptive system result in an index for each dimension. The indices for the EQ-5D dimensions are reported as vectors with 11111(full health) for both versions of the EQ-5D, and 33333 and 55555 for worst health for the 3L and 5L version respectively. These vectors are in turn used to derive utilities from the predetermined tariffs obtained using the time trade-off (TTO) method from the UK population (Dolan, 1997). Subsequently, the utilities are used to calculate a QALY measure when the duration in a particular health state is considered using the specification below:

$$QALY = T \times U \tag{1}$$

Where *T* is time in a particular health state (years), and *U* is utility for that health state.

When QALY estimates are made, they are compared to costs in a form of an incremental cost effectiveness ratio (ICER). The results allow comparisons across interventions to be made through a measure of cost per QALY gained thereby aiding decision-making as to whether an intervention is worth undertaking based on the acceptable willingness to pay for each QALY gained from an intervention (Hailey et al., 2002; Lorgelly et al., 2010).

<u>QALY limitations</u>

Despite the widespread use of the EQ-5D approach to elicit utilities for QALY calculation, there is some evidence that the QALY framework presents limitations in economic evaluations. It is restricted to measuring and maximizing health with total disregard to other outcomes of an intervention which may matter to individuals, especially those that go beyond health (Tsuchiya et al., 2001; Brouwer, 2009; Lorgelly et al., 2010; Devlin and Lorgelly, 2016). This implies that there is great possibility that QALYs underestimate the relative consequences of interventions (Lorgelly et al., 2010).

As this thesis directly relates to public health interventions which, in most cases, consist of health and non-health related outcomes, it is likely that their outcomes are undervalued in economic evaluations because of the restrictive nature of the QALY framework. The other non-health outcomes that matter to individuals or society can only be considered if individual or societal preferences are considered alongside QALYs. For example, the non-health outcomes of the WIAT case study used in this study include changes in individual behaviours related to visits to woodlands, pleasure in views of the woods; the enhanced environment in terms of quality of the woodland environment, safety and maintenance; and the social support for environmental use such as community activities and engagement. The interesting question that remains is how to consider all these non-health outcomes in an economic evaluation. This is what, among other issues, this thesis attempts to address.

The other limitation of a QALY measure relates to interventions that are preventive in nature. For example, outcomes of preventive interventions may take a long time to manifest and a QALY is highly dependent on age and life context which are not constant during that time. This limitation may directly apply to public health interventions (Phillips and Thompson, 1998). Furthermore, a QALY measure has been found to be inadequate in valuing emotional and mental health related interventions (Phillips and Thompson, 1998). The five dimensions of the EQ-5D are known to be incapable of fully capturing complex or severe mental health related problems and evidence as to whether it is fit for purpose in mental health related studies is mixed (Brazier, 2010; Luyten et al., 2016; Shah et al., 2016). This may be attributed to the incapacity of respondents with severe mental health problems to complete the questionnaire.

The next section discusses the cost-benefit analysis (CBA) approach. This approach can address the weaknesses of the CUA approach but is rarely used in healthcare because of its methodological concerns of directly eliciting preferences from individuals or society. Again, more details are presented on this approach because of its capability to capture and value outcomes of an intervention other than health. This is of relevance to this thesis which looks at economic evaluations of public health interventions which mostly have broad outcomes.

2.5.5 Cost-benefit analysis (CBA)

The CBA approach is distinct from the other economic evaluation methods of analysis in that it compares the benefits with the costs of an intervention on a monetary scale (Briggs, 2009; Edlin et al., 2015; Giles, 2015). As such, it is possible to make judgement of whether an intervention is worthwhile within the healthcare sector and or across other sectors of the economy (Edlin et al., 2015). For this reason, it is considered to have a broader focus, hence suitable to deal with allocative efficiency of resources (Donaldson, 1998a; Reed Johnson, 2012; Kobelt, 2013).

To assess the value of outcomes of an intervention, the CBA approach uses willingness to pay (WTP) estimates which are elicited using two methods: the revealed preference method (RP), also known as the market preference method; and the stated preference method (SP) (O'Brien and Viramontes, 1993; Healey and Chisholm, 1999; Lancsar and Louviere, 2008; Albaladejo-Pina and Díaz-Delfa, 2009). These two WTP elicitation methods for the CBA approach are explained in some detail below:

I. The revealed preference method (RP)

The RP method elicits individuals' WTP for a good or service by examining their actual real-life behaviour (Fujiwara and Campbell, 2011). This approach is rarely used in healthcare because of lack of readily available healthcare data in a real-life situation (Viney et al., 2002; Lancsar and Louviere, 2008). One key reason for the absence of real-time healthcare data is the presence of public or private health insurance schemes which, in most cases, obscure the actual market prices for healthcare resources. As a result, individuals are unaware of them. The other reasons for the scarce use of RP methods are: the agency relationship between patients and doctors which causes some bias on observed individual preferences; and the absence of market data on new interventions that are not yet introduced in the market (Lancsar and Louviere, 2008).

Examples of RP methods include hedonic pricing and travel cost methods (Fujiwara and Campbell, 2011). Hedonic pricing method uses the price differential between otherwise similar goods. For example, if there are two identical houses on market, one has a view of the park while the other does not, the house with the view is priced higher than the one without any view. The price differential can reveal information on individuals' WTP for the non-market priced "view of the park". On the other hand, travel cost method, which is mostly used in environmental economics, uses time and travel cost expenses incurred by individuals to visit or access a site to estimate their willingness to pay. This willingness to pay represents the "price" of access to the site.

Despite the usefulness of the RP in revealing individuals' WTP for the non-market priced goods, its limitations stated above render the SP methods to be the most preferred. The SP method elicits WTP values based on hypothetical scenarios (what individuals would do) as opposed to what they are observed to do in real-world (Lancsar and Louviere, 2008). The next section looks at the SP method in detail.

II. The stated preference method (SP)

The SP approach uses specially constructed surveys to elicit individuals' WTP estimates for a particular good or service (Fujiwara and Campbell, 2011). There are two broad categories of the SP method: the first category is the contingent valuation (CV) method which focuses on the valuation of a non-market priced good as whole using direct elicitation techniques of monetary values such as open-ended; bidding game; payment card; and single or double bounded dichotomous survey techniques (Pearce et al., 2002; Bridges et al., 2011). An open-ended survey uses questions like "what is your maximum WTP?"; a bidding game survey uses several rounds of stated preference discrete choice questions or bids followed by an open-ended WTP question; a payment card survey uses visual aid with large number of monetary amounts and respondents tick their WTP amount; and lastly, a single dichotomous survey technique uses 'yes' or 'no' to a single WTP amount and a double-bounded dichotomous choice with a 'yes' or 'no' to a single WTP amount or then a 'yes' or 'no' to a higher or lower WTP amount (Pearce et al., 2002).

The CV method has been criticised for its attempt to directly monetize the outcomes especially when they relate to health or life. This has been considered as an unethical by others and the implication on the ability to pay is considered to discriminate those who cannot afford to pay (McIntosh et al., 1999). These concerns have resulted in the CBA approach to be rarely used in practice rendering the CUA approach to be dominant in economic evaluation of healthcare interventions (Dolan and Edlin, 2002; Briggs, 2009; McIntosh et al., 2010).

The second category of the SP approach is the stated preference discrete choice experiment (SPDCE). This approach indirectly elicits individuals' WTP values of specific attributes of a non-market priced good using the marginal rate of substitution as explained later in this chapter (Pearce et al., 2002; Bridges et al., 2011; Fujiwara and Campbell, 2011). This indirect elicitation of the WTP values addresses the limitations of the CV method related to ethics and ability to pay. More importantly, the SPDCE approach accommodates all types of attributes, health and non-health without being restrictive to any dimension. The SPDCE approach is a relatively new methodological development of the CBA approach and is of interest in this thesis, hence the next section discusses it in more details.

2.6 The stated preference discrete choice experiment (SPDCE)

The SPDCE approach is based on two theories: the Lancaster (1966) theory of economic value; and the McFadden (1974) Random Utility Theory (RUT). The theory of economic value posits that the value of any non-market priced good can be determined from a bundle of its characteristics or attributes rather than its consumption per se (Lancaster, 1966). For example, the total value of a woodland derives from its characteristics rather than its use. Preferences would change if the any of the characteristics are altered. On the other hand, the Random Utility Theory (RUT) proposes that individuals choose goods which give them the highest level of satisfaction (utility) (McFadden, 1974). This behavioural rule of individuals when making choices is commonly referred to as "*utility maximizing behaviour*" (Hensher et al., 2015 p.66).

When the Lancaster (1966) theory is considered together with the RUT of McFadden (1974), it is possible to estimate the value of a good using logistic regression and analyse the choices that individuals make between different bundles of characteristics of a good (Kjær, 2005; Hanley et al., 2006; Ryan et al., 2008a; Mentzakis et al., 2011; Londoño and Ando, 2013).

To be able to do this, the SPDCE approach uses a specially constructed survey which consists of hypothetical scenarios with attributes of a good which vary in terms of their levels. This survey is presented to respondents and extent to which respondents are prepared to trade-off one set of attributes or levels against one another is assessed. When one of the attributes is cost, it is possible to indirectly estimate WTP values of the attributes or levels using the marginal rate of substitution (MRS) (Carlsson, 2011; Greiner et al., 2014; Hensher et al., 2015). The MRS is calculated as the ratio of the statistically significant coefficients of the attributes or levels of interest divided by the negative of the coefficient on the cost. The sum of the willingness to pay for each attribute is the value of the good being evaluated.

Central to the SPDCE approach is its design process. Designing a SPDCE involves five key stages: first, the identification of attributes; second, the assignment of levels to the attributes; third, the development of an experimental design which defines the choice sets that would be presented to respondents; fourth, the development and administration of questionnaires to collect data; and fifth, the data input, analysis and interpretation of responses from the survey (Louviere et al., 2000; Lancsar and Louviere, 2008; Mentzakis et al., 2011; Kløjgaard et al., 2012). The reliability of the SPDCE results is, generally, very much dependent on how these five stages have been conducted. The next section presents the five stages of the SPDCE process.

2.6.1 Identification of attributes

The first stage of a SPDCE is to identify attributes of the good under evaluation that individuals value or consider important. These attributes can either be qualitative or quantitative (Lancsar and Louviere, 2008). Currently, there is generally very little guidance and no prescribed way for undertaking this process (Kjær, 2005; Coast et al., 2012). This lack of guidance and consensus has resulted in poor reporting on how attributes and their associated levels have been developed in various SPDCE studies (Coast et al., 2012).

Despite this, there is some agreement on the importance of qualitative work when identifying the attributes such as using: focus groups; expert interviews; policy documents; scientific literature; pilot studies; working with experts; literature reviews and theoretical arguments; existing outcome measures; professional recommendations; patient surveys and reviews by other people through debriefing and free text commenting; and rating or ranking exercises to determine appropriate attributes (Kjær et al., 2006; Guttmann et al., 2009; Bridges et al., 2011; Coast et al., 2012; Kløjgaard et al., 2012; Hiligsmann et al., 2013; Kragt, 2013). The use of qualitative work could result in identifying meaningful and important attributes (Coast et al., 2012).

When identifying attributes for a SPDCE, it is generally agreed that the identification process can rarely include all important attributes. In this case, it is important to ensure that the most important attributes are included in a way that is meaningful; easy to comprehend; concise; and relevant so that respondents do not ignore them or make assumptions about other excluded attributes (Kjær, 2005; Coast et al., 2012; Kløjgaard et al., 2012; Kehlbacher et al., 2013).

During the attribute identification stage, the question that arises is how to determine the number of attributes to be included in a SPDCE (how many?). Although there is no specific suitable number of attributes stipulated, using many attributes may have a practical implication of increasing cognitive burden to respondents because of the increased tasks. Respondents may use certain forms of behaviours such as cognitive shortcuts (heuristics) and ignore much of the information presented to them which may result in non-trading off of other attributes or levels (Lloyd, 2003; Kehlbacher et al., 2013).

Many SPDCE studies have followed a pragmatic approach of using a rule of thumb of having a maximum of eight attributes although some studies have reported varied numbers to as many as 15 attributes (Kjær, 2005; Alves et al., 2008; Lancsar and Louviere, 2008; Pfarr et al., 2014). The advice, therefore is to have a manageable number of attributes. Following this advice, it is reasonable to reduce the number of attributes when faced with so many. The following reduction techniques can be used: first, combining mutually dependent attributes into one attribute or including one of the mutually dependent attribute in the introductory text of the choice question while leaving the other one in the choice set (Kjær, 2005; Kløjgaard et al., 2012); second, excluding all casually related attributes and including the attribute that depicts the effect as this could not result in omitted variable bias (Kjær, 2005); and third, using statistical data reduction techniques such as factor analysis which seeks to reduce complexity in a set of data and reveals a smaller set of the independent underlying factors within it which enables the discovery of main themes in participants' responses (Coast and Horrocks, 2007; Goetz et al., 2013). Having identified the attributes, the next stage of the SPDCE approach is to assign the levels to the attributes.

2.6.2 Assignment of levels

The starting point for assigning levels to attributes is the current baseline 'status quo' (Hanley et al., 2001; Street and Burgess, 2007; Pfarr et al., 2014). As is the case with attribute identification, the assignment of levels to attributes is also improved by qualitative work to make them appropriate and realistic.

During this stage, the decision on the number of levels to assign to attributes has to be made and can be generally complex (Hensher et al., 2015). There is no need to have the same number of levels for all attributes but it is important to note respondents in a SPDCE survey tend to give more value to attributes with more levels, hence, having the same number of levels to all attributes can minimize this problem (Ratcliffe and Longworth, 2002; Kløjgaard et al., 2012).

The number of attribute-levels can determine the type of effects to estimate in a model such that with two levels, only linear effects can be estimated while more than two levels can allow an estimation of non-linear effects (Kløjgaard et al., 2012; Pfarr et al., 2014; Hensher et al., 2015). It follows, therefore that the more the levels are assigned to attributes, the higher the chance that accurate effects can be estimated. However, too many levels can be problematic as can lead to fatigue effects to occur when respondents evaluate the choice options (Alriksson and Öberg, 2008; Pfarr et al., 2014; Hensher et al., 2015). Many SPDCE studies have limited the number of levels to three or four per attribute (Bridges et al., 2011).

It is also essential to carefully consider the attribute-level ranges in a SPDCE as this has considerable impact on the SPDCE design. Inappropriate level ranges may result in over or underestimated SPDCE results which could be misleading (Kjær, 2005). Studies suggest that wider ranges result in smaller standard errors, hence statistically preferable than narrow ranges although sometimes too wide ranges can be problematic as they can lead to dominant alternatives to govern the SPDCE (Rose and Bliemer, 2008; Rose and Bliemer, 2009; Choicemetrics, 2014). Dominant alternatives in a SPDCE indicate that one alternative is particularly very attractive compared to the other to the extent that no particular useful statistical information is provided (Johnson et al., 2007). When the levels have successfully been assigned to the attributes, the next stage involves the construction of an experimental design.

2.6.3 Experimental design

An experimental design is typically a constructed matrix of values based on some statistical specification. This matrix of values represents attributes and levels and is used to map the attributes and their associated levels into sets of alternatives which respondents choose from in a SPDCE survey

(Rose and Bliemer, 2009; Johnson et al., 2013). The columns of the matrix represent the attributes and alternatives and the rows represent the choice options. The advantage of an experimental design is that it helps address statistical problems of sample size and stipulates the number of choice sets that have to be presented to respondents in order to have reliable parameter estimates (Rose and Bliemer, 2008).

The key question to be addressed is how best to allocate the attributes and levels in a matrix. This is normally done in three steps during the construction of the experimental design: coding of levels; model specification; and determining the experimental design type.

1. <u>Coding of levels</u>

Coding of levels helps to assign the values of attribute-levels in their matrix location in a systematic manner that obeys some pre-determined statistical dimensions such as orthogonality and attribute level balance, which are explained later on in this chapter (Rose and Bliemer, 2008; Hensher et al., 2015). These coded values are replaced by their actual attribute-levels during questionnaire construction.

There are three most common coding methods in an experimental design for a SPDCE: first, is design coding also known as dummy coding (0,1,2,3…); second, is orthogonal coding (also referred to as effects coding) (-1, 1) for two levels, (-1,0,1) for three levels, (-3, -1, 1, 3) for four levels, (-3, -1, 0, 1, 3) for five levels, and (-7, -3, -1, 1, 3, 7) for six levels; and third, is coding according to the actual attribute and level values (Johnson et al., 2007; Rose and Bliemer, 2008; Hensher et al., 2015). As can be noted from the effects coding above, the procedure for undertaking effects coding in case of even numbers of levels is to assign one level a positive value while the second level is assigned the same value but negative whereas when the number of levels is odd, the median value is assigned 0 value (-1, 0, 1). This process results in having the matrix of values that is diagonal with columns (not rows) of all levels in one attribute adding up to zero when orthogonality (the zero correlation of attributes) has been achieved (Johnson et al., 2007; Ryan et al., 2008b; Hensher et al., 2015). For this orthogonality test to work, Hensher et al. (2015) note that, conventionally, only odd numbers are used in effects coding with the exclusion of -5 and 5.

Orthogonality is one of the important characteristic of a good experimental design as will be discussed later in this chapter.

The effects coding is of interest in this thesis because it is considered to be superior to design or dummy coding in that it avoids confounding between base levels of categorical attributes and the constant during the estimation using logistic regression (Bech and Gyrd-Hansen, 2005; Hensher et al., 2015). This benefit offered by effects coding has, however, been questioned and considered immaterial to the overall SPDCE in a recent study (Daly et al., 2016). Daly et al. (2016) argue that the sensitivities to the differences across levels for given attributes and the comparison of those differences are the most important aspects in a SPDCE, and are equivalent independently of the coding scheme used.

However, the possibility of testing orthogonality renders effects coding to be more meaningful in an experimental design and it is for this reason that this thesis uses effects coding.

2. <u>Model specification</u>

Model specification in an experimental design involves the understanding of the specific choice problem that the experimental design is required for. This entails some important considerations.

First, the number of alternatives required for the study should be determined; (for example, Option A; Option B; and an opt-out alternative of choosing ‘none of these’ of the two options). Including an opt-out option allows respondents to choose freely among the options without being forced to make a choice which might better reflect real decision-making (Lancsar and Louviere, 2008; Hoyos, 2010).

Having determined the number of alternatives to include in a SPDCE, the second consideration is to specify the utility function for the SPDCE model. This is the probability that a respondent will choose a particular option which can be determined by an indirect utility function (U) with the deterministic and random components as specified below using an example of woodlands:

$$\text{Utility for a woodland}=U(woodland\ A) = \beta_0 + \beta_1(some\ support) + \beta_2(a\ lot\ of\ support) + \beta_3(time) + \beta_4(averagequality) + \beta_5(goodquality) + \beta_6(someopportunities) + \beta_7(manyopportunities) + \beta_8(cost) + \varepsilon \quad (2)$$

Where ***U*** is the observed utility estimated as a function of the attribute levels, β_0 is the constant and assumed to be zero, $\beta_1 - \beta_8$ are the mean attribute level utility weights (deterministic component) and ε is an error term (unobservable random component). Utility for a 'none of these' option is zero.

This consideration is clearly demonstrated in an empirical analysis of the case study in chapter seven. It has to be noted that the utility function for the option 'none of these' is zero which implies that it is not useful in the model specification (Ryan et al., 2008b). In the event that this option is added in the model or imposed later, the experimental design still maintains its optimality (Street and Burgess, 2007). However, the opt-out option is critical in the analysis of SPDCE responses, with the degree of complexity of the SPDCE, and some socio-economic characteristics of respondents having an influence on the opt-out choice (Boxall et al., 2009).

The third consideration is the type of logistic regression model to use when analysing the data. There are various types of logistic regression models to use including the multinomial logit (MNL); mixed logit (MXL) which allows for random taste variation (heterogeneity); nested logit (NL) which allows more flexible error distributions. These models are explained in detail later in the chapter, however, it is essential to note that it is good practice and recommended to start with MNL model when analysing pilot SPDCE data in order to obtain prior information which could be used to construct an improved experimental design for the final survey (Rose and Bliemer, 2009; Hensher et al., 2015). All this becomes clearer later in this chapter when discussing the types of experimental designs.

The fourth consideration is the determination of whether the alternatives should be labelled or unlabelled (Lancsar and Louviere, 2008; Hoyos, 2010; Doherty et al., 2013). Labelled alternatives have descriptors such as names, locations, policy and others which may convey additional information to respondents beyond the attributes and their levels (Blamey et al., 2000) while unlabelled alternatives share the same attributes (generic) with varied levels. Using an example of woodlands, labelled alternatives can be: Kelvingrove woodland and Ruchill woodland while that of unlabelled alternatives can be: woodland A and woodland B.

It has been found that responses from labelled alternatives offer familiarity with the context of the choices in a SPDCE hence reducing the cognitive burden on respondents although trading-off of attributes and levels could be compromised (Blamey et al., 2000; Carlsson, 2011; Doherty et al., 2013). On the other hand, unlabelled alternatives permit respondents to focus on trading-off the attributes with less attachment to descriptors of the alternatives which allows the estimation of marginal rate of substitution (MRS) (Carlsson, 2011; Greiner et al., 2014; Hensher et al., 2015).

As will become apparent later, the focus of this thesis is on unlabelled alternatives where each alternative shares the same generic attributes to allow the estimation of WTP values through marginal rate of substitution.

The fifth consideration is that of degrees of freedom which are a number of observations in a sample minus the number of independent (linear) constraints ($\beta - parameters$). This is calculated by the rule of thumb: the number of parameters (attributes plus one (Rose and Bliemer, 2009). For example, if there are five attributes in a SPDCE, five plus one which equals six, is the minimum number of choice sets that a SPDCE can have. However, the formula to check the degrees of freedom binds the number of choice sets and is given as S>= K/(J-1) (Choicemetrics, 2014).

Where S is the choice sets, K is the maximum number of parameters including constants (five plus constant equals six, in this example), and J is the unique observation of whether an alternative is chosen or not, which is two not three (thus either making a choice or not).

$$S \geq \frac{6}{(2-1)} = 6 \qquad (3)$$

The most interesting part of the consideration for the degrees of freedom is that it is looked at together with another consideration of attribute-level balance (each attribute level should appear an equal number of times for each attribute) in an experimental design (Rose and Bliemer, 2009).

This attribute level balance consideration has some experimental design implications such that mixing the number of attribute-levels, for example, having 2, 3, and 5 levels for different attributes in a SPDCE may result in a higher number of choice sets (30 choice sets) while 2,4,and 6 levels may result in the minimum of only 12 choice sets (Rose and Bliemer, 2009; Choicemetrics, 2014). The general advice, therefore, is not to mix too many different number of attribute levels or at least have them all in even or odd numbers in order to have a reasonable number of choice sets (Rose and Bliemer, 2009).

3. Experimental design type

Having decided on the above considerations, it is time to decide on the type of the experimental design to use in the SPDCE. There are different options available as below:

a. Full factorial design

The first option is to use a full factorial design where all possible different choice situations and all possible effects (main and interaction) can be estimated (Rose and Bliemer, 2009). The calculation of a full factorial design for a labelled SPDCE is l^{mk} whilst for a generic or unlabelled SPDCE is l^k where k is the number of attributes each with l levels, and m is the number of alternatives (Viney et al., 2005; Lancsar and Louviere, 2008; Hensher et al., 2015). For example, with an unlabelled SPDCE with five attributes, all of them with three levels, a full factorial design results in 243 choice sets ($3^5 = 243$).

Practically, a full-factorial design has the statistical advantage of ensuring that all attributes are not correlated with each other (orthogonality) and that attribute levels occur with the same frequency (attribute level balance) which allows the estimation and testing of all possible main and interaction effects. However, as noted above, a full-factorial design yields many choice sets which can be cumbersome to be evaluated by respondents, hence, they are considered to be unrealistic (Breffle, 2008; Lancsar and Louviere, 2008; Carson and Louviere, 2010; Johnson et al., 2013; Pfarr et al., 2014). Furthermore, gathering data on all choice sets based on full factorial design becomes practically difficult, and or when done, it can again increase respondents' cognitive burden.

For this reason, full factorial designs are rarely of interest. Different strategies to reduce the number of choice sets in a SPDCE are increasingly being used. These include: reducing the number of levels within the design to have only two extreme level range, also known as best-worst scaling

(Flynn et al., 2007); using fractional factorial designs; blocking the design which involves dividing the numerous choice sets into manageable parts; and combining a fractional factorial design with a blocking strategy (Hensher et al., 2015).

While the aim of employing these strategies is to reduce the number of choice sets in an experimental design, recently, it has been shown that the more the choice sets are presented to respondents, the better the results in terms of error variances. This is attributed to the increased learning curve which reduces uncertainty on the part of the respondents (Carlsson et al., 2012; Regier et al., 2014).

b. Fractional factorial designs

Out of the above choice set reduction strategies, the most commonly used strategy is the fractional factorial design which is capable of estimating main effects in the SPDCE model (Louviere et al., 2000; Rose and Bliemer, 2009; Choicemetrics, 2014; Pfarr et al., 2014).

When using fractional factorial designs, it is essential to note that the current practice is to follow a two-staged process: first, to construct an initial design based on the principle of orthogonality (a purely statistical specification that ensures that attributes in the design are not correlated with each other) (Rose and Bliemer, 2009; Can and Alp, 2012; Domínguez-Torreiro, 2014). The initial assumption of an orthogonal design is zero prior information about the strength and or direction of individual preferences (Bliemer et al., 2008; Rose and Bliemer, 2009; Hensher et al., 2015). This design is used for pilot surveys whose results (coefficients) are used as prior information in the second stage.

The second stage involves another design which is known as an optimal or efficient design which uses the prior information obtained in the first stage. This optimal or efficient design goes beyond looking at orthogonality and seeks to optimise the statistical efficiency of the SPDCE model in terms of reducing the standard errors of parameter estimates (Scarpa and Rose, 2008). The reduction of the standard errors in the estimated SPDCE parameters results from the prior information used and has implications on the sample size to be used in a SPDCE as discussed later in this chapter.

c. Other experimental designs

Other approaches when using fractional factorial designs include: first, the Bayesian approach which randomly draws numbers from the prior information values assumed by the researcher. These random numbers are progressively and cumulatively added up in order to generate an optimal experimental design (Bliemer et al., 2008; Breffle, 2008; Choicemetrics, 2014). Second, the use of ad hoc designs which are randomly selected from a full factorial design. As their name suggests, ad hoc designs are generally discouraged because they do not rely on any formal statistical theory, therefore, may be inefficient and poorly conditioned (Breffle, 2008; Carson and Louviere, 2010; de Bekker-Grob et al., 2012).

The SPDCE of this thesis that is presented in chapter seven adopts the current practice of constructing an initial fractional factorial experimental design which is orthogonal for the pilot surveys, followed by an efficient fractional factorial design for the final survey (Rose and Bliemer, 2009).

It is recognised that generating an experimental design which deals with all the considerations discussed earlier can be complex without using any software (Johnson et al., 2013). For this reason, some econometric packages or special SPDCE experimental design programs are commonly used (Carlsson, 2011). Examples include R packages (Aizaki and Aizaki, 2015), and Ngene software (Choicemetrics, 2014). Recently, a Stata module known as DCREATE for creating efficient designs has been developed (Hole, 2015). Ngene software is widely used and it is because of that reason that the SPDCE in this thesis uses it. It has the capability of generating both orthogonal fractional factorial experimental design and efficient fractional factorial experimental designs (Ryan et al., 2012b).

The new ‘Dcreate’ Stata program for constructing experimental designs might be preferable in future compared with Ngene software simply because Ngene software is limited to experimental design only. On the other hand, Stata can be used for both constructing the experimental designs and analysis of SPDCEs. Furthermore, it is also commonly used in the analysis of other economic evaluation techniques in healthcare.

<u>Qualities of a good experimental design</u>

While it is important to ensure that attributes in the initial experimental design have zero correlations (orthogonality), other qualities of a good experimental design as proposed by Huber and Zwerina (1996) include: attribute-level balance which requires that all levels of each attribute should appear with equal frequency across choice sets in order to obtain information about each attribute without prejudice on one another; minimal overlap of attribute-levels which means that the probability of repeated attribute-level within a choice set is minimized in order to provide maximum information about respondents trade-offs; and lastly, utility balance which means that the alternatives in choice sets should be close in utility space for respondents in order for them to have equal chances of being chosen (Kanninen, 2002; Johnson et al., 2007; Breffle, 2008; Lancsar and Louviere, 2008; Carson and Louviere, 2010).

It should be noted that the above qualities are critical in an experimental design. However, in many cases, it is impossible to create a design that satisfies all the four qualities at once as some of them may conflict with each other or indeed one quality may be detrimental to the whole SPDCE (Huber and Zwerina, 1996). For example, while minimal attribute-level overlap is one of the desirable qualities of a good experimental design, on the contrary, its presence in a design can have the advantage of improving response efficiency (less attribute non-attendance) in that it simplifies the respondents’ choice tasks by reducing the number trade-offs of attributes that have to be evaluated in a choice set (Street et al., 2008; Johnson et al., 2013).

The absence of attribute-level overlap in an experimental design implies extremely difficult choice making decisions whereas many overlaps would mean easy choice making decisions. The question that arises is how to determine the acceptable degree of minimal attribute-level overlap given that there is no well-established guidance.

Appendix 8 choice task number 2 shows an example of a choice set without any attribute-level overlap while choice task number 3 shows a choice set with attribute-level overlap on the cost attribute.

To achieve minimal attribute-level overlap in an experimental design, many studies have used an approach known as a fold-over method where original experimental design profiles are paired randomly with their mirror image to ensure minimal overlaps of attribute levels within a choice set. (Louviere et al., 2000; de Bekker-Grob et al., 2012; Johnson et al., 2013). However, this method can be complex to implement in a SPDCE with more than two attribute-levels. Other methods have included the use of rotated designs which creates profiles of alternatives in each choice set by rotating each attribute-level one place to the right or by wrapping around to the start of the sequence (0,1,2,3 profile becomes 1,2,3,0 profile) (Johnson et al., 2013); and the use of a special type of a sequential orthogonal design known as optimal orthogonal in the differences (OOD) design proposed by Burgess and Street (2005).

The OOD design maximizes the differences in the attribute-levels across alternatives hence forcing trading of all attributes in the choice set. It is more suited, therefore, for unlabelled choices where attributes are common across alternatives. This design maintains orthogonality as well as ensuring that attributes that are common across alternatives do not take the same level in the choice set, hence, there is no attribute-level overlap as shown in Appendix 8, choice task 1 (Choicemetrics, 2014).

The advantage of the OOD design is that it can provide much information on trade-offs of the attributes because respondents are forced to trade on all attributes in the choice sets, which would allow the estimation of MRS for WTP values. On the other hand, however, as noted earlier, when there are no attribute-level overlap in a choice set, respondents face an extremely difficult choice making decision task (Johnson et al., 2013). Each attribute in a choice set is set to be different across alternatives which has a drawback of tending to promote some lexicographic choice behaviour which would potentially promote a particularly dominant attribute to govern the SPDCE (Choicemetrics, 2014).

Recently, studies have emerged that have looked at and weighed the disadvantages and advantages of emphasizing the achievement of minimal attribute-level overlap in an experimental design, with the former outweighing the latter (Flynn et al., 2016; Flynn, 2016).

Another example of problems created by the achievement of the good qualities of an experimental design are the limitations presented by having an orthogonal design. It is known that orthogonal designs have the advantage of being easy to construct and allowing independent estimation of each attributes contribution to the variations of the levels because of the zero correlations. However, they present a problem of failing to identify a dominant alternative in choice sets (Rose and Bliemer, 2009; Choicemetrics, 2014). For this reason, some studies have reported against their use (Hoyos, 2010; Pfarr et al., 2014). The example in Appendix 8 depicts a dominant alternative (woodland B).

Perhaps, this justifies the current two-staged practice of constructing experimental designs discussed earlier where the second stage emphasizes on the use of optimal or efficient designs rather than orthogonal designs. Efficient designs do not aim at or emphasize on achieving orthogonality but reducing the standard error of the parameters estimates of the SPDCE through use of prior information obtained from initial orthogonal designs

(Scarpa and Rose, 2008; De Ayala et al., 2012; Rose and Bliemer, 2013; Choicemetrics, 2014; Pfarr et al., 2014). Efficient designs are also known to be relatively robust to incorrect prior information (Flynn et al., 2016). Hence, the recommendation of the current practice of using orthogonal designs for pilot surveys and efficient designs for final surveys is more likely to have less bias or may reveal the cause of bias (Flynn et al., 2016).

Turning back to the Ngene software to create orthogonal experimental designs, it is important to note that there are two types of orthogonal designs in Ngene: sequential and simultaneous designs (Choicemetrics, 2014). The sequential design has orthogonality within each attribute. An OOD design discussed earlier, is an example of a sequential orthogonal design. Simultaneous design, however, has orthogonality within each attribute and across attributes. Sequential designs are more suited for a SPDCE with unlabelled alternatives and result in designs with fewer choice sets (Choicemetrics, 2014). This is the type of the orthogonal design that has been used in the initial SPDCE pilot surveys in this thesis.

As regards optimal or efficient experimental designs, their efficiency or optimality are evaluated by measures such as: D-error (design-error) and B-statistic (Bangdiwala's statistic) (Choicemetrics, 2014). A small D-error indicates that the design is efficient and would enable the estimation of parameters with low standard errors. Efficient designs are potentially cost saving as they require a small sample while being able to offer quality information (Louviere et al., 2008). An optimal design should ideally have an efficiency of 100%. This is practically difficult to achieve and the recommendation is to have a design that is nearly optimal although there is no formal definition of what ‘nearly optimal’ means (Street et al., 2005).

It should be recalled that it is very difficult to achieve all the qualities of a good design. Again, it is essential to note that highly efficient designs are associated with higher levels of difficulty related to the ability of completing choice tasks on the part of respondents, hence, they compromise the robustness of the SPDCE results (Rose and Bliemer, 2009; Flynn et al., 2016; Flynn, 2016). For these reasons, some sacrifices should be made to achieve some desired level of efficiency. The next step after the construction of an experimental design is the development of the questionnaire and its administration.

2.6.4 Questionnaire development and administration

Having constructed the experimental design, its output must be framed into a questionnaire instrument to be presented to a sample of respondents. When developing the questionnaire, the recommendation is to consider including one or two examples of choice sets that are not generated by the experimental design and are not intended to be used for analysis but to act as consistency and reliability tests (Kjær, 2005; Carlsson et al., 2012).

Firstly, a consistency test is a theoretically dominant choice task on attribute-levels which is used to check the rationality of the respondents. Appendix 8 shows an example of a consistency check (choice task number 3 and 15). Secondly, a reliability test is simply a re-insertion of a choice set from the experimental design to somewhere later in the questionnaire. This is shown in Appendix 8 choice task number 11 which is re-inserted as choice task number 19. This is used to check replicability of measurement over time (stability) to ensure generalizability of results.

Removing failures of these tests from the analysis is considered to be inappropriate because of the difficulty in determining the reasons for failure (de Bekker-Grob et al., 2012). Hence, the recommendation is to include them all because the SPDCE models used for analysis have proved to be robust to such failures (Ryan and Gerard, 2003).

<u>SPDCE questionnaire format</u>

With regard to the format of the SPDCE questionnaire, Bennett (1999) recommends having an introduction on the subject of research and the researchers; an explanation of the context of the survey; the importance of participation and confidentiality; and the inclusion of an example of the choice task to help respondents to understand the choice tasks at hand. In addition, respondents should also be told about their time commitment on the survey and where to direct queries, in case of any. Some guidance on how to proceed answering the choices questions should also be given.

It is also important to include supporting questions such as introductory, warm-up and attitudinal questions; debriefing questions; and socioeconomic characteristics (Krupnick and Adamowicz, 2007; Hensher et al., 2015). The responses to these supporting questions do not directly form part of analysis in the logistic regression model because they normally do not vary within a choice but can be added to the model as interaction terms with the attributes (Ryan et al., 2012b). The supporting questions may also provide further insights into the nature of or characteristics of respondents which may help clarify or explain some decision strategies used by respondents when making choices (Pearce et al., 2002; Kjær, 2005; Krupnick and Adamowicz, 2007; Carlsson, 2011; Kreye et al., 2014; Hensher et al., 2015).

Ethical approval

Another important consideration prior to any data collection is seeking ethical approval even when the SPDCE itself does not endanger respondents in any way and or even when the data is anonymised. Since respondents' time is involved, it is considered morally correct to have ethical approval. As discussed in chapter seven, the SPDCE in this thesis sought ethical approval from the University of Glasgow ethics committee as shown in **Appendix 7**.

Sample size

The next step is to decide on the appropriate sampling frame for eliciting preferences and the sample size. To date, there is no consensus or guidance on the appropriate sample size for a SPDCE. There exists no definitive statistical formula to determine the sample size, partly because of many complexities related to the whole undertaking of a SPDCE (Marshall et al., 2010).

Previous studies have shown that sample sizes of 40-100 respondents may be sufficient for reliable statistical analysis (de Bekker-Grob et al., 2013). Orme (2006) proposes a total of 300 respondents for robust quantitative research and a minimum of 200 per group for subgroup analysis (Marshall et al., 2010; Rose and Bliemer, 2013).

It has to be recognized that while a large sample size may provide robust results and give the statistical power of a SPDCE, practically, large sample sizes are costly and difficult to obtain and a poor experimental design may further compromise the ability to retrieve meaningful statistical parameter estimates (Rose and Bliemer, 2009). A small sample size, on the other hand, may lessen the reliability of the parameter estimates.

It is for this reason that efficient designs have the potential benefit of reducing confidence intervals of parameters in a SPDCE model hence permitting the use of reduced sample sizes (Kerr and Sharp, 2009). The argument put forward for use of small sample size when an efficient design has been obtained is that efficient designs result in larger decreases in standard errors than those obtained when a larger sample size is used. The gains of improvements to the standard errors for each additional respondent occurs at a diminishing rate until the effect becomes of little statistical significance on the parameter estimates (Rose and Bliemer, 2009).

The general rule of thumb used for calculating the minimum sample size of a SPDCE is that proposed by Orme (1998) (Johnson et al., 2007; Marshall et al., 2010; Rose and Bliemer, 2013):

$$N \geq 500 \times \frac{L^{max}}{J \times S} \tag{4}$$

Where L^{max} is the largest number of levels for any of the attributes used in the SPDCE study, J is the number of alternatives, and S is the number of choice tasks that each respondent faces. For example, if the SPDCE has three as the maximum number in any of the attributes, and three alternatives (A, B and opt-out), and if the experimental design has 18 choice sets plus two choice sets for reliability and consistency test (20 choice sets in total), then, using the formula (4), the minimum sample size is calculated as:

$$N \geq 500 \times \frac{3}{3 \times 20} = 25$$

This is so far the best guidance in the absence of empirical evidence on SPDCE sample size in healthcare (Marshall et al., 2010). The justification for this sample size calculation is that it would yield observations which would be enough to estimate a SPDCE model (Hensher et al., 2015). The sample size sought for the SPDCE in this thesis was 500. This was considered sufficient for robust results because it is well over and above the recommended set rules of thumb discussed above.

The problem of relying on rules of thumb, however, is that such rules cannot be strictly accurate and reliable (de Bekker-Grob et al., 2015). It is because of this that, recently, de Bekker-Grob et al. (2015) have attempted to develop a general approach of determining the minimum sample size requirement for any SPDCE. This new approach requires information about the significance level; the statistical power; the statistical SPDCE model type to be used; the initial prior information about attribute parameters (coefficients); and the SPDCE design considerations such as the number of choice sets, the number of alternatives per choice set, the number of attributes, and the combination of levels in each choice set. It is still unclear if this new development will be widely adopted.

Piloting

It is recommended to pilot-test the initial questionnaire from an orthogonal design on a relatively small sample in order to obtain prior information which can be used to update the design to an optimal or efficient design for the final study (Rose and Bliemer, 2009). In addition, piloting provides an opportunity to pre-test the questionnaire and gain an indication of its feasibility and the quality of the data obtained (Hoefman et al., 2014).

Questionnaire administration

Moving on now to SPDCE data collection, generally, five main methods exist: face to face interviews; telephone interviews; mailed questionnaires and; internet –based interviews or and a combination of any of them (mixed method-drop-off survey where a questionnaire is mailed prior to a visit by the interviewer or mixed method-mail and telephone survey where the questionnaire is mailed prior to a telephone call by the interviewer) (Pearce et al., 2002; Kjær, 2005).

While face to face interviews generate very high response rates of more than 70% (Pearce et al., 2002); bring the interviewer with the interviewee together; and ensures clarity of the questions to respondents, there is potential of bias where the interviewer may influence choices, and face to face interviews are costly in terms of money and time (Kjær, 2005). It is because of these reasons that they are rarely used for SPDCEs in healthcare.

On the other hand, telephone interviews offer sharing of time and not space but are considered to be cheaper whilst still offering high response rate of about 60-75% (Pearce et al., 2002). The drawback though is that SPDCEs are complex and require an understanding of the scenarios. Therefore, the recommendation is to mail the questionnaire in advance of the telephone interview or use mailed questionnaire method (Kjær, 2005).

The mailed questionnaire method is common in many SPDCE studies including healthcare because they are relatively cheap to administer, and give respondents the flexibility to respond at their convenient time. The major problem with mailed questionnaire method has been low response rate of about 25-50% and sampling bias (Pearce et al., 2002; Kjær, 2005).

Recently, internet-based interviewing has become popular with the proliferation of computer use. SPDCE questionnaires are administered online. This method is considered to be of low cost and simple although some sections of the individuals do not feel confident using computers or may reject their use altogether, and or the use of email and internet may preclude a random sample (Pearce et al., 2002; Kjær, 2005).

The SPDCE in this thesis used an internet-based online survey as discussed in chapter seven. The major advantage of online surveys for SPDCEs is that they are flexible to respondents in terms of response time; they ensure the independent treatment of each choice set presented to respondents at each click of the button so that each choice set is not compared to any other set in the survey; and are relatively quick hence cost saving (Pearce et al., 2002; Hensher et al., 2015).

It is widely acknowledged, however, that online surveys are problematic as they require respondents to be computer literate, which may be a hindrance (Pearce et al., 2002; Kjær, 2005; Shah et al., 2015). However, the benefits may outweigh the disadvantages.

Once the SPDCE data are collected, analysis should take place. The next section discusses the data input, analysis, and interpretation of the SPDCE. This is the final stage in the undertaking of the SPDCE approach.

2.6.5 Data input, analysis, and interpretation

In a SPDCE, the same respondent is presented with several choice-sets to complete at a point in time. In order to be able to perform data analysis, all data is set-up as a panel so that each row of the dataset represents one alternative for one respondent (Long and Freese, 2014). It is advisable to set-up an initial data structure using a dummy data-set prior to obtaining the actual SPDCE data. This enables the researcher to be able to see if the analysis of data would be feasible (Ryan et al., 2008b). During the data setting-up stage, decisions are made on the type of coding to use for the variables for model analysis; how to treat missing and or incorrectly filled responses; what type of choice model to use; and what software to use for analysis (Champ, 2003; Ryan et al., 2008b; Burton et al., 2014).

This is followed by the development of set of rules referred to as codebook for coding the entry of data. As will become apparent in chapter seven, data structure for the SPDCE in this thesis was set-up as a panel, initially in Microsoft Excel, with a view of using Stata version 13 (StataCorp, 2013) for data analysis.

<u>Effects and dummy coding</u>

As during the construction of an experimental design, coding of the SPDCE variables is also required during the analysis stage. Two types of coding can be used: effects and dummy coding (Louviere et al., 2000; Bech and Gyrd-Hansen, 2005; Hensher et al., 2015; Hauber et al., 2016). It is important to note that both dummy and effects coding produce similar results both in terms of model goodness fit and coefficients of the payment vehicle although the estimated coefficients for the categorical variables differ which would result in different WTP values (Hasan-Basri and Karim, 2013; Daly et al., 2016; Hauber et al., 2016). The decision to use one or the other depends on the researcher and on the ease of interpreting the estimates from the model (Louviere et al., 2000).

The limitation of dummy coding, however, is that it confounds the base attribute-levels with the overall attribute-levels especially when a SPDCE has an “opt-out” option included in the analysis and, for this reason, effects coding is preferred (Bech and Gyrd-Hansen, 2005; Mercer and Snook, 2005; Bridges et al., 2011; Hensher et al., 2015; Ryan et al., 2012b). Hence, currently, the recommendation for coding categorical or qualitative attribute-levels is to use effects coding (-1, 1, or 0 rather than just 1, or 0 for dummy variables) with the base level coded as -1. The reason is that effects coding avoids the base attribute-levels being absorbed in the zero of the alternative specific constant (ASC) in the logistic regression model during analysis.

Recently, this advantage offered by effects coding has been criticised and considered immaterial in SPDCEs because confounding at attribute base levels is not a cause for concern in SPDCE. What matters is the comparison of differences across attribute-levels for given attributes which both dummy and effects coding are able to provide (Daly et al., 2016).

It is essential to note that continuous variables are generally modelled with their actual values input, with the estimated parameters interpreted as the value of unit change in that continuous variable.

<u>SPDCE data analysis models</u>

Turning now to the SPDCE data analysis, several software applications are available for the estimation of the SPDCE data. Some of them are NLOGIT; Sawtooth; Biogeme, SAS and Stata (Lancsar et al., 2017). The SPDCE in this thesis was analysed using Stata 13 (StataCorp, 2013). There are also several models that can be used to estimate the respondents' preferences. These models form part of the considerations during the experimental design discussed previously. They include: the multinomial fixed or random effects logit (MNL); the generalized extreme value (GEV); the probit; and the mixed logit (MXL) (Train, 2009). The next section discusses these models in detail.

a. <u>The multinomial logit (MNL) model</u>

According to the ISPOR guideline, the starting point for analysing a SPDCE data is the MNL model, also commonly known as the conditional logit (clogit) (Hauber et al., 2016). It is based on the RUT of McFadden (1974) explained earlier in this chapter (de Bekker-Grob et al., 2012; Mengoni et al., 2013). Throughout this thesis, the terms MNL and clogit are used interchangeably.

The clogit model was originally developed by Luce (1959) with the assumptions that the error term of utility specification for one alternative is unrelated to the error term of utility for another. This assumption is known as the independence from irrelevant alternatives (IIA). The IIA assumption implies that adding or deleting an alternative does not affect the odds among the remaining alternatives (Long and Freese, 2014). This IIA assumption was further developed by McFadden (1974) who showed that the distribution of error term of utility is not correlated over alternatives and that the variance is the same for all alternatives (Train, 2009). This phenomenon is generally referred to as the independently, identically distributed (IID) distribution, sometimes called Gumbel and type 1 extreme value (Train, 2009).

The IIA and the IID assumptions of clogit model are considered to be restrictive because they imply that clogit models can only be used in three situations: 1) when variation in preferences relates to observed characteristics (systematic) and not unobserved characteristics (random); 2) when there can be proportional substitution across the alternatives given the researcher's specification of representative utility; 3) when the researcher needs to capture the dynamics of repeated choices since it assumes that the unobserved factors are independent over time in repeated choices situations (Train, 2009).

Nevertheless, clogit model is the widely used model for analysing SPDCE data (Hensher and Greene, 2003; Rose and Bliemer, 2009; Long and Freese, 2014; Hensher et al., 2015). This is partly because it is easy to interpret and allows the capture of the dynamics of repeated choices since it assumes that the unobserved factors are independent over time in repeated choice situations (Train, 2009; Vojáček and Pecáková, 2010; Long and Freese, 2014).

However, in reality, the clogit model may not hold for some situations because of its independence assumption; some unobserved factors for one alternative may relate to other alternatives, and these unobserved factors may persist over time (Train, 2009). Other considerable limitations are: firstly, it cannot represent random taste variation because it assumes that respondents have the same preferences or that their preferences depend on some observable characteristics; it is restrictive and does not allow any pattern substitution due to the IIA assumption; and it cannot be used with panel data when unobserved factors are correlated over time for each decision maker.

It is because of these drawbacks that variants and extensions of clogit models which are more flexible and less restrictive models have been developed to address some of the limitations of clogit models. Specifically, these models have the capability to account for correlated errors from multiple responses from each individual or heterogeneity in preferences across the sample (Train, 2009; de Bekker-Grob et al., 2012; Hauber et al., 2016). Examples of these models are generalized extreme value (GEV), the probit and the MXL models

b. The generalized extreme value (GEV) model

The GEV models relax the IIA assumption limitation. According to Train (2009), the GEV model is one of the models developed largely to overcome the limitations of the independence assumption of the clogit model. This model is more general and constitutes a large class of models that exhibit a variety of substitution patterns and allows correlation of unobserved factors over alternatives. It becomes a clogit model if this correlation is zero.

c. The Probit model

The probit model is another type of a SPDCE model which assumes that the error term of utility is distributed jointly normal over alternatives and over time. It has the limitation of relying on the normal distribution for all unobserved components of utility which does not suit all situations, and as such can give misleading predictions (Train, 2009).

d. The Mixed logit (MXL) model

This is also known as the random-parameter logit (RPL) model (Hauber et al., 2016). It assumes that the probability of making a choice from alternatives depends on the attributes of the alternatives and individual-specific variations in taste. This means that the MXL model relaxes the assumption that preferences are the same and allows them to vary implying that different respondents may have different preferences (Train, 2009).

The MXL model has the ability to approximate any random utility model and allows the error term of the utility function to follow any distribution without restrictions (McFadden and Train, 2000; Hensher and Greene, 2003; Train, 2009). Other forms of mixed logit models include the latent class (LC) models where each respondent is assumed to belong to a class where preferences vary across but not with classes and have the ability to investigate the probability of belonging to a given group (Train, 2009; Vojáček and Pecáková, 2010; de Bekker-Grob et al., 2012; Mengoni et al., 2013; Hauber et al., 2016).

MXL models are considered to be the most promising state of the art SPDCE models and it is recommended, therefore, to start with the clogit model in a SPDCE pilot in order to obtain the coefficients (prior information) which could be used to improve on the experimental design for the final SPDCE survey which could allow the use of the MXL model (Hensher and Greene, 2003; Rose and Bliemer, 2009).

The use of any of these models for analysis is dependent on the assumptions made about the distribution of the random term ε of the utility model specification. There is no consensus on the best method for analysing a SPDCE which often results in inconsistencies and lack of credible justification.

The International Society for Pharmaco-economics and Outcomes Research (ISPOR) has recently developed a guideline and checklist for the statistical analysis of SPDCE data. Key to the guideline is that there is still no clear consensus on the best method. What the guideline and checklist recommends is to understand the properties of the SPDCE data and the properties of the available methods. Then, it is important to make a justification of a chosen method, describe the analysis in detail and interpret the results of the model (Hauber et al., 2016). The acronym '*ESTIMATE*' provides a checklist of the necessary steps as summarized and explained in the reproduced Table 2-1 below:

ESTIMATE	Recommendation
Estimates	Describe the choice of parameter *estimates* resulting from the model appropriately and completely, including • Whether each variable corresponds to an effects-coded level, a dummy-coded level, or a continuous change in levels • Whether each variable corresponds to a main effect or interaction effect • Whether continuous variables are linear or have an alternative functional form
Stochastic	Describe the *stochastic* properties of the analysis, including • The statistical distributions of parameter estimates • The distribution of parameter estimates across the sample (preference heterogeneity) • The variance of the estimation function, including systematic differences in variance across observations (scale heterogeneity)
Trade-offs	Describe the *trade-offs* that can be inferred from the model, including • The magnitude and direction of the attribute-level coefficients • The relative importance of each attribute over the range of levels included in the experiment • The rate at which respondents are willing to trade off among the attributes (marginal rate of substitution)
Interpretation	Provide *interpretation* of the results taking into account the properties of the statistical model, including • Conclusions that can be drawn directly from the results • Applicability of the sample, including subgroups or segments, to the population of interest • Limitations of the results
Method	Describe the reasons for selecting the statistical analysis *method* used in the analysis, including • Why the method is appropriate for analyzing the data generated by the experiment • Why the method is appropriate for addressing the underlying research question • Why the method was selected over alternative methods
Assumptions	Describe the *assumptions* of the model and the implications of the assumptions for interpreting the results, including • Assumptions about the error distribution • Assumptions about the independence of observations • Assumptions about the functional form of the value function
Transparent	Describe the study in a sufficiently *transparent* way to warrant replication, including descriptions of • The data setup, including handling missing data • The estimation function, including the value function and the statistical analysis method • The software used for estimation
Evaluation	Provide an *evaluation* of the appropriateness of the statistical analysis method to answering the research question, including • The goodness of fit of the model • Sensitivity analysis of the model specification • Consistency of results estimated using different methods

Table 2-1: A checklist for SPDCE data analysis. Source: Hauber et al. (2016)

<u>SPDCE results interpretation</u>

Whatever method of analysis is used, the results can be used to determine: whether the attributes are important through: the statistical significance of their coefficients; the direction of effect as shown by the sign of the estimated coefficients; and the relative importance of the parameter as shown by the size of the estimated coefficient (Ryan et al., 2012b).

This wraps up the discussion of the SPDCE, a relatively new and improved preference elicitation technique for cost-benefit analysis approach. The next section looks at the role of economic evaluations in decision-making.

2.7 Decision-making and economic evaluations

As explained earlier in this chapter, the purpose of conducting an economic evaluation is to provide information that results from a comparison of the costs and outcomes of two or more alternative interventions to aid decision-making in healthcare resource allocation. Two conceptually distinct but simultaneous decisions are made in healthcare: first, whether the new intervention should be adopted given results of an economic evaluation; or second, whether additional evidence is required to support the adoption of an intervention (Claxton et al., 2012). The basis of deciding whether a particular intervention should be adopted given its outcomes compared with the cost of resource use is provided by the decision rules of the method of analysis used in an economic evaluation (Morris et al., 2012).

On the other hand, the basis for deciding whether additional information is required to support the adoption of an intervention is determined through the expected value of perfect information (EVPI) framework (Claxton et al., 2012; Morris et al., 2012). The EVPI is the difference between the expected net-benefit of an intervention when perfect information is available and the existing information that is full of uncertainty (Claxton et al., 2012). The next section looks at the decisions rules employed to determine the adoption of an intervention for the different methods of analysis in economic evaluations of healthcare.

2.7.1 Cost-minimization analysis decision rule

When the cost-minimization analysis approach is used in an economic evaluation, the decision rule is that the intervention with the lowest cost should be adopted. However, as noted earlier, the cost-minimization analysis approach is not a preferred method of analysis in economic evaluations of healthcare because it only considers the cost-side, hence delivers a partial economic evaluation which does not look at the outcome side of the intervention. It has since been considered not to be helpful in aiding resource allocation decision-making (Briggs and O'Brien, 2001; Dakin and Wordsworth, 2011).

2.7.2 Cost-utility analysis decision rule

In terms of the cost-utility analysis, which is considered as a special type of cost-effectiveness analysis, the standard decision rules for considering the adoption of an intervention are expressed in a form of the incremental cost-effectiveness ratio (ICER). The ICER is the ratio of the expected cost difference (incremental costs) over the ratio of the expected health outcomes difference (incremental effect) between the intervention and control groups, for example, as shown below:

$$ICER = \frac{cost_{interv} - cost_{control}}{effect_{interv} - effect_{control}} = \frac{\Delta C}{\Delta E} \tag{5}$$

Where $cost_{interv}$ is the expected mean of the cost of intervention in the intervention group and $cost_{control}$ is the expected mean of the cost of intervention in the control group; while $effect_{interv}$is the expected mean effectiveness in the intervention group and $effect_{control}$ is the expected mean outcome in the control group; and ΔC represent the difference in cost between the two groups; and ΔE represent the difference in effect between the two groups.

The ICER results in four possible outcomes which are better depicted using a graph known as a cost-effectiveness (CE) plane. These outcomes can fall into the four quadrants (North West-NW; South West-SW; North East-NE; and South East-SE) of the CE plane as shown in Figure 2-1 below:

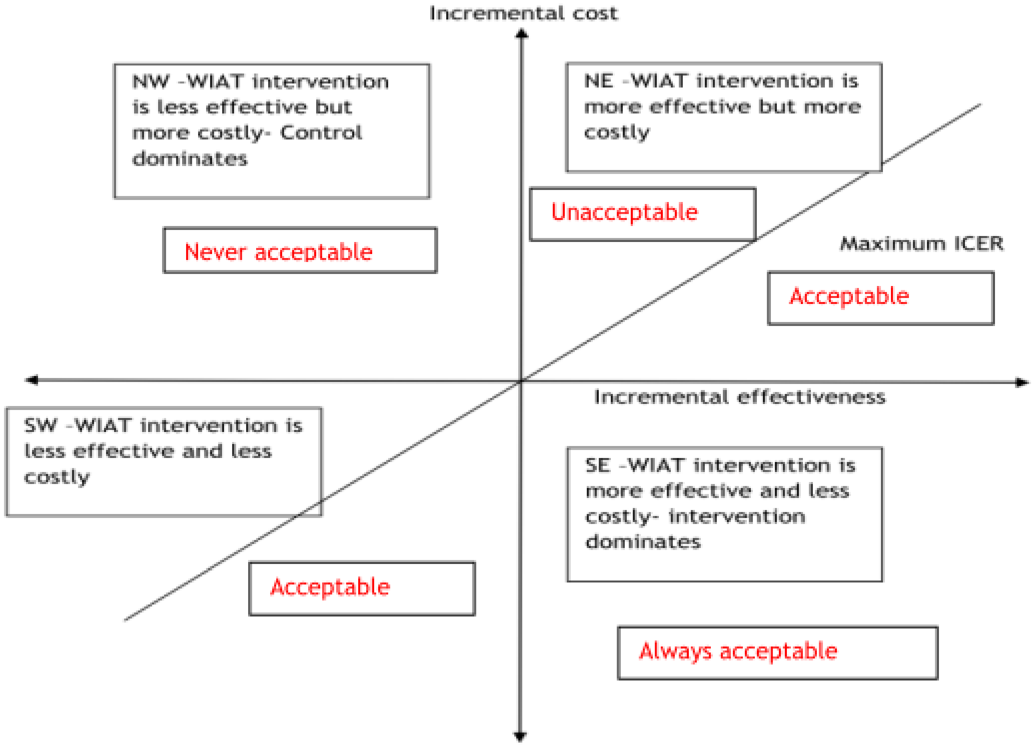


Figure 2-1: Cost-effectiveness plane, adopted from Briggs et al. (2006) and Parkin et al. (2015).

If the ICER for the intervention group compared with the control group falls in the SE quadrant where costs are negative and effects are positive effects, the intervention is considered more effective and less costly, hence cost-effective (dominant, achieving better outcomes at lower cost). If the ICER falls in the NW quadrant where the costs are positive and the effects are negative, the intervention is said to be more costly and less effective and never considered cost-effective (dominated, achieving poorer outcomes at higher cost).

However, if the ICER falls in the NE quadrant with positive costs and positive effects or in the SW quadrant with negative costs and negative effects, the trade-off between costs and effects needs to be examined by comparing to specific thresholds of WTP (λ) (Fenwick et al., 2006).The maximum WTP (λ) threshold is shown as a slope of a line from the origin of the CE plane. The intervention would be considered as cost-effective if the ICER is lower than the WTP threshold $\frac{\Delta C}{\Delta E} < \lambda$ for ICERs in the NE quadrant and higher than the WTP threshold $\frac{\Delta C}{\Delta E} > \lambda$ for ICER in the SW quadrant. In this case, the four quadrants of CE plane are interpreted using a dichotomy. Thus, any intervention falling above the maximum ICER or ceiling ratio for WTP (λ) in Figure 2-1 above is not acceptable.

The ratio approach used in the CE plane, however, presents problems of interpretation. The positive ICERs can belong to either the NE or the SW quadrant and the negative ICERs can be for the NW or SE quadrants such that the ICERs per se are not informative about the cost-effectiveness of an intervention without additional information (Briggs et al., 1997; Briggs and Fenn, 1998; Zethraeus et al., 2003; Fenwick et al., 2004; NICE, 2013).

The other problem presented by ICERs relates to the statistical analysis. When the effect size is zero which when dividing the incremental cost by the incremental effect results to infinity ($\frac{\Delta C}{0}$), hence moving away from normal sampling distribution (Elliott and Payne, 2005; Gray et al., 2010).

<u>Cost-effectiveness acceptability curve (CEAC)</u>

To resolve this problem, cost-effectiveness acceptability curve (CEAC) and the net monetary benefit (NMB) approaches have increasingly been used (Stinnett and Mullahy, 1998). The CEAC approach is a graphical depiction of the proportion of density where the intervention is cost-effective for a range of possible WTP (λ) values for a unit of improvement in health outcomes (Briggs et al., 1997; Fenwick et al., 2004; NICE, 2013). Figure 2-2 below shows an example of the cost-effectiveness acceptability curve:

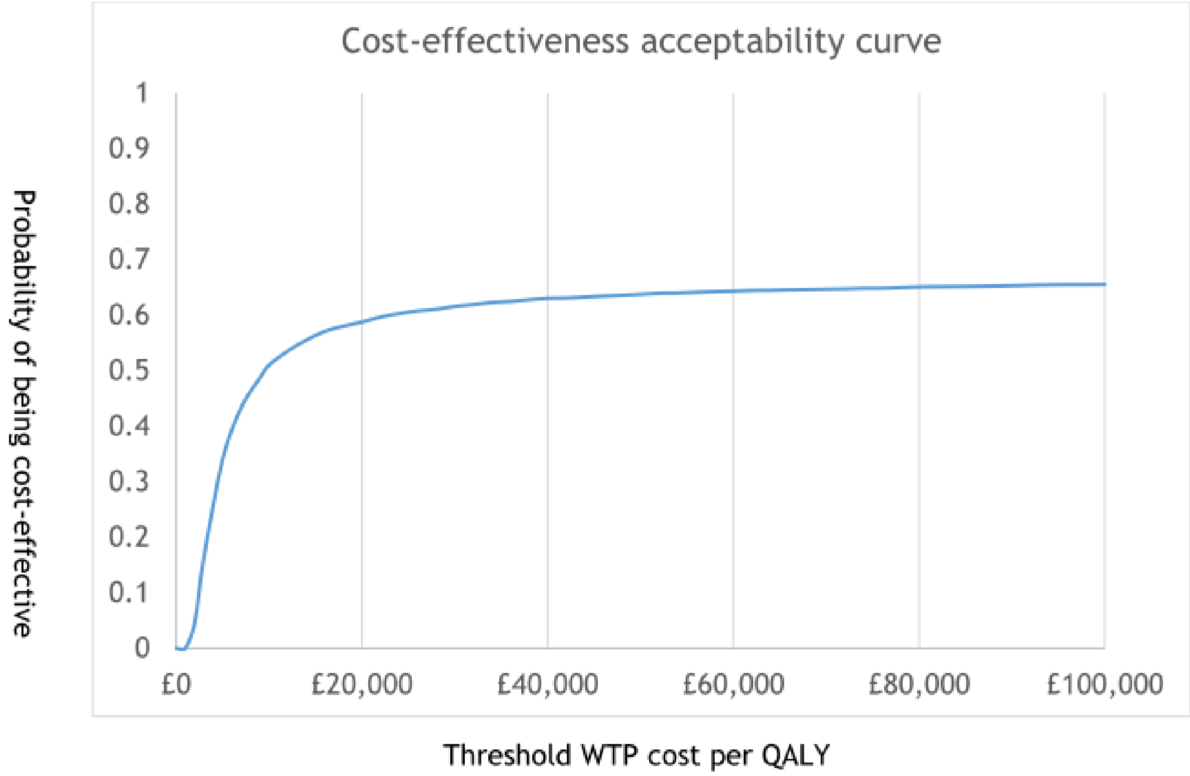


Figure 2-2: Cost-effectiveness acceptability curve example.

The CEAC graphically displays the proportion of the estimates generated by bootstrapping that would be acceptable below or within the threshold of between £20,000 and £30,000, in the case of the UK (NICE, 2013; Ride et al., 2014). This way, the probability of a given intervention being cost-effective is calculated for different levels of WTP (λ) values. The concept of bootstrapping is explained later in this chapter. The CEAC approach represents uncertainty and is an alternative to confidence intervals around ICER and is recommended by NICE (NICE, 2013).

The net monetary benefit (NMB) framework

The other solution to the ICER problems is the use of the net monetary benefit (NMB) approach. The NMB is basically the difference between the monetized incremental effectiveness ($\lambda \times \Delta Effect$) and monetary incremental cost (ΔCost) where λ is the willingness to pay (WTP) or ceiling ratio (Edlin et al., 2015). This is expressed as:

$$\Delta NMB = \Delta Effect \times \lambda - \Delta cost \tag{6}$$

When this is used instead of the ICERs to determine the cost-effectiveness of an intervention, the decision rule is that when the NMB is greater than zero then the intervention is cost-effective while if it is less than zero, then the intervention is not cost-effective (Morris et al., 2012). The drawback of the NMB approach is that it requires the willingness to pay value (λ) to be known or estimated in order to monetize the incremental effects and bring them on the same monetary scale as costs (Edlin et al., 2015). Currently, in the UK, the willingness to pay value (λ) itself is the subject of debate as questions still remain as to how it can be arrived at and whether it is an appropriate range (McCabe et al., 2008; Claxton et al., 2015).

Having discussed the decision rules for cost-utility analysis, a special form of cost-effectiveness analysis, it is acknowledged that decisions on the cost-effectiveness of an intervention are based on existing information about resource use and outcome. There is uncertainty surrounding this information. The adoption of an intervention should not solely rely on this information if there is an opportunity for additional information to support its adoption. The expected value of perfect information (EVPI) approach is used to determine whether additional evidence is required to support the adoption of an intervention following cost-effectiveness analysis and estimate the actual value of obtaining additional information is estimated. The EVPI is simply what remains after the expected net benefit with perfect information is subtracted from the expected net benefit with existing information (Claxton et al., 2012).

The decision rule then becomes that the decision-maker should select the intervention that maximizes the net-benefit for a given value of uncertainty. The EVPI provides the maximum WTP for the additional information to inform healthcare decisions after it is known how uncertainty resolved in model.

2.7.3 Cost-benefit analysis decision rule

When a CBA approach is used in an economic evaluation, the decision rule is that an intervention should be undertaken if the expected total benefits outweigh expected costs, thus, if the net benefit is positive, after propagating uncertainty surrounding these values (Morris et al., 2012; Edlin et al., 2015). This decision rule is based on two theoretical principles discussed earlier, which do not apply to real situations: first, an intervention can only be considered worthwhile if it offers net improvement in the welfare of the society. Thus, when it makes one or more individuals better off while making no individual any worse off; or second: if an intervention allows those who benefit from it to compensate those who become worse off as a result, while still being better off themselves. These principles are known as the actual Pareto improvement principle and the principle of potential Pareto improvement, or the compensation principle, respectively (Coast, 2004; McIntosh et al., 2010).This way, everyone in society will be better off, while if the costs are more than the outcomes, financing such an intervention would inevitably make someone in society worse-off.

In real-life, it is very difficult or practically impossible to observe these two principles because of lack of explicit markets for healthcare. It is for this reason that hypothetical techniques to elicit WTP values for outcomes of an intervention such as SPDCE are commonly used. The next section presents other important considerations in economic evaluations.

2.8 Perspectives in economic evaluations

In an economic evaluation, an intervention may be attractive to some stakeholders while may be unattractive to others depending on many factors. The costs and outcomes of an intervention may affect different sectors of a society in a different way. For example, those who invest in an intervention may not necessarily be the beneficiaries and the impact of an intervention may depend on the method of analysis.

This being the case, it is important to be explicit about the motivation for conducting an economic evaluation and the questions that the evaluation can inform. This is generally referred to as clearly choosing the perspective of an economic evaluation (Morris et al., 2012). The choice of the perspective is normally based on a number of factors, for example, who is funding the intervention, who is going to benefit from or use the results of the intervention (Wonderling et al., 2005). Furthermore, the choice of the perspective dictates which costs and outcome should be evaluated and also helps to make conclusions or some value judgements about the best intervention or the cost-effectiveness of a given intervention (Hoch and Dewa, 2005; Stoto and Cosler, 2005; McIntosh and Luengo-Fernandez, 2006b; Morris et al., 2012).

In an economic evaluation of healthcare, three important perspectives are identified: first, the funder perspective; second, the health perspective; and lastly, the societal or economic perspective (McIntosh et al., 2014).

The funder perspective is that of a decision maker whose objectives are clearly outlined and the costs and outcomes are evaluated in alignment with the objectives. Undoubtedly, conducting an evaluation in alignment with given objectives can be limiting and unsuitable in some cases. For example, this can be problematic in a public health intervention if the decision-maker's objective is to maximize the outcome of an intervention in terms of health in the face of a limited healthcare budget. It can result in underestimating the outcomes of a public health intervention by focusing only on health and ignoring other non-health related outcomes.

Despite this constraint, it is a commonly used perspective in health economic evaluations in the UK with QALYs as a measure of the outcomes of an intervention (NICE, 2013; SMC, 2015). In this case, costs are valued based on the National Health Service (NHS) or personal social services (McIntosh et al., 2014).

On the other hand, the health perspective focuses on the overall long-term consequential impact of an intervention on health. An economic evaluation, in this case, measures and values the health-related outcomes. Like the decision-makers' perspective, the health perspective is restricted to only long health outcomes with a total disregard of other outcomes that go beyond health.

Lastly, is the societal perspective. In contrast with the other perspectives discussed above, the societal perspective takes a broader view which tries to value multidimensional outcomes resulting from a single intervention separately (McIntosh et al., 2014). This perspective is relevant to public health interventions because of the need to capture and value the broad outcomes that consist of health and non-health related. The advantage of using the societal perspective for public health interventions is that it is flexible to allow other methods of analysis to be used such that part of the analysis, especially for the health outcomes could take the decision-makers' or the health perspectives (Walter and Zehetmayr, 2006; Payne and Thompson, 2015). One way of presenting the various perspectives when used in a single economic evaluation could be the use of a cost-consequences analysis (CCA) approach discussed earlier in this chapter.

2.9 Comparators and study population

Once the identification, measurement and valuation of both costs and outcomes have been done, a full economic evaluation in healthcare seeks to make comparisons and a trade-off between two or more alternative interventions in terms of their costs and outcomes using any of the three methods of analysis: CEA; CUA; and CBA discussed earlier in this chapter. The comparative analysis can be between the new intervention versus the standard one or the status quo or a 'do nothing' option. In this thesis, the comparative analysis of the costs and outcomes is between the intervention and control group as presented in the WIAT case study in chapter three.

Turning to the issue of the relevant study population, it is essential to define the relevant population for the intervention (Payne and Thompson, 2015). As regards the WIAT study, the relevant study population was defined to include settlements of 20,472 in the intervention sites (FCS, 2011b). It is essential to note that the target population, in this case, could be any eligible community and this could have implications in the cost-effectiveness of the intervention.

2.10 Time horizon

The other important issue to consider when conducting economic evaluations in healthcare is the time horizon of the intervention over which to track costs and outcomes. Generally, the choice of the time horizon depends on the research question being addressed and can range from a few weeks to several years. What is important is that the time horizon has to be long enough to reflect all the expected costs and outcomes (HAS, 2012). It is essential to recognize that some public health interventions can have inter-generational outcomes such that they need to be followed up for a long period (Park, 2014). This can, sometimes, be infeasible or costly to undertake.

When choosing the time horizon, it should be ensured that both the resource consumption and the chosen outcomes of an intervention are observable in this period. For example, it will be noted during the empirical analysis of the WIAT case study used in this thesis, in chapter five and six, that the time horizon is two years. Thus, between wave one and wave two; and wave two and wave three.

2.11 Discounting

Discounting is another important consideration in economic evaluations of healthcare. It is defined as the adjustment of the costs and outcomes of an intervention in order to reflect three key aspects: firstly, the fact that marginal increases in future consumption are valued less as real incomes increase over time (pure time preference); secondly, the possibility of a catastrophic risk such as death and other adverse events which can curtail the realization of future expected utility; and thirdly, the uncertainty of the future in general (risk) (Morris et al., 2012; Paulden et al., 2016),

Discounting allows all flows of costs and outcomes over time (which tend to occur at different points in time) to be expressed on a common basis in terms of their net present value (NPV). This is achieved using a discount rate.

While there is an agreement that both costs and outcomes should be discounted, establishing a suitable discount rate has proved to be controversial and problematic in healthcare economic evaluations (Drummond et al., 2005; Westra et al., 2012). In order to maintain quality and comparability of health economic evaluations, it is noted that different countries have different discounting guidelines and the actual discount rate differs from country to country (Westra et al., 2012).

Literature on discount rates for economic evaluations in healthcare reveals five different approaches that are undertaken and each approach that is chosen has an impact on the final value of health outcomes of an intervention (Westra et al., 2012). These approaches are: constant (stationary) also known as straight-line; hyperbolic; proportional; stepwise and time-shifted discounting approaches.

A constant discounting approach has both the costs and outcomes discounted at the same rate. The discount rate is, in most cases, determined by the return on risk free government bonds or the real interest rate (normally between 3% and 5%) in line with the society's expectation. For this reason, it is considered to have its basis in welfare economics (Westra et al., 2012). This is the approach that is used in the UK with the discount rate currently at 3.5% for both costs and outcomes (NICE, 2008; NICE, 2013).

Other researchers have suggested that individuals tend to discount the near future higher than the distant future probably because of positive anticipation (Gowdy, 2007). For this reason, the constant method of discounting is considered to underestimate future values. As such, the hyperbolic and proportional discounting approaches aim to reflect this time preference of the society. This time preference is basically the inclination of individuals towards current as opposed to future consumption. The implication of these two approaches is that future benefits are given less weighting at an increasing rate than current benefits.

The step-wise approach, on the other hand, uses a constant rate during a specific period of time and this rate is lowered as time progresses (Westra et al., 2012). Lastly, under time-shifted approach, the discounting is only carried out from the moment a risk reduction caused by an intervention takes place. This approach, however, has the problem of not taking into account the time preference for individuals (Gowdy, 2007; Westra et al., 2012).

2.12 Uncertainty in economic evaluations

While economic evaluations are a useful undertaking in healthcare to aid decision-making, estimates of costs and outcomes are subject to some uncertainty. This uncertainty is normally categorized as: first, stochastic uncertainty, also known as first order uncertainty which relates to the variability in individuals' response to the effects of an intervention despite having the same probabilities and outcomes (heterogeneity) (Polinder et al., 2011; Briggs et al., 2012). For example, individuals in a sample might respond differently to the intervention where some might experience the intervention while others in the same sample might not when exposed to the same intervention. This type of uncertainty can be dealt with through subgroup cost-effectiveness analysis in economic evaluations.

Second, is parameter uncertainty which is also referred to as second order uncertainty. It can arise for two reasons. Firstly, it can come from sampling variation around variable estimates. While the aim of sampling is to achieve representativeness of the population in the sample obtained, chances are that the sample obtained is not representative or has some random errors which can affect the certainty of evidence. This is typically resolved using probabilistic sensitivity analysis (PSA) through the use of 95% confidence intervals, cost-effectiveness planes, cost-effectiveness acceptability curve (CEAC) and expected value of information (EVPI) discussed earlier (Briggs et al., 2012). The second reason for parameter uncertainty can be the lack of consensus about value judgements on the values of some parameters, for example, the appropriate discount rate to be used (Briggs, 2000). When this is the case, a one way or multi-way deterministic sensitivity analysis which is explained later in this section, can be used to resolve it (Polinder et al., 2011; Briggs et al., 2012).

Third, is structural or model uncertainty. As it is well known that economic evaluation models are a simplification of reality. Arguably, there is a possibility that some elements, parameters or characteristics are not included in the model which could result in decision uncertainty (Briggs et al., 2006). Furthermore, the process of identifying the parameters to be used in the model and the choice of or preference for a particular model can lead to different results giving rise to some methodological uncertainty (Edlin et al., 2015). As such, there is need to account for it in order to assess confidence in a chosen course of action and the EVPI approach can help work out the value of collecting additional information to help better decision-making (Briggs et al., 2012).

Having discussed the concept and types of uncertainty surrounding the costs and outcomes estimates in economic evaluations of healthcare, the following section looks at the commonly used approaches of dealing with it.

2.12.1 Dealing with uncertainty

The usual way of dealing with uncertainty in economic evaluations is to undertake a sensitivity analysis which consists of a set of techniques that analyse how sensitive the outcomes of an economic evaluation model are to changes in the model.

According to Edlin et al. (2015), this sensitivity analysis can fall under five categories: one-way, with one parameter in the model that is varied while the rest are held constant; multi-way, where more than one parameters in the model are varied and all others are held constant; threshold analysis, where one parameter in the model is varied to identify the point at which the decision implication change; analysis of extremes, where one or more parameters in the model are set at their lower and upper values to see the impact on predicted costs and outcomes; and lastly, probabilistic sensitivity analysis (PSA) which is a statistical sensitivity analysis that uses probability distributions to obtain credible ranges and the likelihood of any given value being observed.

These approaches to uncertainty can broadly be summarized into two: the deterministic sensitivity analysis when the parameters take a value that is known, and then this known value is changed to explore uncertainty; and the probabilistic sensitivity analysis (PSA) when all the parameters in a model are not known with certainty, the values in the model are replaced by their probability distributions to reflect the expected values and the uncertainty around that expectation for each value.

Probabilistic sensitivity analysis (PSA)

Of interest in this thesis is the PSA. The PSA approach to dealing with uncertainty is considered to be the most predominant approach in economic evaluations of healthcare and is also well known for obtaining credible ranges for parameter values in a model hence recommended (Briggs et al., 2006; Baio and Dawid, 2011; Polinder et al., 2011; Edlin et al., 2015). The PSA is undertaken through Monte Carlo simulations which repeatedly create random data from a mathematically defined probability distribution of parameters with a mean and variance. This is also known as parametric bootstrapping. Any amount of bootstraps above 1000 is considered enough to allow the calculation of 95% confidence intervals from the bootstrapped replicates of data using approaches such as the percentile method which uses the lower and upper percentile (0.025 and 0.975) respectively from the simulations in order to disperse uncertainty in the parameters (Briggs et al., 2006).

Parameter distributions

To conduct a PSA, the first step is to determine the type of distribution to assign to each parameter of the model. The commonly used distributions in economic evaluations are as follows:

First, is the Normal distribution where the average of the data is represented by a mean (μ) with a standard deviation (δ) that describes the amount of variation of the normal distribution. These two values are required to calculate Normal distribution which can be fitted using the method of moments on the basis of the central limit theorem (Edlin et al., 2015). A random parameter on the normal distribution has values between negative and positive infinity ($-\infty$, $+\infty$) (Briggs et al., 2006). This type of distribution can be used when the outcome of interest relates to effectiveness or utilities.

However, a Normal distribution becomes problematic when the properties of a certain parameter differ with those of normal distributions. For example, when a probability parameter is only constrained to extend between zero and one whereas the normal distribution can go beyond one and zero at the extreme. This problem, however, can be dealt with by using a more appropriate distribution such as a Beta distribution (Edlin et al., 2015).

A Beta distribution is another type of distribution used in a PSA. It is a unimodal distribution which lies between zero and one (Edlin et al., 2015). It consists of two parameters: Alpha for the count of events that occur; and Beta for the non-events count (α, β). As with Normal distribution, it can be used when the outcome is about effectiveness or utilities. Despite its use in probability parameters, it is only applicable to binomial probabilities. Where a multinomial parameter is involved, a Dirichlet distribution is used instead (Briggs et al., 2006).

A Dirichlet distribution is a multivariate generalisation of a Beta distribution with the same two parameters (α, β). This distribution arises when the probability of transitioning from one event in a model splits into two or more alternative events. It is generally applied when the outcome of interest is effectiveness. It has to be noted that many software packages including Excel, do not have a Dirichlet function (Edlin et al., 2015). One way to go around this problem is to use Beta distribution in a step by step (sequentially) for each split alternative event to ensure that the total probability equals one. This can be difficult.

Another commonly used distribution is the Gamma distribution. This is generally used for parameters that are continuous, highly skewed and constrained between 0 and ∞. It consists of two parameters: Alpha and Beta (α, β) and is generally used when the outcome is cost or utilities (Briggs, 2000; Edlin et al., 2015).

LogNormal distribution is the other type of distribution that is used when the outcome of interest relates to cost or effectiveness (Briggs, 2000). This is formed by taking the exponential of a Normal distribution hence characterised by a mean (μ) or a standard deviation (δ).

Lastly, other distributions in economic evaluations of healthcare include Weibull, Gompertz and Exponential are also common but often used in survival analysis (Edlin et al., 2015). Edlin et al. (2015) note that a Weibull distribution uses two parameters: Lambda and Kappa (λ, ĸ) and is applied when the outcome of interest is effectiveness. On the other hand, a Gompertz distribution uses two parameters: Gamma and Lambda (ϒ, λ) and is applied when the outcome of interest is effectiveness. And, the exponential distribution, also known as the negative exponential distribution uses Lambda (λ) and is applied when the outcome of interest is effectiveness.

After conducting a PSA which pairs the generated estimates of cost and outcomes, uncertainty around the point estimates is quantified, displayed graphically and analysed using the cost-effectiveness (CE) plane or more intuitively using the cost-effectiveness acceptability curve (CEAC) (Fenwick et al., 2004; Fenwick and Byford, 2005). The CEAC shows the probability of an intervention being cost-effective at given willingness to pay thresholds, for example in the case of the UK, of between £20,000 and £30,000 (Claxton et al., 2005; Baio and Dawid, 2011; NICE, 2013; Ride et al., 2014).

So far, this chapter has presented the theoretical framework of standard economic evaluations in healthcare which included the definition of economic evaluation, its importance, costing, and the different methods of analysis which depend upon the choice of the perspective in an economic evaluation. The next section looks at this connection between the choice of the method of analysis and the perspective in an economic evaluation by examining two viewpoints that exist in health economics: welfarism and extra-welfarism.

2.13 Welfarism versus extra-welfarism

There is a connection between the choice of the method of analysis and the perspective of an economic evaluation which culminates into two broad competing viewpoints in healthcare: the welfarist and extra-welfarist viewpoints (Brouwer, 2009; Morris et al., 2012).

2.13.1 Welfarism

Welfarism is defined as a systematic analysis of the desirability of an intervention, solely in terms of the utility obtained by individuals (Morris et al., 2012). The implication of this definition is that the welfarist viewpoint in economic evaluations has the objective of allocating healthcare resources to interventions that maximize individuals' utility in the face of budget constraints and the sum of the individual utilities constitutes welfare of those individuals or society (Birch and Gafni, 1996; Gyrd-Hansen, 2005; Brouwer, 2009; Buchanan and Wordsworth, 2015).

As such, welfare as a product of utilities of individual members of society is based on the following notions: first, that it is only the individuals themselves that can judge whether their utilities have improved or not (individual sovereignty). This means that judgements made by others about what is good for an individual is irrelevant; and second, that individuals make choices based on preferences to improve or maximize their utility, and that any intervention or policy must be judged on the resulting or consequent effects on their utility. The implication of this notion is that the motivation or intention for the intervention or policy does not matter but outcomes do (consequentialism) (Brouwer, 2009; Morris et al., 2012; Hurley, 2014; Drummond et al., 2015).

It can be argued, however, that the notions presented above cannot be true in the strictest sense or can be irrelevant in healthcare. For example, Morris et al. (2012) highlight the following problems with welfarism: first, the assertion that individuals make rational choices and that are utility maximizers is considered by some as irrelevant in healthcare where third parties such as healthcare providers or experts such as doctors are in a better position to decide on what is best for individuals; second, the assumption that welfare comprises of utility only is considered flawed because other things might matter in the welfare of individuals; third, its basis on individualism can be seen to exclude the role of community values such as making a contribution to some common good which might not necessarily increase one's utility; fourth, the use of utility as a measure of well-being is questionable which has led to the concept of other measures such as capability approach (Coast et al., 2008b), a measure of well-being based on an individual's functioning as discussed in chapter four.

Questions also remain on how to aggregate individuals' utilities especially that the relative desirability of goods depends on the trade-offs people make in reality, which is not at all objective (Morris et al., 2012). To attempt to resolve the problems of welfarism presented above, two theoretical principles of the actual Pareto improvement and the potential Pareto improvement discussed earlier are used. The former posits that an intervention can only be considered worthwhile if it makes one or more individuals better off while making no individual any worse off; and the latter considers an improvement to the social welfare if an intervention allows those who benefit from it to compensate those who become worse off as a result, while still being better off themselves (McIntosh et al., 2010; Morris et al., 2012; Payne and Thompson, 2015).The implication of these principles is that everyone in society will be better off if the outcomes of an intervention outweigh the costs while if the costs are more than the outcomes, financing such an intervention would inevitably make someone worse-off.

It can be noted that central to welfarism is the maximization of individuals' welfare using their utilities, hence linked to the societal perspective because it considers only individual preferences in the economic evaluative space of interventions (Tsuchiya and Williams, 2001; Claxton and Cookson, 2012; Payne and Thompson, 2013). The CBA approach as a method of analysis in economic evaluations has been associated with welfarism. The reason is that it places monetary values on the outcomes of an intervention through willingness to pay of individuals or society elicited from their preferences. The approach is rarely used in healthcare despite having its basis in welfare economic theory. This is because of concerns on the monetization of health outcomes and its direct elicitation techniques of willingness to pay values which imply ability to pay but dominates in other areas of social policy such as environment and transport (Gafni, 2006; Weatherly et al., 2014).

2.13.2 Extra-welfarism

Questions have been asked about the appropriateness and usefulness of welfarism described above which is restricted to individual utilities. This has resulted in an alternative viewpoint known as extra-welfarism (Drummond et al., 2005; Richardson et al., 2005; Brouwer et al., 2008). In contrast to welfarism, extra-welfarism argues that the information on which to base judgement about the results or output of healthcare should be broader than simply utilities of individuals in a society and should be based on the extent an intervention contributes to health itself of individuals (Gyrd-Hansen, 2005).

The focus of extra-welfarism is clearly on maximising health as an outcome of interest in healthcare other than utilities given a finite healthcare budget (Gyrd-Hansen, 2005; Payne et al., 2013; Richards and Hallberg, 2015). For this reason, it is associated with the funder or decision-maker perspective. Cost-effectiveness analysis which includes cost-utility analysis is linked to the concept of extra-welfarism with the outcomes measured in health-related terms and expressed as ratios of incremental costs to incremental outcomes. An intervention is considered efficient if the incremental cost-effective ratio (ICER) which is the additional cost per QALY gained is below or within the range of willingness to pay amount set by the decision maker, which in the UK is between £20,000 and £30,000 (NICE, 2013).

Health, as an outcome of interest in extra-welfarism is considered inadequate for some interventions, especially public health interventions, whose outcomes are multiple and varied, and often go beyond health (Buchanan and Wordsworth, 2015). Furthermore, there is an assumption in extra-welfarism that the health outcome measure of a QALY is the same for all individuals with the same disease or condition which practically means individuals with the same disease or condition are homogeneous (Weinstein, 1988; Weinstein et al., 2009). However, this is highly debatable. For example, in public health interventions, the impact may vary among individuals or groups in terms of addressing inequality. In contrast, welfarism considers variations in utilities which are relative to individual preferences or conditions.

In theory, it can be considered that extra-welfarism is permissive to allow other outcomes other than utilities, hence aims to broaden the economic evaluative space in healthcare (Claxton et al., 2007). In practice, however, questions remain as to how to do it, because extra-welfarism does not prescribe what other things or extra information to consider when undertaking economic evaluations in order to capture other outcomes other than utilities (Brouwer et al., 2008; Brouwer, 2009; Culyer and Cookson, 2012; Hurley, 2014).

Perhaps, the difficulty in incorporating these other outcomes other than utilities in economic evaluations, has led to the continued narrow focus on health as the outcome of interest for extra-welfarism. This is a clear departure from its theoretical broader focus of permitting other outcomes other than utilities (Coast, 2009). As such, it can be concluded that extra-welfarism has simply removed the restrictive outcome space of utility and replaced it with another restrictive outcome space of health (Coast, 2009; Morris et al., 2012; Hurley, 2014).

Another contention regarding extra-welfarism relates to the willingness to pay threshold set by decision-makers to determine the worthiness of an intervention. In the UK, for example, the between £20,000 and £30,000 threshold for a QALY gained from an intervention is arbitrary without any basis. Currently, suggestions are that it should be lowered; be different for subgroups of people or circumstances; or be based on a country's gross domestic product (GDP) per capita (Claxton et al., 2015; Gray and Wilkinson, 2016).

These concerns on extra-welfarism have led to a renewed interest in welfarism, especially on the use of cost-benefit analysis (CBA) method of analysis which is grounded in welfare economic theory. The only concerns of CBA mainly relate to its willingness to pay (WTP) direct elicitation methods which monetise health outcomes and are considered unethical or imply ability to pay. There have been suggestions that the way forward should not be focusing on the limitations of these WTP methods but instead to improve its methodological framework (Gafni, 2006; Schlander, 2010; Reed Johnson, 2012).

Recently, there has been some methodological development on preference elicitation technique used to obtain WTP estimates, known as the stated preference discrete choice experiment (SPDCE) discussed earlier in this chapter. This technique indirectly obtains willingness to pay values from individuals through the marginal rate of substitution (MRS) also discussed earlier. This new technique could possibly render CBA approach credible and perhaps could be widely adopted (McIntosh et al., 2010; de Bekker-Grob et al., 2012).

Currently, the debate on the two viewpoints in economic evaluations remain unabated with some researchers seeking to establish equivalence; and theoretical differences of the two viewpoints; or superiority of one viewpoint over another. The next section discusses these debates and presents the position taken by the proposed methodology in this thesis on these two viewpoints.

2.13.3 Welfarism and extra-welfarism-equivalence

A vigorous debate exists on whether welfarism and extra-welfarism are equivalent (Bala et al., 2002). Numerous studies have attempted to unite these varying viewpoints in economic evaluation of healthcare by comparing CBA and CEA approaches (Phelps and Mushlin, 1991; Johannesson, 1995; Bleichrodt and Quiggin, 1999; Hansen et al., 2004; Kenkel, 2006). The unifying argument put forward is that CBA and CEA can be considered similar mathematically but with different reporting style when looked at from the perspective that only worthwhile programs should be implemented on the basis of cost-benefit ratio rather than just having the benefits that outweigh the costs (Phelps and Mushlin, 1991; Briggs, 2009). The justification of this argument is that if all societal costs are considered in a CEA and a cost per QALY willingness to pay threshold is used, then it becomes a CBA with a monetized QALY (Johannesson, 1995) as exemplified in a net monetary benefit (NMB) framework shown in (6). The decision rule becomes that when NMB is greater than zero, the intervention should be adopted while if it is less than zero, then the intervention does not offer value for money (Morris et al., 2012; Edlin et al., 2015).

The other argument on the equivalence of welfarism and extra-welfarism is that they both use utilities. For example, extra-welfarism uses utilities in the construction of its measure of health of a QALY (Hurley, 2014). This suggests that it does not demonise the use of utilities but argues that utilities on their own are an insufficient basis for value judgement, hence the need for other things alongside utilities as implied in the term “extra”-welfarism (Brouwer et al., 2008; Gray and Wilkinson, 2016). However, as noted earlier, extra-welfarism does not prescribe what these other things are, neither does it offer any guidance on how to treat or include them alongside utilities in economic evaluations.

2.13.4 Welfarism and extra-welfarism-theoretical differences

Other researchers have argued that welfarism and extra-welfarism are theoretically different, irreconcilable and address totally different resource allocation questions in healthcare by looking at CBA and CEA (Donaldson, 1998b; Dolan and Edlin, 2002; Gyrd-Hansen, 2005).

In CBA, an intervention is considered worthwhile if the monetary valuation of outcomes exceeds the costs, hence addresses allocative efficiency (Brazier et al., 2007). This being the case, CBA has an advantage of being broader in scope as a tool for decision making especially for public health interventions, if all the broad outcomes are valued in a monetary metric. For this reason, CBA holds out the promise of allowing comparisons of interventions in terms of the costs and outcomes of an alternative intervention; and a comparison of the magnitude with other interventions or in other sectors to be made rather than just within the healthcare sector because of the monetary metric for both costs and outcomes (Belfield and Levin, 2010).

On the other hand, the CEA focuses on health as a single outcome measure and does not address questions of 'worthiness' of an intervention. Hence, as regards resource allocation, CEA addresses questions of technical efficiency where a specified health gain is produced at the lowest possible cost given the budget constraint. Therefore, it can be considered to be well suited to the task of allocating a fixed budget to competing programmes so as to maximize the chosen effectiveness measure of health (Drummond et al., 2005). On this front, CEA is considered limited in scope and therefore not suitable for evaluating interventions with outcomes other than health (Hall et al., 2004). The argument put across is that CEA does not permit decision makers to say whether healthcare spending is too high or too low but rather shows how any given spending can be arranged to maximize the health outcomes attained by an intervention (Brazier et al., 2007; Gray et al., 2010).

Furthermore, CEA presents comparison problems in contrast with CBA. It only compares interventions that produce similar units of outcomes (Drummond et al., 2005; Gray et al., 2010).

There have also been attempts to establish superiority of welfarism over extra-welfarism by comparing CBA and CEA approaches. Since CBA accords most with welfare economic framework compared with CEA, it has been considered to be a superior and a theoretically sound approach (Dolan and Edlin, 2002; Buchanan and Wordsworth, 2015). On the other hand, CEA is considered not to be consistent with the economic theory but others have considered it to be superior because it is easy to implement and is commonly used (Donaldson, 1998b; Bleichrodt and Quiggin, 1999; Gyrd-Hansen, 2005; Gafni, 2006). The superiority of CEA that arises from ease of use and widespread use rather than having a basis in economics has been questioned.

2.13.5 Welfarism and extra-welfarism-integration

Given the above discussion, rather than attempting to make welfarism and extra-welfarism appear to be mathematically similar or theoretically different or even consider one viewpoint to be superior over another, this thesis proposes a possible way forward. It proposes making the two viewpoints of extra-welfarism and welfarism complement or add value to each other in an economic evaluation. This is especially relevant when undertaking economic evaluations of public health interventions with broader outcomes other than health where the multiple and varied outcomes could be combined on the same monetary scale.

As has been introduced in chapter one, this is possible through what this thesis terms as ‘an integrated approach’. Using the WIAT case study which has broad outcomes like those of a public health intervention, firstly, only those health-related outcomes that are measured by the EQ-5D questionnaire are valued using the widely acceptable CUA associated with extra-welfarism. Then, only those non-health related outcomes that are identified for the WIAT study captured and valued through the SPDCE which is linked to welfarism. Overall results are initially presented in a cost-consequences analysis (CCA) which is not restricted to any viewpoint in economic evaluations. The CCA allows the listing of multiple and varied outcomes without being restricted to a single metric. Subsequently, these broad outcomes (health and non-health) are combined in a complementary manner using the net monetary benefit (NMB) approach in a CBA framework.

The process of mapping the WIAT main study questionnaire items that were considered to measure the non-health outcomes to the attributes and levels of the SPDCE to be able to value the incremental changes or improvements resulting from the intervention has limitations. These are brought about by the WIAT study design and the nature of its data as will be apparent in chapter seven. What is key however, is that the thesis demonstrates that the broadening of the economic evaluative space capable of considering both health and non-health outcomes using an integrated approach is feasible and argues that this is particularly suitable for a public health intervention.

The details and the empirical analysis of the WIAT case study using the integrated approach are described in chapter eight. This way, the extra-welfarist CUA would address the question of achieving the maximum health outcome given the willingness to pay threshold by the decision-maker. At the same time, the welfarist CBA framework that uses the willingness to pay values from individuals through the SPDCE for the incremental improvements or changes in the attributes and attribute-levels resulting from the intervention would reveal whether the intervention is worthwhile from the aspect of the identified non-health related outcomes. The overall results would be assessed using the CBA framework which when the NMB is greater than zero would mean the intervention is cost-effective while if it is less than zero, then the intervention is not cost-effective.

2.14 Summary

This chapter has focused on the theoretical framework of the traditional economic evaluation in healthcare. It began with a discussion on the definition; importance; costing; methods of analysis; and perspectives of economic evaluations. The chapter went further to look at how economic evaluations are used in decision-making; and aspects relating to important considerations in economic evaluations such as comparators and study population definition; time horizon; discounting; and uncertainty. Then, it was necessary to discuss welfarism and extra-welfarism in relation to the proposed integrated approach of this thesis because the quality and usefulness of economic evaluations largely depend on the perspective and method of analysis used.

2.15 Conclusion

This chapter has set out the theoretical grounding for conducting an economic evaluation. Key to the methodology proposed in this thesis is the discussion on welfarism and extra-welfarism. During the discussion of these viewpoints, it became apparent that there was a consensus on the limitations that each method of analysis has. Many studies are attempting to address these limitations, including this thesis, especially for economic evaluations of public health interventions.

The next chapter, introduces the relationship between nature and well-being of individuals. The focus is on the positive effects of green spaces in order to fully understand broad outcomes of the WIAT case study used for empirical analysis in this thesis.

Chapter 3: Nature and well-being of individuals

3.1 Introduction

The aim of this chapter is to help understand the Woods In and Around Towns (WIAT) study which has been used for an economic evaluation empirical analysis in this thesis. The WIAT case study is an environmental improvement intervention aimed at enhancing access to woodlands for individuals to have contact with nature which may result in positive health benefits and other outcomes beyond health (Silveirinha de Oliveira et al., 2013). The identified outcomes of the WIAT study are broad consisting of health and non-health related. These broad outcomes present valuation problems in standard economic evaluation of healthcare because conventional methods of analysis are not capable of considering outcomes beyond health (Weatherly et al., 2009).

To understand the WIAT case study and its objective, this chapter begins by defining nature, and makes a distinction between wild and managed nature. It goes on to discuss how contact with nature may result in the positive health effects on individuals. The mechanisms behind the relationship between contact with nature and good health are explained, supported by evidence from observational and experimental studies. Then, the chapter presents the mixed findings on the association of nature and health. This is followed by a discussion on the positive association of nature and mental well-being of individuals, which is the focus of this chapter. Then, the WIAT case study is introduced in detail, with the conceptual model on its impacts on individuals. Lastly, a summary and conclusion of the chapter are presented.

3.1.1 Meaning of nature

Nature generally refers to the physical features and processes that are not man-made that people ordinarily can perceive which include: flora and fauna; still and running water; qualities of air; and weather and landscapes (Hartig et al., 2014). Nature is usually considered to be synonymous with natural environment or wilderness. It can span both green and blue spaces (large body of water) with varied characteristics and topographies.

It can be noted from the definition of nature given above that there is an implication of non-human interference. However, in practice, nature can be managed to include artificial nature designed by humans in the built environment such as indoor plants, street trees, gardens, and urban parks that are designed, built, regulated and looked after to appear natural (Africa et al., 2014). This is reflected in the definition of nature provided by Bratman et al. (2012) which considers nature as '*areas containing elements of living systems that include plants and non-human animals across a range of scales and degrees of human management, from a small urban park through to relatively "pristine wilderness"'* Bratman et al. (2012 p. 120). This thesis adopts this latter definition of nature to include managed nature.

3.1.2 Contact with nature

There is evidence that suggests that contact with nature may have positive health effects on individuals (Frumkin, 2001; Mitchell and Popham, 2007; Hartig et al., 2014). Contact with nature can happen in different ways: direct contact or views; or indirectly through photographs, films or virtual reality (Hartig et al., 2014). The outcomes from contact with nature vary among individuals in a given population and also across populations and may be produced through various mechanisms discussed later in this chapter (Richardson et al., 2012; Hartig et al., 2014; WHO, 2016).

While nature may confer positive health benefits to individuals through contact, there is also evidence that some factors could hinder or prevent individuals' contact with nature. These include: fear of being attacked; fear of strangers, violence, and kidnapping while out in nature; feelings of insecurity in living environments, especially in very strongly urbanised areas where feelings of insecurity related to criminality are present; fear of being infected by certain diseases which particularly exist in natural environments; exposure to air pollutants and environmental allergens such as pollen; and exposure to natural disasters such as hurricanes and earthquakes (Maas et al., 2009a; Bratman et al., 2012; WHO, 2016). It is suggested that women, older people, and children are more likely to be affected by these factors, hence, are less likely to get contact with nature (Farrall et al., 2000; Jorgensen et al., 2002; Brownlow, 2005; Jansson et al., 2013). Furthermore, it has been found that factors such as access and proximity to natural environments are positively correlated with increased contact with nature (Kaczynski and Henderson, 2007; Sugiyama and Thompson, 2008; Schipperijn et al., 2013). For this reason, poor infrastructure in terms of footpaths, signage, lack of essential facilities and long distant green spaces may prevent easy access and result in less or no contact with nature. Individuals' contact with nature could also be thwarted by the increased demand for natural land which is being released for development to households, firms, and government, as a result, natural environments are disappearing (Barbosa et al., 2007; Choumert, 2010; Vandermeulen et al., 2011). It is essential to note, however, that the evidence on these limiting factors appears to be strong with regards to wild nature (wilderness) compared with managed nature (Susan and Henk, 2012).

Of interest in this chapter, are the positive health effects of the subset of nature in the built environment, generally referred to as urban green spaces. Urban green spaces is a broad term for any “green spaces”, “public open spaces”, or “parks” in an urban setting (Lee and Maheswaran, 2011; PHE, 2014). These terms are often used interchangeably within literature and are loosely used as synonymous in this thesis. It is acknowledged, however, that there may be subtle qualitative differences between them in practice.

Definition of green spaces

There are various definitions of green spaces in published studies. They include areas with natural vegetation such as grass; plants or trees; the built environment features like urban parks; less managed areas like nature reserves; and woodlands and allotments which provide habitat for wildlife but can be used for recreation (Lachowycz and Jones, 2012; Conedera et al., 2015). Scottish Natural Heritage (2008 p. 2) provides a much broader meaning of green spaces as '*any vegetated land or water within or adjoining an urban area*'. This definition encompasses natural and semi-natural habitats; green corridors such as paths, disused railway lines, rivers and canals; amenity grassland, parks and gardens; outdoor sports facilities, playing fields and children's playing areas; cemeteries and allotments; accessible countryside that immediately adjoins a town; and derelict, vacant and contaminated land (Scottish Natural Heritage, 2008).

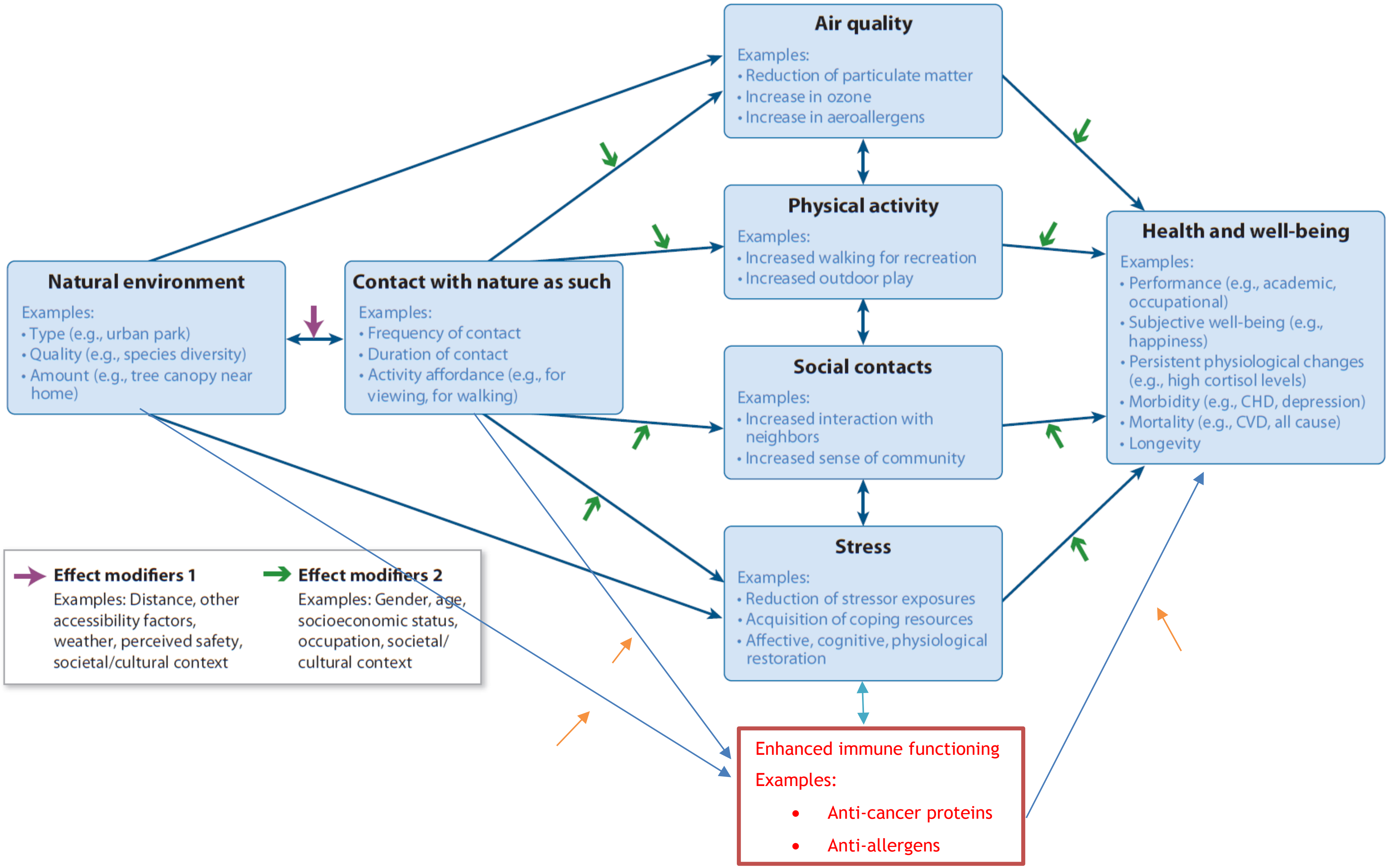
3.1.3 Mechanisms behind green spaces and well-being

Contact with nature may result in health benefits through various mechanisms which may be engaged at the same time and influence one another, may be connected to each other in one way or the other, and may have a synergetic effect to each other depending on the type of green spaces and mode of contact (Townsend, 2006; David et al., 2008; Lee and Maheswaran, 2011; Ward Thompson et al., 2012; Hartig et al., 2014; WHO, 2016). The mechanisms that have received much attention are: improved air quality, greater social interactions emanating from visits to green spaces; increased participation in physical activity; enhanced immune functioning resulting from the positive emotional reactions triggered by contact with nature; and stress reduction (Groenewegen et al., 2012; Lachowycz et al., 2013; Richardson et al., 2013; Hartig et al., 2014; Kuo, 2015; Nieuwenhuijsen et al., 2016). Groenewegen et al. (2006, p.2) have labelled all the positive effects of green spaces that are triggered by these mechanisms as the '*Vitamin G effect'* where *'G'* stands for the green spaces in our surrounding.

3.1.4 Conceptual framework linking green spaces and well-being

Cognizant that there are many mechanisms which point to the link between green spaces and well-being of individuals as discussed above, Hartig et al. (2014) provide a general conceptual framework that shows how contact with green spaces is associated with the well-being of individuals. This conceptual framework includes four mechanisms: improved air quality; increased social interactions; increased participation in physical activities; and stress reduction (Hartig et al., 2014). Recently, enhanced immune functioning has been suggested as another important pathway (Kuo, 2015). For this reason, Hartig et al. (2014) conceptual model has been modified to include enhanced immune functioning as an additional pathway as shown in Figure 3-1 below:

Figure 3-1: Conceptual framework linking nature and health or well-being. Source: Modified from Hartig et al. (2014).



3.2 Evidence on the mechanisms behind green spaces and well-being

Evidence on the mechanisms that link green spaces and well-being of individuals originates from experimental and observational studies with observational studies dominating the evidence base (van den Berg and van den Berg, 2012; Hartig et al., 2014). The next section provides evidence on each of the mechanisms. This evidence is not from a systematic review, hence, not comprehensive. Nevertheless, it suggests that the association of green spaces and well-being of individuals is relatively strong.

3.2.1 Improved quality of air

Human health can be positively affected by green spaces through the ambient quality of the air (Hartig et al., 2014). Green spaces such as trees or other vegetation may absorb or adsorb pollutants in the air including gases and particulate matter, thereby improving the perception of air quality (Nowak et al., 2006; Hartig et al., 2014; Madureira et al., 2015; WHO, 2016). For example, a study in Shenyang, a heavily industrialized city in northeastern China, found that green spaces were associated with reduced levels of air pollutants and reduced atmospheric carbon dioxide (Liu and Li, 2012). Furthermore, a recent study by Madureira et al. (2015) concluded that green spaces improved the perception of air quality of the urban residents in Portugal and France.

3.2.2 Increased social interaction

Social interaction is another important mechanism behind the association of green spaces and well-being of individuals but has not been exhaustively studied. Possibly, because the link between social interaction and well-being is complex and difficult to explore (Hartig et al., 2014). Most evidence comes from observational studies which only reveals associations between green spaces and the protective effects of social relationships on health, such as social inclusion and social opportunities and health behaviours like walking groups (WHO, 2016). For example, green spaces have been found to promote social interactions and a sense of community which may foster health behaviours such as walking groups or community activities, among others, which may result in reducing stress and depression (Swanwick et al., 2003; Kim and Kaplan, 2004; Hordyk et al., 2015). These kind of activities allow people get and stay involved in common spaces informally and can result in social networking, also known as social capital (David et al., 2008; Herzele and Vries, 2011; Lee and Maheswaran, 2011; Xiaolu and Md Masud, 2012). The social capital that is developed through the informal contacts that are facilitated by green spaces can be twofold: bonding social capital; and bridging social capital (Townsend, 2006). The bonding social capital refers to the social connectedness of people with similar social identity. On the other hand, bridging social capital implies mutual relationships across differences in ethnicity, age, class, and social identity. They immensely contribute to the sense of safety and adjustment for individuals (Lee and Maheswaran, 2011).

Evidence from Netherlands by Maas et al. (2009b) suggests that residents in areas with green spaces feel less lonely even when they are not in close contact with neighbours or friends. In addition, quantity and quality of green space has been found to be associated with perceived social cohesion (Herzele and Vries, 2011; de Vries et al., 2013). This is possibly because green spaces can provide the aesthetics of nature which attract people to visit them, and in turn, be in more frequent contact with others and may lead to cohesive communities (Groenewegen et al., 2012).

Recently in Canada, it was found that urban green spaces facilitated social cohesion among newly arrived immigrants who engaged in various activities in the community (Hordyk et al., 2015).

Another positive aspect of green spaces on the social interaction perspective relates to children. It has been found that green spaces offer children an opportunity to interact and widen their social circle with their fellow children and families through facilities such as playfields, parks and others (Townsend, 2006). For this reason, green spaces are considered to offer an opportunity to children to broaden their exposure, develop a sense of diversity, stimulate their ingenuity and imagination, which in turn, may improve their cognitive ability (Xiaolu and Md Masud, 2012). For example, in Zurich, Switzerland, urban green spaces were found to contribute significantly to making cross-cultural contacts and friendships among children and youth (Herzele and Vries, 2011).

3.2.3 Increased physical activities

Participation in physical activities is another possible mechanism linking green spaces and well-being of individuals (Lachowycz and Jones, 2012). This mechanism is well studied but often produces contradictory results. Most evidence is from observational studies (Hartig et al., 2014). For example, evidence from Kaczynski and Henderson (2007) and Coombes et al. (2010) seem to suggest that the closer someone is to green spaces in terms of proximity, the more likely it is they would participate in some type of physical activity in that green space, through walking, for example. Other studies have also demonstrated consistent associations between increased physical activity and some factors of the built environment such as proximity to green spaces and quality of green spaces (Takano et al., 2002; Giles-Corti et al., 2005; Kaczynski and Henderson, 2007; Tzoulas et al., 2007; Sugiyama and Thompson, 2008;

Ward Thompson et al., 2012; Richardson et al., 2013; Schipperijn et al., 2013). Experimental studies using Global Positioning System (GPS) and accelerometer data have also found a positive association between green spaces and moderate to vigorous physical activities (Hillsdon et al., 2006; Almanza et al., 2012).

In contrast with the above findings, It has also been found that the amount of green spaces in the living environment is hardly related to the level of physical activity (Maas et al., 2008). This implies that having more green spaces would not result in increased levels of physical activities among individuals. Furthermore, a systematic review by Lachowycz and Jones (2011) concluded inconsistent and mixed findings across studies on the link between green spaces and physical activities.

This uncertainty, perhaps, stems from the limitation in assessing the type and amount of physical activities that takes place in green spaces; different measures used for green spaces and that green spaces' effects on individuals may be dependent on mediating factors such as gender, age, socioeconomic as well as other variables (Richardson and Mitchell, 2010; Donovan et al., 2011; Herzele and Vries, 2011; Annerstedt et al., 2012; Dadvand et al., 2012; Richardson et al., 2012; Ord et al., 2013; Wheeler et al., 2015).

The possible link between green spaces and well-being of individuals resulting from physical activities could be that the aesthetic characteristics of green spaces attract people outdoors and such outings ordinarily entail some physical activity, usually walking (Greenspace Scotland, 2008; Herzele and Vries, 2011; Groenewegen et al., 2012; Hartig et al., 2014). These physical activities may result in relaxation and reduced stress although causality is unclear (Salmon, 2001). A psycho-biological theory is put forward as a possible explanation of the causality between physical activities and well-being of individuals. This theory posits that physical activity helps the production of endorphins which are the '*brain's feel good'* neurotransmitters and that the movement caused by physical activity helps shed some tensions through the focus placed on a single task, resulting in improved mood, calmness and cleared mind (Salmon, 2001). In addition, it has also been suggested that physical activity could make the brain more resistant to future stressors, thereby maintaining the '*feel good'* status (Schoenfeld et al., 2013).

However, questions remain if at all physical activity is the main mechanism explaining the association between green space and health. For example, Ord et al. (2013) found that the availability of green space in a neighbourhood was not associated with physical activity which led to the suggestion that, perhaps, direct effect of perceiving a natural environment could offer a possible explanation.

Despite these mixed findings on the physical activity as a mechanism behind green spaces and well-being of individuals, the role of physical activity in green spaces to the well-being of individuals is generally acknowledged (David et al., 2008; Bowler et al., 2010; John, 2011; Xiaolu and Md Masud, 2012). For example, Wolf (2010) found that green spaces generally offer de-stressing through physical activity. Recently, it was also found that physical activities in natural environments reduces the risk of poor mental health more than physical exercise in any other environment and that different types of environments may enhance the psychological well-being differently (Mitchell, 2012). Di Nardo et al. (2012) also, found that positive perception of, and access to green spaces may result in 24%

chance of individuals' involvement in physical activities, which may possibly, contribute to the mental well-being of individuals.

3.2.4 Enhanced immune functioning

Enhanced immune functioning is also suggested to contribute to the association of green spaces and well-being of individuals. Some studies have shown that contact with green spaces produces immune responses such as anti-cancer proteins and some anti-allergens in children exposed to allergens from green spaces (Li et al., 2008; Lynch et al., 2014). This has been supported by the "biodiversity hypothesis" which suggests that reduced contact with nature and biodiversity in general may have negative impacts on the health of individuals, particularly on immunity to allergies and chronic inflammatory diseases (Hanski et al., 2012).

Recently, another study also points the link of green spaces and well-being of individuals to the immune system. It is suggested that the positive emotional reactions triggered by contact with green spaces go a long way to boost the immune system, resulting in various health benefits to individuals (Kuo, 2015).

3.2.5 Stress reduction

Contact with green spaces can have positive mental health effects such as improved relaxation and restorative atmosphere, thereby reducing stress (Hartig et al., 2014). It has been found that green spaces reduce exposure to challenging environmental conditions by creating a gap to stressors and decreasing the perceptual prominence of these stressors (Hartig et al., 2014). While other senses are equally of great importance, this mechanism relies much on vision (Grahn and Stigsdotter, 2010). The attractiveness of green spaces offers recreational opportunities which results in joy, excitement and relaxation among individuals, and in turn, help reduce stress (Xiaolu and Md Masud, 2012). In addition, green spaces may act as identity symbols for cities; as a result, they attract restorative activities such as tourism, holidays and offer venues for symbolic activities which can be refreshing to individuals (Xiaolu and Md Masud, 2012). The role of positive perception of green spaces in bringing about wellness related to mental health is further confirmed in a paper entitled '*Feel blue? Touch green! Participation in forest/green spaces management as a treatment for depression*' by Townsend (2006 p.1). In this paper, it is suggested that contact with nature can offer relief from stress and mental fatigue.

Observational studies in Sweden and Denmark have further found that access to a garden or green areas near homes is associated with lower perceived stress (Nielsen and Hansen, 2007; Herzele and Vries, 2011). A study done by Kuo (2001) on the restorative effects of green spaces on cognition and concentration showed that people who lived nearby trees and grass managed major life issues more effectively because of reduced mental fatigue than those without nature in their proximity. Again, a pre and post relocation longitudinal study of low-income urban children showed that children whose homes were improved the most with greenness tended also to have high cognitive capabilities (Wells and Evans, 2003).

Also, linked to positive perception of green spaces are sounds of nature such as birds and water. These sounds have been associated with well-being of individuals (Townsend, 2006). There are also some ameliorating effects in the green spaces which come from reduced heat stress. On this aspect, a study in Italy and UK provided confirmatory evidence that longer and frequent visits to green spaces alleviate the perception of thermal discomfort (Lafortezza et al., 2009).

Recently, an experimental study by Aspinall et al. (2013) which aimed at establishing the relationship between the environment and behaviour and emotions showed evidence of lower frustration, engagement and arousal, and higher meditation when participants were exposed to green spaces. This study used a new technology, electroencephalography (EEG), to access the cortical correlates of emotional states of individuals in contact with the environment. The study concluded that green spaces could be a mood-enhancing environment for walking or for other forms of physical and reflective activity (Aspinall et al., 2013). The next section looks at the theories that help explain the link between green spaces and mental well-being of individuals.

<u>Theories behind green spaces and mental well-being</u>

Two theories exist that help explain the link between green spaces and mental well-being of individuals: firstly, the psycho-evolutionary theory of stress reduction (PET) developed by Ulrich (1983); and secondly, the attention restoration theory (ART) by Kaplan and Kaplan (1989). The next section gives a brief discussion of these theories.

1. <u>Psycho-evolutionary theory of stress reduction (PET)</u>

This theory emphasizes the positive physiological and emotional changes that occur while viewing a scene of nature after a challenging situation or threat which resulted in high stress levels (Hartig et al., 2003). It was developed by Roger Ulrich (1983). The PET proposes that '*nature may allow psycho-physiological stress recovery through innate, adaptive responses to attributes of natural environment such as spatial openness,* *the presence of pattern or structure and water features which trigger positive emotional reactions related to safety and survival'* (Bowler et al., 2010 p.2). The underlining principle of this theory is that the perception of particular qualities and contents in a scene can support psycho-physiological stress recovery (Hartig et al., 2003; Health Council of the Netherlands, 2004). This implies that this theory focusses on stress recovery through affective responses to the environment that are visually evoked and instantly trigger feelings of liking, accompanied by change in psycho-physiological activation (Roe and Aspinall, 2011).

The PET holds in both clinical and non-clinical settings. A clinical setting example is the study by Ulrich (1984). This study compared recovery from surgery among patients in hospital rooms that had a view of trees while the other rooms had a view of a brick wall. The results showed that patients who had a view of trees used less analgesics, had fewer negative comments in nurses' notes and had short post-operative hospital stay compared with counterparts who had a brick wall view. The conclusion was that a view of natural elements serves as a distraction that evokes positive emotions, counter-acts stress and enhances pain management (Ulrich, 1984). The implication, therefore, is that the way patients' beds are arranged in a clinical setting could be very critical to their recovery as it could lead to psycho-physiological stress recovery. Raanaas et al. (2012 p.2) refer to this kind of set-up as '*bedscape*'.

A non-clinical setting example includes a recent observational study of pregnant women in England which showed a beneficial relationship between green spaces and depressive symptoms, with those living near greenest green space 20% less likely to report feeling depressed (McEachan et al., 2015). Some experiments in non-clinical settings have also reported psycho-physiological stress reduction associated with views to nature versus no views; and videotapes of natural settings versus urban settings (Raanaas et al., 2012). For example, an experiment in the UK showed that walking in green spaces was associated with enhanced relaxation and restoration compared with walking in areas where there were no green spaces (Aspinall et al., 2013). Some experimental studies have analysed the effect of being in natural environments (outdoor) and well-being. Results have concluded that green spaces may enhance coping mechanism of major life issues; and may improve children's mental ability (Kuo, 2001; Wells and Evans, 2003; Tzoulas et al., 2007; Giuseppe et al., 2012; Xiaolu and Md Masud, 2012). Other experimental studies which have used cortisol pattern as a biomarker of chronic stress have also demonstrated that contact with green spaces was associated with reduced stress, particularly in deprived areas (Ward Thompson et al., 2012; Gidlow et al., 2016).

2. Attention restoration theory (ART)

On the other hand, the ART proposes '*that nature provides the particular environmental stimuli to allow restoration from attention fatigue that occurs during the performance of cognitive tasks that require prolonged maintenance of directed attention*' (Bowler et al., 2010 p.2). This implies that exposure to nature offers restoration from the fatigue of prolonged mental work. The theory was developed by Kaplan and Kaplan (1989).

In urban life, living in high density areas exposes people to stimuli that demands a great deal of attention leading to mental fatigue (Peschardt and Stigsdotter, 2013). This kind of stimuli is known as hard fascination (Aspinall et al., 2013). In order to recover from such fatigue, the restorative environment provides stimuli that is compelling, through soft fascination and these stimuli do not require any mental effort (Irvine et al., 2013; Peschardt and Stigsdotter, 2013). The theoretical premise behind the whole recovery is facilitated by four factors that require natural environments for recovery to occur (Hartig et al., 2003; Abraham et al., 2010; Fan et al., 2011). These factors are: the sense of being away that renders a psychological temporary escape from the routine mental contents; the soft fascination from the natural environment which is the effortless and involuntary form of attention or curiosity which is sustained by the third factor; the scope or extent of how coherent and orderly the environment is. All these three factors are together matched with the fourth factor; which is an individual's compatibility of his inclinations, thus, opportunities provided by the setting and whether they satisfy an individual's purposes (Hartig et al., 2003; Maller et al., 2006; Hansmann et al., 2007; Aspinall et al., 2013).

Herzog et al. (1997) summarize the whole restorative experience through the ART in two stages: attention recovery and reflection. These two stages can be expanded into four successive stages for a restorative experience to pass through: the first stage involves clearing the head of distracting thoughts; the second one is recovery from directed attention fatigue which depletes attention capacity; the third is the process of contemplation or cognitive quietness; and the fourth stage is an immense sense of restoration (Roe and Aspinall, 2011; Susan and Henk, 2012).

Most evidence on restorative effects of green spaces is from experimental studies which reveal that viewing or being in contact with green spaces results in restorative physiological responses including reduced blood pressure (Hartig et al., 2003; Ottosson and Grahn, 2005; Lee et al., 2011; Lee et al., 2014). For example, an experimental study that sought to examine restorative cardiovascular responses to walking in green spaces versus urban environments found that walking in green spaces may promote cardiovascular relaxation compared with walking in urban environments (Lee et al., 2014). Another experiment using salivary cortisol as a biomarker of stress found that visits to green spaces even for a short term resulted in reduced stress and higher perceived restorativeness (Tyrväinen et al., 2014).

While the processes of the psycho-evolutionary and attention restoration theories differ, the two theories can be considered to bear some similarities which all point to the biophilia hypothesis (Wilson, 1984; Wilson and Kellert, 1993; WHO, 2016). This hypothesis claims that individuals possess an inherent preference for nature and that over the course of millions of years, human beings have adapted to respond positively to nature in order to thrive and for survival (Burls, 2007; Mason, 2009; Salingaros, 2015). The characteristics of the green spaces must be preferred and considered safe to trigger the positive emotional reactions which result in the positive effects on health and well-being.

3.3 Green spaces' health benefits and mixed findings

While much evidence converges to validate that green spaces offer broad and varied positive effects on individual's well-being (Mitchell and Popham, 2007; Mitchell and Popham, 2008; Bates and Marquit, 2011; Hartig et al., 2014), some studies have also emerged to offer contradictory, mixed, and unexpected findings on this relationship. For example, an observational study in Calgary, Canada which sought the association of spatial access to green spaces and child obesity concluded null findings (Potestio et al., 2009). Again, Richardson et al. (2012) found no evidence of association between green spaces and reduced mortality from heart disease, diabetes, lung cancer and automobile accidents at city level in the USA. However, this was partly attributed to the weak design of the study and the fact that green cities in the USA tend to be more spread out, hence, high dependence on the use of cars is a part of the lifestyle of people. Thus, some of the benefits that green spaces offer are masked. Another study in Denmark on access and use of green spaces and the impact on obesity showed an association of access to a garden or short distances to green spaces from homes with lower likelihood of obesity (Nielsen and Hansen, 2007). However, Coombes et al. (2010) did not find this association. Probably, this was due to the complex and diverse influences on bodyweight which, among others, include dietary behaviours (Coombes et al., 2010).

There are also other studies that have only reported associations of green spaces and well-being for certain groups, for particular areas, for specific types of green spaces, and for specific duration while other studies have found no relationship on these at all (Hillsdon et al., 2006; Pinder et al., 2009; Lachowycz and Jones, 2012; Wheeler et al., 2015). A possible explanation on these mixed findings could be that various studies use different measures of green spaces and that green spaces' positive health effects on individuals' well-being may be dependent on mediating factors such as gender, age, socioeconomic as well as other variables (Richardson and Mitchell, 2010; Donovan et al., 2011; Herzele and Vries, 2011; Annerstedt et al., 2012; Dadvand et al., 2012; Richardson et al., 2012; Wheeler et al., 2015).

Questions remain, therefore, whether the positive effects of green spaces are applicable to wider society and for long term. A general conclusion that can be made from the above discussion is that the positive health effects of green spaces on the well-being of individuals cannot be generalized (Tzoulas et al., 2007). These mixed findings may have important implications for study designs and sensitivity analysis of mediating factors in the studies about the relationship between green spaces and well-being of individuals. Despite the above conflicting outcomes on the role that green spaces play in the well-being of individuals, it is widely acknowledged that most experimental and observational studies show consistently strong and significant correlation between green spaces and improved mental health, particularly in deprived areas (David et al., 2008; Lafortezza et al., 2009; Bowler et al., 2010; Wolf, 2010; John, 2011; Ward Thompson et al., 2012; Alcock et al., 2014). This remains the focus of this thesis and the next section discusses how green spaces could be used as a preventive intervention to reduce mental health-related problems at population-level.

3.4 Green spaces and mental well-being

The World Health Organisation (WHO) defines mental health as a state of well-being in which an individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and is able to make a contribution to his or her community (WHO, 2010). It has been found that some mental health problems result from a life generally full of stressful events with often very little opportunities for mental restoration (Stigsdotter and Grahn, 2011). This stress impacts negatively on the general mental health of individuals. According to Annerstedt et al. (2010), stress is the result of the interactions between individuals and the environment that is perceived as straining or exceeding individuals' adaptive capacities, and threatening their well-being.

Worldwide, it is estimated that about 450 million people suffer from some mental illness (Dean et al., 2011). In general, mental health related diseases are estimated to represent 12% of the global disease burden with women twice likely to be depressed compared with men (Annerstedt et al., 2012).These figures are alarming and set to increase in the coming years (Maller et al., 2006).

In developed countries, mental health problems are particularly a big concern. About 75% of people in the developed world live in urban areas (Thompson Coon et al., 2011; Saulle and La Torre, 2012; White et al., 2013). Evidence suggests that part of the reasons for increased mental health problems in these countries is increased urbanization which result in: firstly, urban traffic noise which may cause non-auditory stress effects such as high blood pressure, cognitive deficits like poor memory, lack of concentration and poor attention; secondly, sedentary life style; and thirdly, poor sleep quality, among other problems (Gidlöf-Gunnarsson and Öhrström, 2007; Bratman et al., 2012; Dallat et al., 2013; Africa et al., 2014).

Interestingly, mental health inherently affects the physical health of people and vice versa (Stigsdotter and Grahn, 2011). In the UK alone, it is estimated that a quarter of the General Practitioner (GP) consultations are about mental health related problems (Marselle et al., 2012). This implies considerable economic costs to the National Health Service (NHS) in terms of treatment, social services and to the economy, in general, in terms of lost productivity (Shearer and Byford, 2015). All sectors of the UK economy account for a total mental health-related expenditure of about £22.5 billion (2007 estimate) (McCrone et al., 2008). In Scotland, the annual economic cost of mental health-related illnesses has been estimated at £10.7 billion (2009/2010 estimate), almost half of the UK's total expenditure on mental health-related problems (Silveirinha de Oliveira et al., 2013). Furthermore, it is estimated that 5% of men and 9%

of women in Scotland are likely to experience some mental health related problems in any given year (McCrone et al., 2008).

The above statistics are alarming and imply that poor mental health is a major public health concern. For this reason, improving mental health and well-being of the society is a public health priority. With the evidence of the positive health effects of green spaces on the mental well-being of individuals abound, there is an increased interest among public health policy makers to explore the use of green spaces to provide the positive mental health benefits at population-level (David et al., 2008; Lafortezza et al., 2009; Bowler et al., 2010; Pearce et al., 2010; Wolf, 2010; John, 2011; Ward Thompson et al., 2012; White et al., 2013; Tyrväinen et al., 2014; WHO, 2016). This implies that green spaces could be used as an upstream, population-wide preventive intervention to reduce mental health-related problems which could be less costly compared with the cost of treatment. The next section presents the Woods In and Around Towns (WIAT) case study in Scotland which explores the use of environmental improvements through physical and social interventions to enhance access to natural environments in deprived communities in order to improve the mental well-being of individuals.

3.5 The Woods In and Around Towns (WIAT) study

The WIAT study is part of the wider WIAT programme, a project of the Forestry Commission Scotland (FCS) that targets urban and post-industrial areas in the cities and towns of Scotland. These areas are classified as deprived and closely linked to poor health. Other initiatives in these deprived areas similar to the WIAT programme and also managed by the FCS include: the Forestry for People (F4P) which targets groups or communities to use and realise the potential contribution of local woodlands to health, learning and community cohesion (Scottish

Government, 2011); and programmes working with communities on woodland management; woods for health such as greening hospital grounds, creating attractive environments to improve health and life expectancy of people in Scotland; and woods for learning such as outdoor learning, teaching and playing (FCS, 2011b).

Research has shown that deprived areas are less likely to access good quality green spaces, hence, cannot have the positive health effects offered through contact with nature (Pearce et al., 2010). This being the case, the wider WIAT programme's core objective is to regenerate neglected woodlands, create new woodlands and support people to use and enjoy the woods. This is done through removing barriers that prevent individuals from visiting and benefiting from woodlands in order to improve their quality of life which would subsequently reduce health inequalities (FCS, 2011b).

The WIAT programme focuses on woodland within 1km of settlements of over 2000 people (FCS, 2011b). It was launched over a decade ago, in 2005, and the programme has run three phases. The first phase aimed to increase awareness of the benefits of urban woods and green networks; to identify priority areas for targeting resources; and to create woods on derelict and underutilized or land associated with new development (FCS, 2005; FCS, 2008).The second and third phase had the long term focus of further creating new woodlands; manage neglected woodlands; and involve communities in the use of their local woodland (FCS, 2008).

Currently, the WIAT programme is in its fourth phase with the focus on increasing access to woodlands for individuals to have contact with nature which may result in positive health benefits. This would help reduce health inequalities in areas of high social deprivation in Scotland (FCS, 2011b; FCS, 2015). The basis of delivering this fourth phase is the relatively strong evidence base on the positive association between green spaces and well-being of individuals (FCS, 2015).

Part of this fourth phase of the WIAT programme is a natural experiment which has been used in this thesis as a case study for an economic evaluation empirical analysis. The WIAT study specifically looks at the impact of woodlands on the psychological well-being and stress levels of people living in deprived communities in parts of Scotland. The study selected three paired intervention and control sites based on some neighbourhood characteristics such as: the woodlands should be in the area that cover a minimum of four hectares; and within the worst 30% of socio-economic deprivation in Scotland as assessed using the Scottish Index of Multiple Deprivation (SIMD); and the woodland sites should not have received environmental intervention investment or direct promotion within the last five years (NIHR, 2012; Silveirinha de Oliveira et al., 2013).

The intervention sites received the physical intervention in wave two which involved regenerating and improving woodlands through physical changes such as clearance of rubbish and any signs of vandalism; improved signage, access paths and trails, and marked entrances as depicted in Appendix 3 to Appendix 5. This was, then, followed by the social intervention in wave three which aimed to promote the woodlands as safe through community engagement; leafleting; led-walk programmes and other community-based programmes in order to increase access which could in turn increase physical activity and provide some health benefits to individuals (Silveirinha de Oliveira et al., 2013; Ward Thompson et al., 2013).

The study started in 2012 and was expected to finish in 2016. It is a cross-sectional and unbalanced panel survey of residents within 1.5km of the three paired sites (control and intervention). Table 3-1 below shows the paired sites for the WIAT study:

Intervention sites		Control sites	
1.	Haugh hill/Pollok	1.	Millikenpark
2.	Linwood	2.	Newarthill
3.	Mayfield	3.	Glenrothes

Table 3-1: Intervention and control sites for the WIAT study

All the above paired six sites lie within the Lowlands Forest District which covers the central belt of Scotland from the west to the east coast as shown on the map in Figure 3-2 below:

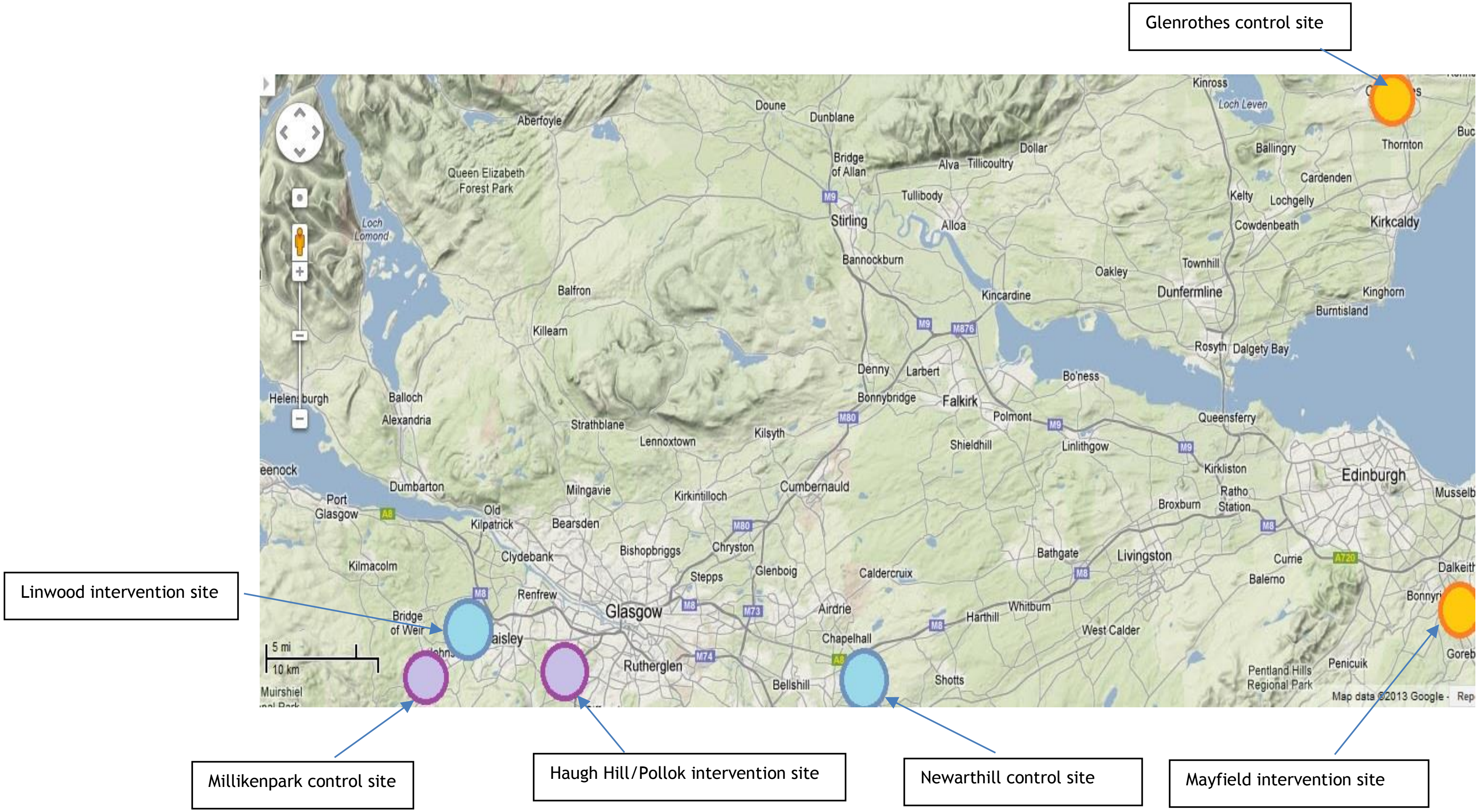


Figure 3-2: Intervention and control sites for the WIAT study

The control sites are used to determine any changes or improvements in the communities' physical activities and mental health attributed to the intervention. There were three waves of data collection. The first wave was in 2013 (wave one) which was the baseline before any intervention was undertaken, then the second wave was in 2014 (wave two) after the physical changes to the woodland, and wave three was in 2015 following the social intervention. The next section presents the conceptual model of the WIAT study.

Conceptual model of the WIAT study

The WIAT study's expected outcomes are broad consisting of health and non-health. Figure 3-3 below is a conceptual model depicting these outcomes following the physical and social intervention:

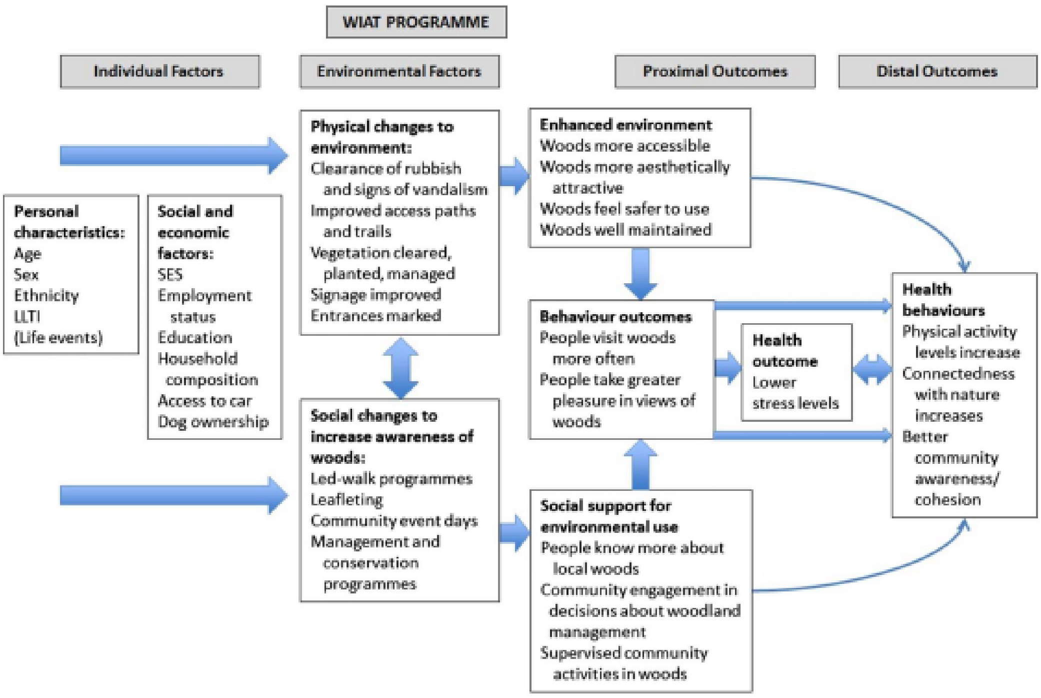


Figure 3-3: Conceptual model of the impacts of the WIAT study. Source: Silveirinha de Oliveira et al. (2013).

The primary health outcome is lower stress levels while the secondary health outcomes include health behaviours such as increased physical activities; connectedness with nature; and community awareness or cohesion. On the other hand, the non-health related outcomes consist of the enhanced environment, behavioural changes in terms of woodland use and social support for environmental use.

A pilot study which paired an intervention and control site in some deprived area of Glasgow and satisfied the inclusion criteria in the WIAT study, showed some evidence of increased visits over time in the intervention site compared with the control site; increased use and attitude to woodlands as venues for physical activity; and improvement in the quality of life of residents around the WIAT study area measured through improved perception of the quality of the physical environment (Ward Thompson et al., 2013). However, questions remain whether the WIAT intervention could improve the health-related quality of life as measured by the EQ-5D questionnaire and whether it is cost-effective.

The WIAT study, therefore, is used in this thesis as a case study for an empirical analysis to establish its impact in terms of health-related quality of life resulting from accessing woodlands as measured by the EQ-5D tool. Subsequently, a standard economic evaluation is undertaken to determine its cost-effectiveness in terms of cost per quality-adjusted life year (QALY) discussed in chapter two.

Given that the outcomes of the WIAT study are broad as revealed by its conceptual model in Figure 3-3, consisting of health and non-health related, valuing the other outcomes beyond health would be problematic as they would not fit in the traditional QALY framework of economic evaluation (Weatherly et al., 2009; Smith and Petticrew, 2010). This thesis goes further to develop a broader economic evaluation space capable of considering both the health and non-health related outcomes of a public health intervention whose outcomes are broad, like those of the WIAT intervention as demonstrated in chapter seven and eight of this thesis.

3.6 Summary

This chapter introduced and defined nature; its positive health effects on the well-being of individuals, particularly green spaces which is a subset of nature have been discussed; The chapter has also presented the mechanisms behind the link between green spaces and well-being of individuals together with the conceptual framework. Then, the chapter discussed the evidence on the mechanisms behind the association of green spaces and the well-being of individuals. Conflicting findings on the health benefits of green spaces have also been discussed. Then, the chapter proceeded to look at the positive health effects of green spaces particularly on mental well-being. This was followed by the presentation of the WIAT case study together with the conceptual model of its impacts on individuals. The impacts of the WIAT intervention on individuals as presented in the conceptual model are broad consisting of health and non-health related outcomes. It was noted that these broad outcomes would clearly present valuation challenges. Conducting an economic evaluation would also be problematic because the non-health outcomes would not fit in the standard economic evaluation framework of the QALY. Hence, the objective of this thesis which is to develop a broader evaluative space for economic evaluations of public health interventions that considers both the health and non-health related outcomes of a public health intervention.

3.7 Conclusion

Given the evidence that there is a positive association between green spaces and the well-being of individuals, especially in deprived areas, the use of the WIAT case study can be considered sensible because about 68% of the Scottish population are estimated to live within the areas targeted by the WIAT study, hence the woods in this area have the potential to offer the health benefits to a greater proportion of the population (FCS, 2011b). The next chapter discusses economic evaluations for public health interventions; the challenges faced when undertaking their economic evaluations; and the approaches that have been proposed in economic evaluation literature to deal with these challenges. It then presents the integrated approach that this thesis is proposing to value both the health and non-health outcomes of a public health intervention on the same monetary scale using the net monetary benefit (NMB) framework.

Chapter 4: Overview of economic evaluations of public health interventions

4.1 Introduction

Recent years have seen an increased interest in economic evaluations of public health interventions to ensure efficient allocation of limited resources given the continued pressure on healthcare, as well as other sectors' budgets (Griffin et al., 2010; NICE, 2012). The pressure on fixed budgets is exacerbated by the aging population and the complex health needs of the society (Schoen et al., 2009; Oliver et al., 2014). As with other interventions in healthcare, public health interventions consume healthcare or public-sector resources which could, otherwise, have been allocated elsewhere. This implies that they are also associated with an opportunity cost, hence the importance of an economic evaluation which could also ensure that the intervention does more good than harm to individuals compared with the status quo (Brousselle and Lessard, 2011; Morris et al., 2012; Trueman and Anokye, 2013).

In general, economic evaluations of public health interventions can aid decision-making on the efficient allocation of resources aimed at improving public health in the face of fixed budgets. However, public health interventions usually have broader outcomes which include health and non-health that make standard economic evaluation particularly challenging.

As defined in chapter one, public health is the art and science of preventing disease, prolonging life and promoting health through the organized efforts of society (WHO, 2014). Public health interventions, therefore, are a collective social effort to promote health and prevent diseases through population surveillance, regulation of determinants of health and the provision of key health services with an emphasis on prevention (Ruger and Ng, 2014). Further to these definitions, this chapter is organised as follows: it begins with a discussion on the mechanisms of public health, commonly referred to as determinants of health and the rationale for public health interventions. It proceeds to provide an overview of economic evaluations of public health interventions and their challenges. Then, it discusses some of the approaches that have been proposed in health economics literature to deal with these challenges. It goes further to present the integrated approach proposed by this thesis which could offer a broader economic evaluative space for public health interventions. This approach considers both the health and non-health related outcomes of public health interventions on the same monetary scale using the net monetary benefit (NMB) approach.

There are several arguments put forward in favour of public health interventions. These include: that their benefits could reach out to the majority of the population; and that it is morally right to have a healthy society (Rayner and Lang, 2012; Rayner and Lang, 2015). While clinical interventions focus on individual’s factors to achieve a healthy society, public health interventions focus on determinants of health which include the social, economic and physical environment, as well as individual factors (Squires et al., 2016). Many different models and frameworks have been used to describe these factors with most of them presenting similar elements. Perhaps, the most well-known model in the UK is the one developed by Dahlgren and Whitehead (1991). This model describes three general factors which determine health: first, are personal factors such as age, sex, genetics, biology, behaviour, risk factors and lifestyle; second, are community factors which include local influences like home, neighbourhood, workplace as well as the wider society consisting of education and healthcare system; and third, are environmental factors such as the physical, built, biological, and cultural environment (Dahlgren and Whitehead, 1991; Kelly et al., 2009).

Figure 4-1 below presents the Dahlgren and Whitehead (1991) model with three layers responsible for one's health. Individuals are placed at the centre surrounded by community and the general conditions which include the environment.

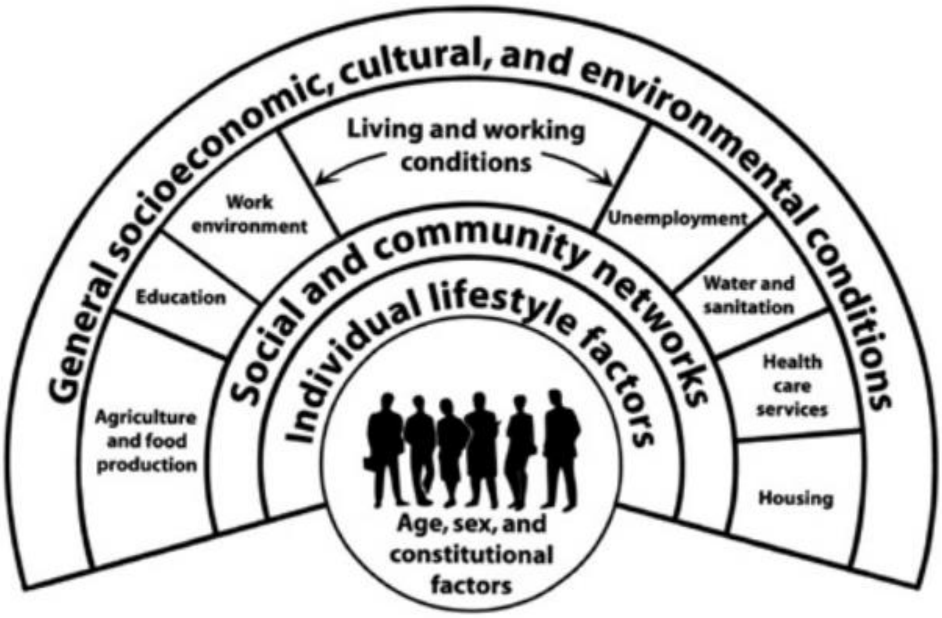


Figure 4-1: Determinants of health, Source: Dahlgren and Whitehead (1991).

The Dahlgren and Whitehead (1991) model in Figure 4-1 above, implies that the general health of individuals is largely driven by other factors including social and community networks and the general socio-economic, cultural and environmental factors including their interactions, other than individual characteristics. It can be argued that public health interventions could address health inequalities caused by its determinants and could be of low cost through: first, those interventions aimed at bringing about long-term improvements to health through structural changes such as the environment; second, those interventions aimed at improving living and working conditions of the society; third, the interventions that strengthen social and community support; and lastly, those interventions that influence individual life styles and attitudes (Dahlgren and Whitehead, 1991; Kelly et al., 2005; Barrett et al., 2014; WHO, 2014).

The WIAT case study intervention used in this thesis is an example of an intervention that brings changes to the environment to change individual behaviours and perception which could result in improved mental well-being at population-level, thereby reducing health inequalities. It could be considered to be reflected in the outer layer of Dahlgren and Whitehead (1991) representing the environmental factors. Public health interventions related to environmental improvements have the advantage of having a long-term perspective, being all inclusive in nature to benefit anyone, providing broad outcomes of both health and non-health related (Rayner and Lang, 2012; Silveirinha de Oliveira et al., 2013; Rayner and Lang, 2015). This could have implications for economic evaluation in terms of methods of analysis to use to consider the broad outcomes, and time horizon, among others.

4.2 Methodological challenges

In general, the objectives of public health interventions and its outcomes extend beyond health (Wanless, 2004). They are primarily concerned with improving health and also reducing health inequalities (Griffin et al., 2010). For these reasons, their economic evaluations are scarce possibly due to the complexity of handling the broad outcomes. The standard economic evaluation framework of a QALY, as discussed in chapter two, is not capable of valuing the non-health related outcomes.

One of the most significant current discussions in health economics is the challenge posed by undertaking economic evaluations of public health interventions (Weatherly et al., 2009; Curtis, 2014; van Mastrigt et al., 2015). Conducting their economic evaluation implies a move beyond the narrower concerns of the standard economic evaluation framework that focuses on, and aims to maximize, health outcomes in the face of some health sector budget constraint, while completely neglecting other non-health related outcomes. Comparisons between alternative interventions in standard economic evaluations are made based on a single measure of the quality adjusted life year (QALY). This measure of outcome is assumed to be the same for everyone who accrues it. This assumption, however, cannot apply to public health interventions whose impacts are multiple and varied which imply difficulties in making comparisons between interventions. In addition, the outcomes of most public health interventions go beyond health, especially those relevant to addressing health inequalities. The non-health outcomes affect individuals or groups differently compared with the QALYs which are assumed to be the same for everyone who gains them (Weinstein, 1988; Weinstein et al., 2009; Curtis, 2014). Given that the outcomes of public health interventions are broad, the standard and favoured economic evaluation approach of the QALY framework has been found to be inadequate or unsuitable for valuing the broad health and non-health outcomes of public health interventions (Weatherly et al., 2009; McIntosh et al., 2010). The next section looks specifically at methodological challenges of using standard economic evaluation framework for public health interventions.

There are several methodological challenges that are highlighted in literature on conducting economic evaluations of public health interventions (Weatherly et al., 2009; Curtis, 2014; van Mastrigt et al., 2015). These challenges relate to the complex nature of public health intervention; the difficulty in measuring and valuing all relevant broad outcomes; the difficulty in choosing the economic evaluation perspective and viewpoint for analysis; the fact that costs and outcomes may span multiple sectors; the difficulty in attributing outcomes to the intervention; concerns about the time horizon for an economic evaluation; the difficulty in determining the acceptable cost-effectiveness threshold and lastly; how to deal with equity considerations in an economic evaluation as the effect of an intervention may vary among individuals or groups (Weatherly et al., 2009; Weatherly et al., 2014; van Mastrigt et al., 2015). These challenges are discussed in detail in the next section.

4.2.1 Complexity of public health interventions

Most public health interventions are considered to be complex because they contain several interacting components and often have multiple outcomes that consist of health and non-health related outcomes (Craig et al., 2008; Smith and Petticrew, 2010). This complexity has become a common explanation for the dearth of economic evaluations of public health interventions, especially on how to handle the health and non-health outcomes (Campbell et al., 2000; McIntosh, 2006; Craig et al., 2008; Shiell et al., 2008; Griffin et al., 2010; Curtis, 2014; Lawson et al., 2014; Rabarison et al., 2015).

The traditionally, commonly used and favoured economic evaluation framework of interventions has maximizing health as an objective (Drummond et al., 2009). Traditionally, the ‘health’ outcome is measured by instruments such as the EQ-5D questionnaire and valued in terms of Quality-Adjusted Life Years (QALYs). This approach is recommended by bodies such as the National Institute for Health and Care Excellence (NICE) and the Scottish Medicines Consortium (SMC) (Drummond et al., 2007; NICE, 2013; SMC, 2015). The QALY simultaneously captures gains from reduced morbidity (quality of life) and reduced mortality (quantity of life) and has been a primary measure in economic evaluations in healthcare. However, the QALY framework becomes problematic for an economic evaluation of a public health intervention, which by its nature can have multiple and varied outcomes (health and non-health related) because good health is largely driven by other factors as described by Dahlgren and Whitehead (1991) model.

These broad outcomes may not all be captured in a single unit of a QALY, hence the QALY framework is considered inadequate or unsuitable because it neglects the non-health related outcomes. This neglect is generally acknowledged in literature (Lancsar and Louviere, 2008; Payne et al., 2013; McIntosh et al., 2014). Furthermore, the traditional QALY framework often measures the outcome of an intervention on the affected group or individuals. However, a public health intervention may have outcomes which spill-over to other individuals or groups, as such, it may be necessary to measure and value the impact of an intervention on these individuals or groups not directly targeted by the intervention. As can be noted from the above discussion, an economic evaluation of public health interventions needs to be sufficiently comprehensive to encompass the complexity, which is a difficult task and or costly to undertake (Rychetnik et al., 2002; Weatherly et al., 2009).

4.2.2 Outcome measurement and valuation

The economic evaluation of public health interventions further presents the problem of outcome measurement and valuation (Weatherly et al., 2009; Smith and Petticrew, 2010). Shiell (2007) argues that the value of the outcomes of public health interventions can only be captured if one can track and measure the multiple outcomes together with their multiplier effects to individuals other than those targeted by the intervention. In addition, the external and spill-over effects of public health interventions are difficult to account for because it is practically impossible to include or involve all individuals who directly or indirectly receive the intervention in an economic evaluation (Evers et al., 2015).

Another view on the challenges of outcome measurement for public health interventions is that of Weatherly et al. (2014) who point out that, unlike the standard healthcare interventions, public health interventions often target populations or communities rather than individuals, as a result their outcomes are relatively small and often very difficult to detect and measure at an individual level. Furthermore, the effects could vary by individual characteristics and settings (Richardson et al., 2010).

Another challenge to conducting economic evaluations of public health interventions is that most public health interventions are context-specific and address a specific issue, hence practically difficult to apply standard economic evaluation approaches to measure and value their outcomes (Shiell et al., 2008; Cookson et al., 2009; Curtis, 2014). For example, the impacts of the WIAT case study used in this thesis are specifically related to mental well-being and physical activities in some deprived areas of Scotland. In this case, the question remains as to whether the use of the QALY framework through a generic EQ-5D questionnaire, solely to allow an economic evaluation, is appropriate compared with the use of a condition-specific tool such as the Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS).

4.2.3 Economic evaluation perspective and viewpoint for analysis

Furthermore, economic evaluations of public health interventions present the problem of choosing the economic evaluation perspective and viewpoint for analysis. The impacts of a public health intervention often go beyond health. As a result, the logic of adopting the traditionally favoured decision-makers' perspective associated with the maximization of health as an outcome in an economic evaluation using the extra-welfarist CUA becomes untenable (Tchouaket and Brousselle, 2013). At the same time, choosing the societal perspective that is broader in focus to include the non-health outcomes through the welfarist CBA approach is problematic because of lack of acceptability of the CBA techniques of valuing health as discussed in chapter two (McIntosh et al., 2010). Questions still remain on the appropriate approach of conducting an economic evaluation of public health interventions (Payne and Thompson, 2015).

4.2.4 Dealing with inter-sectoral costs and consequences

Another problem presented by economic evaluations of public health interventions is how to deal with costs and broad outcomes which may span multiple or inter-sectors of the economy (Weatherly et al., 2009; Smith and Petticrew, 2010). For example, a public health intervention can consume resources which affect multiple sectors other than health and can also have multiple effects affecting and spanning other sectors other than health. The task of identifying these multiple sectors and assessing the amount of effect applicable to each sector can be relatively difficult.

4.2.5 Attributing outcomes to interventions

The difficulty of attributing the outcomes of an intervention, for example, a change in health-related quality of life or improved perception about health, to the intervention is another problem in economic evaluations of public health interventions (Weatherly et al., 2009). As noted earlier, many public health interventions often target populations or wider communities rather than individuals, thus, the outcomes are often minimal and sometimes difficult to detect at population-level (Evers et al., 2015). This makes it difficult to attribute outcomes to the intervention because of the complexity presented by several interrelated components of an intervention (Evers et al., 2015). Turning to the example of the WIAT case study, this challenge has been dealt with at the study design stage given its natural experiment design, with control and intervention groups. The control group is used as the counterfactual to determine the causal inference of the intervention.

4.2.6 Time horizon concerns

The impact of public health interventions often takes a long time to manifest, and could be inter-generational, hence, it becomes problematic to detect any meaningful effects in the short run (Tchouaket and Brousselle, 2013; Mays and Mamaril, 2015). This results in difficulty to demonstrate the economic value for public health interventions in the short run (Kelly et al., 2010; Evers et al., 2015). Long-term follow up of the outcomes of public health interventions is usually practically difficult or costly, especially when targeted at population-level.

4.2.7 Incorporating equity considerations

Lastly, as noted earlier, the focus on health in standard economic evaluations implies maximizing quality-adjusted life years (QALYs) as an objective. One of the key assumptions in standard economic evaluations about a QALY outcome measure is that it is the same for everyone who gains or loses it, no matter the distribution. This assumption is commonly expressed as "*a QALY is a QALY*" (Weinstein, 1988: p1; Weinstein et al., 2009: pS8). However, this concept of uniform distribution of a QALYs is cannot be applicable to public health interventions because their impact of may vary among individuals or groups.

Having discussed the economic evaluation challenges for public health interventions, it can be noted that the challenges are enormous. Currently, there is unabated debate on how to overcome these challenges to reflect the wider impacts of public health interventions which extend beyond health (Coast et al., 2008a; Griffin et al., 2010; Ryan et al., 2012a; Payne et al., 2013; Curtis, 2014; Payne and Thompson, 2015). The next section looks at some approaches that have been proposed in literature to deal with the challenges in economic evaluations of public health interventions discussed above.

4.3 Dealing with these challenges

Currently, approaches that can consider the broad outcomes of public health interventions into economic evaluations are relatively underdeveloped. At the same time, there is no clear way forward on how to conduct economic evaluations of public health interventions (Owen et al., 2011). As noted earlier, the debate on how to deal with the above challenges of economic evaluations for public health interventions is still unabated. Health economics literature and a recent qualitative study provide some suggestions on how to consider the broad outcomes of a public health intervention in an economic evaluation with the purported objective of expanding the economic evaluative space beyond health (Kelly et al., 2005; Coast et al., 2008a; de Bekker-Grob et al., 2012; Payne et al., 2013; Curtis, 2014; van Mastrigt et al., 2015). These include the following:

4.3.1 A "*do nothing*" approach

It is not immediately clear how non-health related outcomes of public health intervention can be incorporated in economic evaluations. For this reason, one approach that has been proposed to handle the problem of broad outcomes in economic evaluations of public health interventions is to "*do nothing*" (Payne et al., 2013). This is, in a way, an acknowledgement that it is difficult to develop a method that captures both health and non-health outcomes of public health interventions, hence there is no need or it is not possible to conduct their economic evaluations (Kelly et al., 2005; Payne et al., 2013).

However, it can be argued that this proposition is problematic because the option of doing nothing is worse, possibly can result in wrong decisions, and does not help decision-making at all. There is need for a basis of allocating scarce resources in healthcare, and indeed in other sectors of the economy because of the opportunity cost implication and or to ensure that interventions do good rather than harm to individuals.

Similar to the "do nothing" approach, the other suggestion has been to conduct standard economic evaluations for public health interventions and exclude or ignore all objectives that go beyond health on the basis that they are outside the realms of an economic evaluation (Richardson, 2009). This suggestion can be problematic as it can result in underestimating the outcomes of a public health intervention by emphasizing on the health outcomes alone.

4.3.2 The social objective framework

Another proposition is to conduct an economic evaluation using the traditional approach but with the inclusion of the social objectives which consider non-health outcomes and address health inequalities along with the objective of maximizing health (Richardson, 2009; Sheill, 2009). However, questions remain and there is a disagreement on how to do this as combining the two objectives may become problematic in terms of formulating decision rules. For example, in the event that an intervention achieves greater health outcome whilst the comparator achieves greater health equality (Weatherly et al., 2014).

4.3.3 A trade-off approach between maximizing health and equity

Cookson et al. (2009) have proposed a trade-off approach that could consider health efficiency and equity in economic evaluations of public health intervention. First, this approach simply entails providing evidence of health equity considerations that are at stake in the standard economic evaluation; second, providing quantitative evidence that the impact the intervention will have on health inequalities using, for example, sub-group analysis or simulation modelling; third, estimating the opportunity cost of the trade-off between efficiency and equity such as QALYs sacrificed by undertaking an equitable option rather than a QALY-maximizing one. Finally, assigning differential weighting to QALYs depending on equity-relevant characteristics as valued by stakeholders using techniques such as stated preference discrete choice (SPDCE) discussed in chapter two. There is, currently, a growing body of quantitative research that is focussing on weighting QALYs to determine if individuals value QALYs equally (Lancsar et al., 2011; Bobinac et al., 2012; Shah et al., 2013).

Perhaps, this approach would resolve the limitations of the social objective framework proposed by Richardson (2009). The Cookson et al. (2009) trade-off approach can be very difficult to implement in the absence of a definition of equity or whether there should be a single definition and the lack of agreement on what should be equitable in an economic evaluation (Richardson, 2009).

4.3.4 The cost benefit analysis (CBA) approach

Another suggestion is to use a welfarist approach of cost-benefit analysis. This approach would value all the outcomes of a public health intervention but in monetary terms. The CBA approach is theoretically considered to be broad in focus with both the costs and outcomes in monetary terms (Kelly et al., 2005; Gray et al., 2010; Marsh et al., 2012; Evers et al., 2015; van Mastrigt et al., 2015). The monetary metric would permit questions of allocative efficiency to be addressed across different sectors of the economy rather than just health (Drummond et al., 2005; Gray et al., 2010). This approach is recommended by the UK's Treasury Green Book as an approach which considers the broader societal costs and outcomes that are comparable because of the same monetary metric (Fujiwara and Campbell, 2011).

However, as has been discussed in this chapter two, the use of the CBA approach in economic evaluations of healthcare is hampered by its methodological challenges related to the direct willingness to pay (WTP) elicitation techniques. There is widespread dislike of placing explicit monetary values on health or life and the implication to pay is considered to favour only those who can afford, hence may be discriminatory (Drummond et al., 2005; Gray et al., 2010; van der Pol and McKenzie, 2010). In addition, the placement of monetary values on outcomes potentially makes an intervention theoretically and practically noncomparable with other healthcare interventions whose outcomes are mostly valued in QALYs or other terms.

One option to address the problems of WTP elicitation techniques which is considered as a methodological improvement of CBA approaches of eliciting WTP values is the stated preference discrete choice experiment (SPDCE) which was discussed in chapter two. The SPDCE method indirectly estimates the WTP values through a marginal rate of substitution process. This is also discussed in chapter two.

4.3.5 The capability approach

Other researchers have cautiously advocated moving away from measures which use health or utility (WTP) to the capability approach (Coast et al., 2008b; Marsh et al., 2012). This approach evaluates an intervention based on an individual's ability to function in a particular way without a prescription of any particular capabilities, hence offering a broad framework for evaluation. It advocates the view that individuals' well-being is best reflected by, and promoted through, their capabilities rather than utility or other aggregate indicators such as Gross Domestic Product (GDP) (Smith et al., 2012).

However, despite the potential of the capability approach in economic evaluations of public health interventions, it has a major drawback of lacking a methodology to operationalize and interpret it so that it can be used in resource allocation decisions of public health interventions (Coast et al., 2015b). Some researchers have further criticised it for relying much on expert opinion about what constitutes well-being (Marsh et al., 2012).

4.3.6 An expanded QALY using cost-utility analysis

This approach of enhancing the QALY framework has been suggested in literature but is rarely used in practice. It can be achieved by incorporating an additional dimension that would capture and measure outcomes other than health on the EQ-5D or other health-related quality of life (HRQoL) instruments resulting in an expanded QALY (van Mastrigt et al., 2015). The possible problem with this approach would be how to value the outcomes captured by this additional dimension. Furthermore, the question that would remain is how to aggregate the other non-health outcomes with the health outcomes since they would, arguably, be valued on different metrics.

4.3.7 A multi-sectoral approach using cost-effectiveness analysis

This approach has been proposed by Claxton et al. (2007). It suggests assessing public health interventions based on its impacts across multiple sectors. The sector specific measure of net benefit of the intervention is assessed against the specific sector's budget. A compensation test could be used as a decision rule to judge the cost-effectiveness of the intervention. If the net benefit of the intervention is positive in all sectors, then it should be adopted. However, if the net benefit is positive in some sectors and negative in others, then it should be adopted if the sectors with a positive net benefit could compensate those with the negative net benefit and still result in a positive funding. While this approach could offer solutions in the absence of appropriate approaches, it is still unclear how this can be done in practice.

4.3.8 Multi-criteria decision analysis (MCDA) approach

Another suggestion is the use of an MCDA approach. This is a decision analytic approach which looks at various alternatives, defines decision criteria akin to attributes in SPDCE, puts weight to the criteria according to importance, then a scoring exercise is undertaken for each alternative to create an overall assessment of value (Marsh et al., 2014). It appears to be a promising tool for interdisciplinary decision-making because of its nature to accommodate different viewpoints. As such, it has been suggested that it would be possible to capture the outcomes of an intervention that go beyond health using multiple criteria in a transparent, and consistent manner (Thokala and Duenas, 2012; Thokala et al., 2016).

However, there is little application of this approach in healthcare. This is due, partly, to lack of guidance up until recently on how to conduct each stage of the approach especially on what and how many criteria to include and how to do the scoring in order to assess the importance of the criteria (Marsh et al., 2014; Marsh et al., 2016). So far, the approach does not have a clear method of accounting for or quantifying uncertainty (Marsh et al., 2014). It has also been argued that the MCDA approach lacks the notion of opportunity cost or sacrifice in its approach when defining weights to the defined criterion (Briggs, 2016). The notion of opportunity cost is fundamental in determining value in economics and its absence in the MCDA approach might not appeal to economists.

Recently, guidelines have been developed on how to conduct an MCDA in healthcare by the International Society for Pharmaco-economics and Outcomes Research (ISPOR) (Marsh et al., 2016; Thokala et al., 2016). Perhaps these guidelines could result in a widespread use of the MCDA approach in economic evaluations of public health interventions.

4.3.9 The subjective well-being (SWB) measure

Another suggestion relates to incorporating the subjective well-being (SWB) measure in economic evaluations of public health interventions. This approach involves measuring how the well-being of individuals varies with the outcomes of an intervention. It uses self-reported stated life satisfaction surveys to uncover the social impact estimates of an intervention (Marsh et al., 2012). Its basis is that individuals are the best judges of their own conditions, as a result, the intervention of any type should aim to maximize the sum of individuals' happiness (Fujiwara and

Campbell, 2011; Marsh et al., 2012; Greco et al., 2016). The respondents are asked to provide a subjective assessment of their overall well-being which is matched with the overall objective measure of the standard economic evaluation. For example, a SWB question can be: ‘On a scale of 0-5, where 0 is “not at all satisfied” and 5 is “completely satisfied”, how satisfied are you with…?’

It is essential to note, however, that the SWB methodological framework is still currently not fully developed to aid decision-making in economic evaluations of healthcare. Furthermore, the SWB measure can be affected by other contextual factors such as weather, mood, order of the survey questions, and who is present during the survey, culture, and experience, among others, which add to its limitations (Schwarz and Strack, 1999; Fujiwara and Campbell, 2011). Despite these limitations, the SWB measure has the potential and its capability is promising in economic evaluations of public health interventions. It may complement the traditional economic evaluation methods of analysis to value outcomes of an intervention that are beyond health or are non-market priced (Fujiwara and Campbell, 2011).

4.3.10 The cost-consequences analysis (CCA) approach

There have also been suggestions to use the cost-consequences analysis (CCA) approach. This approach is discussed in chapter two. It is does not qualify to be an economic evaluation technique in the strictest sense but nevertheless allows the presentation of different elements of an economic evaluation to be listed under the cost and outcome side in a balance-sheet format. This gives freedom to decision-makers to choose relevant information from the disaggregated results to make various resource allocation decisions (Mauskopf et al., 1998; Coast, 2004).

The CCA approach would allow health outcomes of public health interventions to be valued using the standard QALY framework while the non-health related outcomes can be valued using the stated preference discrete choice experiment approach of the cost-benefit analysis (CBA), for example. The problem that can arise, however, is how to combine the multiple and varied outcomes using a single metric. Failure to do this can lead to problems of comparing overall outcomes of an intervention with those of alternative interventions within and across different sectors.

Cognizant of the above discussed methodological challenges and drawing upon the suggested approaches in health economics literature for conducting economic evaluations of public health interventions, this thesis explores the use of an integrated approach. This approach could consider both the health and non-health related outcomes of public health interventions on the same monetary scale through the net monetary benefit (NMB) approach. The next section presents the integrated approach in detail.

4.4 Summary

This chapter has presented an overview of economic evaluations of public health interventions, and the methodological challenges that hinder the undertaking of economic evaluations of public health interventions. Then, suggestions in health economics literature on how to deal with the challenges have been discussed. The next section presents the integrated approach that this thesis is proposing. The integrated approach offers a broader economic evaluative space that can consider both health and non-health outcomes of a public health intervention on a single monetary scale using the net monetary benefit (NMB) framework.

4.5 The integrated approach

4.5.1 Methods

The proposed integrated approach is implemented through a case study of the Woods In and Around Towns (WIAT) study presented in chapter three.

This is an environmental improvement project aimed at improving the mental well-being and increasing physical activity of individuals in some deprived areas in Scotland. The outcomes of the WIAT intervention are broad, consisting of health and non-health and are examples of outcomes of some public health interventions. As previously noted, undertaking an economic evaluation of an intervention with broad outcomes presents methodological challenges.

The integrated approach provides a practical solution to challenges of economic evaluation of public health. In this approach, a standard extra-welfarist cost-utility approach is used to value the health outcomes in terms of QALYs while the identified non-health outcomes are captured and valued through the welfarist stated preference discrete choice experiment (SPDCE). A cost attribute is included in the SPDCE as a payment vehicle to allow the indirect elicitation of the societal willingness to pay (WTP) values. Then, the WIAT main study questionnaire items that are considered to measure the non-health outcomes are mapped to the SPDCE attributes and levels. Subsequently, the SPDCE WTP values are applied to the incremental changes or improvements in the attributes and levels resulting from the intervention. The cost of resources used together with the health outcomes from the CUA and the non-health outcomes valued by the SPDCE are initially presented in a cost-consequences analysis (CCA) format without being restricted to a single metric. Then, a net monetary benefit (NMB) framework is used to monetize QALYs for the health outcomes, effectively deriving a cost-benefit analysis (CBA) framework. This allows the CUA results to be combined with the SPDCE results on the same monetary scale, in a manner that they complement or add value to each, resulting in a broader evaluative space for both health and non-health outcomes. The same decision rule for the standard NMB framework is used: when the NMB is greater than zero, the intervention is considered cost-effective; while if it is less than zero, then the intervention is not cost-effective (Morris et al., 2012; Edlin et al., 2015).

The following specification summarizes the proposed integrated approach:

$$NMB = \Delta effect \times \lambda_1 + \Delta attributes\ or\ \ levels \times \lambda_2 - \Delta costs \qquad (7)$$

Where $\Delta Effect$ is the incremental health effect in terms of QALY; λ_1 is the societal willingness to pay for a unit of health effect gained (QALY) which is between £20,000 and £30,000 for the UK (NICE, 2013), in the event when this WTP estimate is not readily available, any estimate of policy-relevant value of willingness to pay can be used (Glick et al., 2015); *Δattributes or levels* is the incremental change or improvement in attributes or levels resulting from the intervention; λ_2 is the willingness to pay values from the SPDCE; and *Δcost* is the differential cost between the intervention and control group.

Figure 4-4 below presents the conceptual model of the proposed integrated approach:

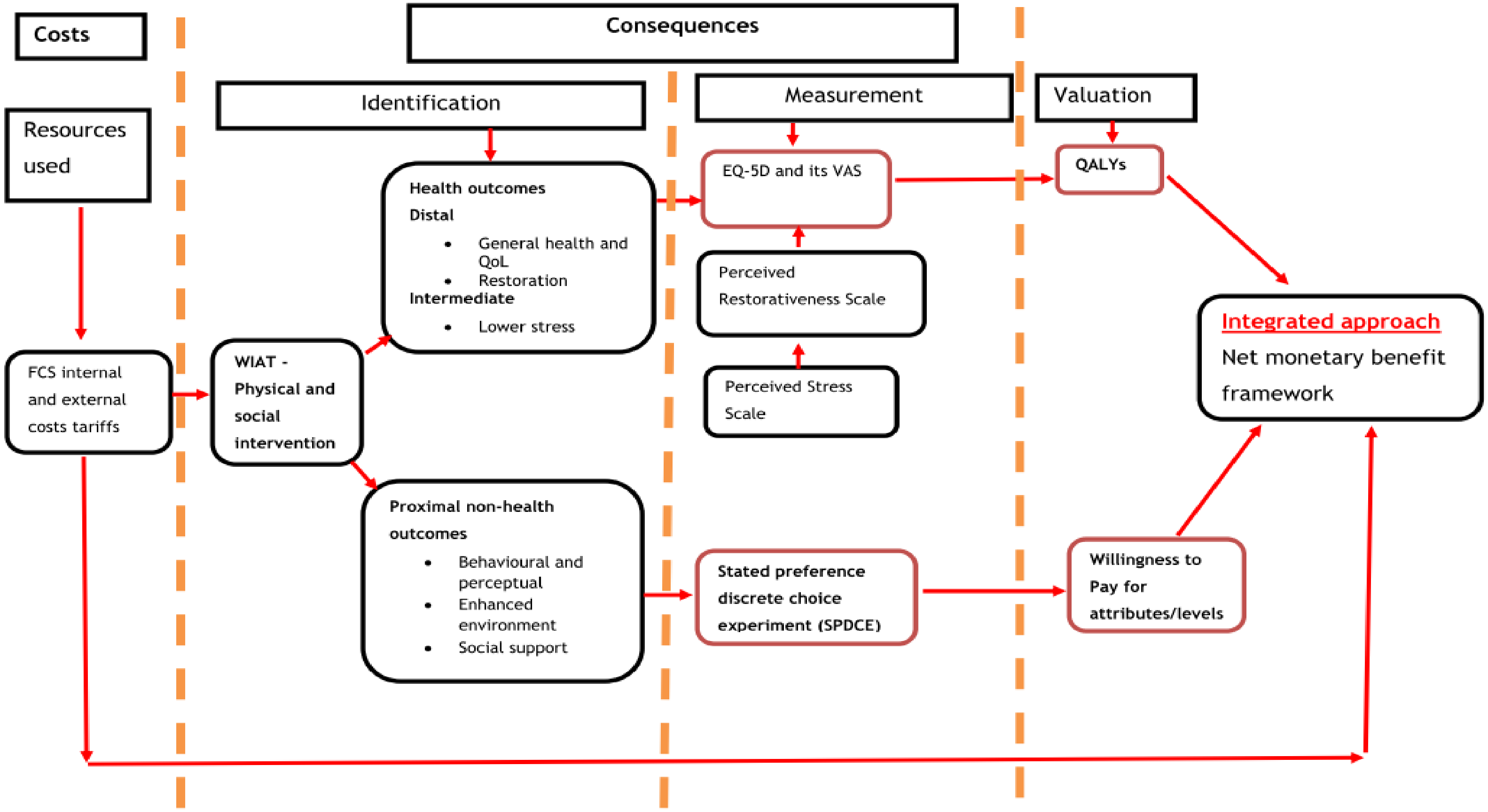


Figure 4-4: The integrated approach for the economic evaluation of the WIAT study.

4.5.2 Results

The integrated approach appears to be feasible and capable of providing a new broader conceptualization and operational approach which considers both health and non-health outcomes on the same monetary scale in an economic evaluation of an intervention with broad outcomes. The monetary metric would allow making comparisons of interventions within healthcare and across other sectors, hence addressing questions of both allocative and technical efficiency (thus, achieving the right mixture of limited resources to obtain maximum possible improvement in outcome and obtaining maximum possible improvement in outcome from limited resources, respectively).

4.5.3 Discussion

The novelty of this approach is the use of the versatile net monetary benefit (NMB) approach, commonly used to resolve incremental cost-effectiveness ratio (ICER) problems in cost-effectiveness analysis. In the integrated approach, this is used to combine the extra-welfarist CUA approach with the welfarist SPDCE approach in cost-benefit analysis framework in a complementary manner which adds value to each other. Hence, offering a broader economic evaluative space for public health interventions capable of considering both health and non-health outcomes. This approach ties well with NICE guidance which permits the use of welfarist approaches in economic evaluations of public health interventions when extra-welfarist approaches are considered to be unsuitable or inadequate (NICE, 2012).

Recently, a similar approach to this integrated approach has been suggested by Wildman et al. (2016) for economic evaluations of assisted living technologies (ALTs). This is in the context of valuing both health and social care on the same monetary scale using the net monetary benefit framework. This thesis has gone further to demonstrate how to operationalize the approach in a different context of environmental improvement intervention using the WIAT case study.

4.5.4 Conclusion

The chapter went further to present the proposed integrated approach which provides a broader economic evaluative space for public health interventions. It can consider both the health and non-health outcomes on the same monetary scale using the net monetary benefit. The next section presents the second part of this thesis which discusses the implementation of the integrated approach using the WIAT case study.

Part 2: Implementation of the integrated approach using the WIAT case study

This second part of the thesis and consists of chapter five, six, seven, eight and nine. It draws upon part 1 which discusses the theoretical aspects of the standard economic evaluation in healthcare; presents the relationship between nature, especially green spaces and the well-being of individuals and introduces the Woods In and Around Towns (WIAT) case study used for empirical analysis; discusses economic evaluations of public health interventions whose outcome characteristics are like those of the WIAT case study. This is followed by the economic evaluation challenges of public health interventions and the novel approach, referred to as the integrated approach, that this thesis is proposing is introduced.

This second part implements the integrated approach by applying those theoretical underpinnings of the standard economic evaluation in healthcare; the understanding of the association between and nature and well-being of individuals, specifically about the WIAT case study; and the suggestions in literature to address the economic evaluation challenges presented by interventions with both health and non-health related outcomes.

Chapter 5: Costing resource use of the WIAT study

5.1 Introduction

This chapter begins to implement the integrated approach as presented in the conceptual model in chapter four using the WIAT case study. It starts by estimating the cost of WIAT study which is driven by the resources used in its delivery. The costing process starts with the identification, quantification and valuation of the resources used.

This chapter addresses the first objective of the thesis that seeks to assess the cost of resources used in implementation of the WIAT study. As pointed out in chapter two, the costing aspect in standard economic evaluations of healthcare is often neglected (McIntosh et al., 2014). Hence, this chapter further addresses this problem.

Costing resource use of the WIAT study was done in two stages: during the physical intervention; and the social intervention. The next section provides the details of what was done at each of these stages to develop the cost estimate of the study.

5.2 Methods

5.2.1 Identification of the costs

The source of funding for public woodlands is tax from the public. Hence the cost of delivering the WIAT programme represents the societal opportunity cost. It was necessary, therefore, to adopt the societal perspective at the beginning of the costing process to ensure that only relevant costs of the resources used in the WIAT project were identified.

Consultations with the Forestry Commission Scotland (FCS) were made to identify the type and unit costs related to the implementation of the WIAT study. The costs were identified to include the FCS's staff time for overseeing and implementing the programme and other costs related to inputs required for the delivery of both the physical and social interventions. Other costs relate to contractors who were involved in carrying out the physical and social intervention work. There were also other costs related to community involvement in the programme during the social intervention but it was difficult to identify each unit and assign the unit cost of involvement to the study. This meant that only costs identified in the WIAT study related to the FCS' staff time and resources used for the development and implementation of the physical and social interventions were included.

5.2.2 Measurement of resource use

As regards the measurement of the resources used in the WIAT study, the resources used were measured in physical units. As discussed in chapter two, there are two approaches that are commonly used for this undertaking in economic evaluations of healthcare: top-down; and bottom-up approaches (Mogyorosy and Smith, 2005; McIntosh et al., 2010). The appropriateness of each approach is dependent on the context of the evaluation.

A top-down approach was adopted for the WIAT study simply because it was practically feasible, transparent and easy to use (Mogyorosy and Smith, 2005; Simoens, 2009). Actual costs were quantified in terms of time in days spent on the study and in terms of the physical quantities of the resources required during the physical and social intervention programmes.

5.2.3 Valuation of the costs to the intervention

With respect to the valuation of the costs of the project, the quantity of units measured was multiplied by their unit cost or price or pay rate per day to obtain the total cost. The pay rates per day were based on the FCS's cost valuation tariffs and were used to compute the total cost of the time used in the WIAT study. These tariffs involve a higher-level cost summary based on the FCS's internal and external costing organised as follows: internal costs which use internal rates per day for various staff grades or various FCS's resources; and external costs which include contracts for external resources or the use of internal rates per day for resources external to the FCS. These tariffs represented the opportunity cost of using a given resource such as time in the implementation of the WIAT study. The use of opportunity cost concept in this context is a theoretically accepted way of valuing costs in economic evaluations, hence was considered to be standard and reliable (Mogyorosy and Smith, 2005; Phillips, 2009; Drummond et al., 2015).

To facilitate the process of capturing, measuring and valuing the costs, a costing model was developed in Microsoft Excel worksheet as shown in Table

5-1 to Table 5-7. In this model, activity logs capturing the time spent on the WIAT study by the FCS staff in different grades were recorded and the physical units of resources used in the physical and social intervention were measured directly and valued using their actual or unit cost. To ensure efficient and consistent reporting of costs in all the intervention sites, the costing model was not modifiable except for the input cells for the pay rate and description of activities.

It is important to highlight that the recording of the cost activities was done as soon as the activities were undertaken or costs were incurred. This helped to minimize recall bias which can result from the lengthy period taken before recording the details (McIntosh et al., 2010; Evers et al., 2015). Subsequently, from time to time, the Excel worksheet was updated for the entire period of the intervention. The following Table 5-1 to Table 5-7 show the valuation of the resources used in the WIAT study.

Internal costs related to time used by the FCS' staff in the physical and social intervention

Physical intervention

Type of site	May-13	Jun-13	Jul-13	Aug-13	Sep-13	Oct-13	Nov-13	Dec-13	Jan-14	Feb-14	Mar-14	Apr-14	**Total**
Haugh Hill/Pollok	£773	£586	£1,502	£2,245	£2,164	£664	£0	£0	£3,119	£401	£288	£321	**£12,060**
Linwood	£1,391	£2,738	£1,767	£1,203	£1,126	£0	£1,060	£729	£3,359	£563	£1,060	£155	**£15,150**
Mayfield	£1,148	£1,481	£3,399	£1,215	£563	£1,965	£1,789	£420	£1,271	£332	£332	£1,023	**£14,936**
													£42,146

Social intervention

Type of site	Apr-14	May-14	Jun-14	Jul-14	Aug-14	Sep-14	Oct-14	Nov-14	Dec-14	Jan-15	Feb-15	Mar-15	**Total**
Haugh Hill/Pollok	£560	£369	£787	£506	£42	£146	£587	£291	£635	£0	£0	£0	**£3,922**
Linwood	£575	£205	£5,303	£326	£2,452	£6,159	£12,094	£0	£3,429	£0	£0	£1,482	**£32,024**
Mayfield	£0	£652	£636	£1,802	£719	£1,199	£2,166	£1,719	£487	£476	£597	£1,603	**£12,052**
													£47,998

Table 5-1: Costing of FCS's staff time for the delivery of physical and social intervention

External costs for contracts for the physical intervention

Mayfield

Job	**Details**	**Net Cost**	**VAT**	**Total**
Tree works	Tree felling to clear path route/ thinning along path route/ chipping of material	£ 6,749.00	£1,349.80	£ 8,098.80
Pre-operation site check	Contract-wildlife site check by Peak Ecology Ltd	£ 510.00	£ 102.00	£ 612.00
Path works	Construction of 990m of new path/installation of stone set entrance feature	£27,252.00	£5,450.40	£32,702.40
Bat survey	Bat survey /supervision felling of dangerous tree	£ 350.00	£ 70.00	£ 420.00
Entrance posts	16 plain entrance posts	£ 397.00	£ 79.40	£ 476.40
Entrance post routing	Site names routed onto 11 of the entrance posts	£ 616.00		£ 616.00
Path repairs	Minor path repair/remedial drainage work following some wash-out	£ 2,040.00	£ 408.00	£ 2,448.00
Total		**£34,351.00**	**7,459.60**	**45,373.60**

Table 5-2: External costs for contracts for the physical intervention for Mayfield

Linwood

Job	Details	Net Cost	VAT	Total
Tree works	10 days with two-man squad/amenity thinning around path network	£ 4,334.00	£ 866.88	£ 5,201.28
Path works	Installation of 6 benches/ 2 picnic benches/gates/creation of 3 ponds, restoration of pond and provision of dipping platform 80m of new path construction/installation of 2 stone set features	£24,694.00	£4,938.80	£29,623.80
Signage	Leaflet (10,000) and 4 Pull up banners	£ 5,529.00	£1,105.80	£ 6,634.80
Miscellaneous	Centrewire Gates	£ 857.00	£ 171.40	£ 1,028.40
Other	Chipper Hire	£ 530.00	£ 106.00	£ 636.00
Natural play	Purchase of power tools for play trail creation	£ 562.40	£ 140.60	£ 703.00
Natural Play	Purchase and delivery of concrete including additional waiting time of mixer	£ 715.68	£ 178.93	£ 894.16
Natural play	Dumper/generator and breaker hire	£ 288.80	£ 72.20	£ 361.00
Natural play	Purchase of timber for natural play	£ 2,955.60	£ 738.90	£ 3,694.50
Replacement gate	Purchase of gate to replace damaged	£ 248.00	£ 62.00	£ 310.00
Total		**£40,714.48**	**£8,381.51**	**£49,086.94**

Table 5-3: External costs for contracts for the physical intervention for Linwood

Haugh Hill/Pollock				
Job	**Details**	**Net Cost**	**VAT**	**Total**
Tree works	Chipper Hire, light thinning around path corridor and path line felling	£ 530.00	£ 106.00	£ 636.00
Path works	New path creation 651m/installation of 2 benches and 2 picnic benches, fence repair, installation of entrance bollards, tarmac and stone set entrance feature	£16,143.92	£3,228.78	£19,372.70
Miscellaneous	Entrance bollards	£ 536.00	£ 107.20	£ 643.20
Total		**£17,209.92**	**£3,441.98**	**£20,651.90**

Table 5-4: External costs for contracts for the physical intervention for Haugh Hill/Pollock

External costs for contracts for the social intervention

Mayfield

Job	Details	Net Cost	VAT	Total
Family Fun Day	Healthy food demo provided by Almond Design	£ 500.00	£100.00	£ 600.00
Family Fun Day	Face painting -Aviatricks	£ 110.00		£ 110.00
Family Fun Day	Circus skills workshop-Flambeau Entertainment (Graham Hardie)	£ 300.00		£ 300.00
Photography Workshops	Three photography workshops -Becky Duncan Photography	£ 1,000.00		£ 1,000.00
Outdoor learning contractor	Outdoor learning sessions -Graham Hardie for Mayfield Nursery 7 days @ £180; 1 half day @ £100	£ 1,460.00		£ 1,460.00
Mayfield leaflet	Research and text for community leaflet on local walks and history, including the Kilns Woodland	£ 308.20		£ 308.20
Mayfield leaflet	The Kilns leaflet design and production	£ 440.00		£ 440.00
Total		**£ 4,118.20**	**£100.00**	**£ 4,218.20**

Table 5-5: External costs for contracts for the social intervention for Mayfield

Linwood

Job	Details	Net Cost	VAT	Total
Volunteer Sessions	6 Sessions lead by TCV	£ 1,500.00		£ 1,500.00
Newsletter	Lennon Design 300 newsletters	£ 148.00		£ 148.00
Photography	Becky Duncan - 2 photography workshops	£ 400.00		£ 400.00
Aspen Workshop	EADHA Enterprises - Aspen workshop -Xmas event 13th Dec	£ 300.00		£ 300.00
Storyteller	Daniel Allison - Storyteller - last event	£ 225.00		£ 225.00
Willow Sculptures	Trevor Leat -Willow Sculptures (central to the theme of the social intervention)	£ 1,500.00		£ 1,500.00
Workshop	Elisabeth/Laryna/Wupperman-FeltCraft workshop-4th event 21st Feb	£ 180.00		£ 180.00
Workshop	Rachan Design – Stone carving Workshop - 4th event 21 Feb	£ 375.00		£ 375.00
Workshop	Green Aspirations-Green woodworking - 4th Event	£ 450.00		£ 450.00
Circus performers	Delighters-Walkabout performance - Event 20th September	£ 425.00		£ 425.00
Photography	Scott Jone Images - Promotional Pics - Event 20th Sep	£ 200.00		£ 200.00
Workshop	Shirley Marzella-Willow Weaving workshop- 20 September	£ 180.00		£ 180.00
Workshop	Green Aspirations-Green woodworking -20 September	£ 450.00		£ 450.00

Workshop	Rachan Design – Stone carving Workshop - 20th September	£ 375.00	£ 375.00
Storyteller	Owen Pilgrim - Storytelling at event 20th of September	£ 225.00	£ 225.00
Photography	Promotional Shots - Event 21st of Feb	£ 240.00	£ 240.00
Felt craft workshop	Elisabeth Laryna Wupperman - Felt Craft workshop - Xmas event -13th Dec	£ 180.00	£ 180.00
Photography	Scott Jone Images - Promotional - Xmas Event - 13th Dec	£ 200.00	£ 200.00
Storytelling	Allison Galbraith - Storytelling at 2nd Event 11th October	£ 225.00	£ 225.00
Contract Ranger	Graham Hardie	£ 8,287.50	£ 8,287.50
Total		**£16,065.50**	**£16,065.50**

Table 5-6: External costs for contracts for the social intervention for Linwood

Haugh Hill/Pollock

Job	**Details**	**Net Cost**	**VAT**	**Total**
Volunteer	6 Sessions lead by TCV	£ 1,500.00		£ 1,500.00
Newsletter	Lennon Design - 150 newsletters	£ 111.00		£ 111.00
Workshops	Becky Duncan - 2 photography workshops	£ 400.00		£ 400.00
Workshop	EADHA Enterprises - Aspen - Xmas event 7th Dec	£ 300.00		£ 300.00
Toilet Hire	Loo King - 1st Event	£ 110.00		£ 110.00
Toilet Hire	Loo King - 2nd Event - 11th October	£ 110.00		£ 110.00
Toilet Hire	Loo King - Xmas Event 7th Dec	£ 160.00		£ 160.00
Toilet Hire	Loo King - 4th Event - 14th Feb	£ 110.00		£ 110.00
Photography	Scott Jone - Promotional Pics - 2nd Event 11th October	£ 200.00		£ 200.00
Photography	Scott Jone - Promotional photos - 4th Event - 14th Feb	£ 200.00		£ 200.00
Workshop	Shirley Marzella – Wreath-for Xmas event - 7th Dec	£ 87.81		£ 87.81
Circus	Delighters - Walkabout - 2nd Event 11th October	£ 425.00		£ 425.00
workshop	Rachan Design - Stonecarving - 4th Event - 14th Feb	£ 375.00		£ 375.00
Workshop	Elisabeth/Laryna/Wupperman/ 4th Event - 14th Feb	£ 180.00		£ 180.00

Workshop	Anna Liebmann - Willow weaving - 2nd event 11th Oct	£ 220.00	£ 220.00
Storytelling	Allison Galbraith - 2nd Event 11th October	£ 225.00	£ 225.00
Sculpture	Owen Pilgrim - Dragon Woodcarving	£ 1,500.00	£ 1,500.00
Workshop	Green Aspirations- woodworking - 2nd event - 11th Oct	£ 450.00	£ 450.00
Workshop	Green Aspirations - woodworking - 4th Event 14th Feb	£ 450.00	£ 450.00
Storytelling	Owen Pilgrim - Storytelling at 1st Event -	£ 225.00	£ 225.00
Storytelling	Owen Pilgrim - Storytelling at Xmas Event 7th Dec -	£ 250.00	£ 250.00
Contract ranger	Graham Hardie	£ 8,287.50	£ 8,287.50
Storytelling	Owen Pilgrim - Storytelling at 4th event 7th Dec -	£ 250.00	£ 250.00
Total		**£ 16,126.31**	**£ 16,126.31**

Table 5-7: External costs for contracts for the social intervention for Haugh Hill/Pollock

5.3 Results

It should be noted that all the cost data for the WIAT study presented in the tables above were collected in a two-year period using the costing model developed in Excel worksheet as explained earlier. This represents the time from wave one to wave two and wave two to wave three. These costs only relate to the intervention sites. There were no costs incurred in the control sites relating to resource use. The first year of collecting cost data was during the physical intervention stage while the second year was during the social intervention stage. The next section presents the overall results of the costing exercise in details according to the two broad categories of the FCS's costing models: internal and external costing.

5.3.1 Internal costs

These costs related to various activities regarding the implementation of the physical and social intervention. They were recorded in terms of time spent on the study by the FCS's members of staff which was multiplied by the pay rate per day for various grades as presented in Table 5-1. These were computed monthly for each intervention site. This resulted in the total of £90, 144 for all the intervention sites. This amount consists of £42, 146 for the physical intervention for the three sites and £47, 998 for the social intervention for the three sites as shown in Table 5-1. These costs translate to monthly average cost of £14,049 (SD £1,725) for the physical intervention and £15,999 (SD £14,461) for the social intervention. The SD (standard deviation) in this case, is an estimate of how monthly costs of the physical and social intervention varied from the average monthly cost within a year. The SD estimate helps in the calculation of standard errors for the cost estimate when fitting the distribution in the probabilistic sensitivity analysis (PSA) in chapter six.

5.3.2 External costs

Turning to the external costs, they consisted of external contracts or the use of the FCS internal rates per day for all resources external to FCS for the implementation of the physical and social intervention. The results in Table 5-2 to Table 5-7 show that these external costs were £115, 113 for the physical intervention for all the three sites (Mean £38, 371, SD £15,457) and £36,410 for the social intervention for the three sites (Mean £12, 137, SD £6,858). The total of external costs becomes £151, 523 for both the physical and social intervention for the three sites. The SD estimates for the external costs are also used in the calculation of standard errors for the PSA in chapter six as mentioned earlier.

Table 5-8 below presents a summary of both the internal and external costs for the delivery of the physical and social intervention of the WIAT study:

	Physical intervention		Social intervention		Total	
Internal costs	£	42,146	£	47,998	£	90,144
External costs	£	115,113	£	36,410	£	151,523
Total	**£**	**157,259**	**£**	**84,408**	**£**	**241,667**

Table 5-8: Internal and external cost of the WIAT project

It can be noted that the external costs were higher than the internal costs (£151,523 and £90, 144 respectively). This implies the nature of the physical and social intervention works which were mostly outsourced rather than done by the FCS's staff. It is also noted from the results in Table 5-8 above that the physical intervention had higher costs compared with the social intervention (£157, 259 and £84, 408 respectively).

A summary of all the costs for each intervention site in wave two and three are presented in Table 5-9 below:

Total cost of delivering the WIAT programme

Intervention site	Description of cost	Wave/Year 2 physical-post physical intervention	Wave/Year 3 social-post social intervention	Total
Mayfield	**Internal costs**			
1	FCS time-physical intervention	£14,936		£14,936
2	FCS time-social intervention		£12,052	£12,052
	External costs			
1	physical intervention contracts	£45,374		£45,374
2	social intervention		£4,218	£4,218
Subtotal		£60,310	£16,270	£76,580
Linwood	**Internal costs**			
1	FCS time-physical intervention	£15,150		£15,150
2	FCS time-social intervention		£32,024	£32,024
	External costs			
1	Physical intervention contracts	£49,087		£49,087
2	Social intervention		£16,066	£16,066
Subtotal		£64,237	£48,090	£112,327
Haugh hill/Pollock	**Internal costs**			
1	FCS time-physical intervention	£12,060		£12,060
2	FCS time-social intervention		£3,922	£3,922
	External costs			
1	Physical intervention contracts	£20,652		£20,652
2	Social intervention		£16,126	£16,126
Subtotal		£32,712	£20,048	£52,760
Total cost		**£157,259**	**£84,408**	**£241,667**

Table 5-9: Costing of resource use for the WIAT programme

Considering the costs of resource use for all the intervention sites, it can be seen in Table 5-9 above that the total cost of delivering the WIAT study was £241, 667 for all the three intervention sites.

Having estimated the total cost of the WIAT study, it is essential to point out that it is the average cost of the WIAT programme per individual that is of interest to the economic evaluation in chapter six (McIntosh et al., 2010; Drummond et al., 2015). The cost-utility analysis (CUA) requires the average cost of the intervention as an input in assessing cost effectiveness.

To estimate the average cost for the WIAT programme, the total cost of WIAT study is divided by the study population, thus, the eligible population of the WIAT project to whom the interventions might have an effect estimated at (n=20,472) (FCS, 2011a; FCS, 2015).

5.4 Discussion and conclusion

So far, this chapter has discussed how the cost of resources used in the development and implementation of the WIAT study were identified, measured, and valued. The aim of the chapter was to address the first research question of this thesis which sought to assess the costs attributed to the delivery of the WIAT study. The estimated average cost of the WIAT programme would be used as an input in the cost-utility analysis in chapter six.

Overall, the results showed that physical intervention costs were almost twice the social intervention costs. Notably, Linwood site had the highest amount of costs. This was partly because of the additional funds that were provided by the Renfrew city council to support the initiative as such there was more spending. The other reason was the use of a contract ranger for the delivery of the social interventions which inflated the cost of the social intervention.

In general, the large variations in the costs for the internal and external costs in the three intervention sites as depicted by the SD could be attributable to general differences that existed in the intervention sites. More work was done in some months, hence higher costs than in other months and some intervention sites required more work to be done than others.

During the measurement of the resources used, it was explained that the use of self-reported activity logging approach by the FCS staff involved in the project using the Excel worksheet costing model was practical and transparent. However, the decision as to which activity to include or exclude depended solely on the staff concerned which may have brought in some uncertainty (Mogyorosy and Smith, 2005). Since the average cost estimate is used in the cost-utility analysis in the next chapter, the uncertainty surrounding this cost value is explored through a probabilistic sensitivity analysis (PSA) which is discussed in chapter two.

The top-down approach used in measuring the cost of resource use could also be problematic in that it is always done in retrospective. This being the case, inevitably, it is prone to some omissions and bias in the recording of data. In addition, as noted earlier, the top-down approach uses high-level summaries such that it cannot group costs into direct and indirect cost categories (McIntosh et al., 2010; Drummond et al., 2015).

As discussed earlier in the chapter, not all relevant costs were captured for inclusion in the economic evaluation. It was difficult to identify and assign a unit cost to the community involvement in the study, for example. It was only feasible to capture and value the cost of resource use related to the Forestry Commission Scotland's staff time and physical resources for the delivery of the physical and social intervention. This is a limitation.

Despite these limitations, the results of this costing exercise are used as an ingredient in the cost-utility analysis framework which informs the cost-effectiveness modelling to determine the worthiness of the programme in chapter six.

It can also be noted from the costing exercise above that the costs of resource use for the WIAT study occurred at different times, hence the need for discounting to make the estimates comparable over different time periods. The discounting is also done in the cost-effectiveness modelling presented in chapter six. Issues relating to the appropriate discounting method and discount rate in economic evaluations have been discussed previously in chapter two.

Chapter 6: The WIAT intervention impact and cost-utility analysis (CUA)

6.1 Introduction

This chapter intends to address the third and fourth research questions of the thesis: to establish the impact of the WIAT intervention; and conduct a cost-utility analysis (CUA), respectively. It discusses how the impact of the WIAT intervention was determined and how the CUA was undertaken to establish the worthiness of the intervention in terms of value for money to aid resource allocation decision-making. As discussed in chapter two, CUA is a special form of cost effectiveness analysis where the numerator of the incremental cost-effectiveness ratio (ICER) is a measure of cost and the denominator is the quality-adjusted life year (QALY) measure. It was also noted in chapter two that the QALY measure consists of quantity and quality of life which is a utility. However, it is essential to note that the term "cost-utility analysis" refers to the use of QALY as a measure of utility.

The focus of this chapter is the health-related quality of life (HRQoL) as measured by the EQ-5D questionnaire. The EQ-5D questionnaire was included in the wider WIAT study to allow the derivation of utilities which are used in the CUA to calculate QALYs. As indicated previously in chapter two, QALY is a conventional measure of health outcomes in standard economic evaluations. The advantage of using QALYs is that they permit comparisons to be made across different health care interventions. For example, the relative value of health outcomes of the WIAT intervention can be compared with that of a vaccine intervention.

In chapter two, it was stated that the EQ-5D questionnaire consists of two tools that measure and provide different information about health: first, is the EQ-5D descriptive system with five dimensions which captures health state information as described by the dimensions; and second, is the visual analogue scale (VAS) which captures the overall health rating of individuals on that day of the survey. Appendix 1, questionnaire item F12 to F16 and F17 shows the EQ-5D descriptive system and the VAS, respectively.

In this thesis, the VAS analysis is only used to make comparisons on the effect of the WIAT intervention on the health-related quality of life with the EQ-5D descriptive system of the five dimensions. It does not, however, add any value to the cost-utility analysis which only focuses on, and use utilities derived from the EQ-5D descriptive system.

The five dimensions of the EQ-5D descriptive system translate into five health states for each respondent in the following order: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. As discussed in chapter two, there are currently two versions of the EQ-5D descriptive system: the old version with three levels (3L); and the new version with five levels (5L). In both versions, the health states are reported as an index on the questionnaire responses with 11111(full health) for both versions, and 33333 and 55555 for worst health for the 3L and 5L version, respectively. This index is then used to derive utilities from the pre-determined values sets obtained from the general public (Dolan, 1997).

This chapter begins with a section on methods which discusses the study design of the WIAT study. Understanding the WIAT study design is important because of the implications on data collection and analysis. The same section discusses how the missing data was dealt with, and how the impact of the WIAT intervention was established. This is followed by the cost-utility analysis. The chapter then proceeds to present the empirical analysis of the WIAT study which includes the results of the impact of the WIAT intervention and its cost-effectiveness analysis. After this, a general discussion and conclusion of the chapter follows.

6.2 Methods

6.2.1 Study design

As explained in chapter three, the wider WIAT study is a natural experiment with an intervention and control group. Its data included the EQ-5D responses which were collected in face-to-face computer-assisted personal interviews (CAPI) in a cross-sectional survey at wave one (baseline), wave two after the physical intervention, and wave three after both the physical and social intervention (Silveirinha de Oliveira et al., 2013).

Before any data were collected, ethical approval was sought from the University of Edinburgh, Edinburgh College of Art Research Ethics and Knowledge Exchange Committee (Silveirinha de Oliveira et al., 2013). The reason for seeking ethical approval from this committee was because the principal investigator for the wider WIAT study is based at the University of Edinburgh. The approval was granted with reference number ref. 19/06/2012.

When the data were collected and set-up, it turned out that some respondents were present in all the three waves; others were only present in one or two waves. This meant that there was an unbalanced panel data, in addition to the cross-sectional data. The unbalanced panel data followed up the same respondents who were in at least two waves from wave one. Even though the focus of the analysis is at population level, taking advantage of this unbalanced panel data could provide a base case analysis alongside the cross-sectional data analysis even though the small sample size could be a limitation.

As was pointed out in chapter three, one of the methodological challenges on linking green spaces and well-being of individuals relates to methods used to measure exposure. The concern is that spatial position of individuals within green spaces does not necessarily imply contact with the woods. However, the only way the health benefits of green spaces including woodlands can be realised is through contact with the woods which can either be direct through visits or views or indirect through virtual reality, films and photographs (Hartig et al., 2014). In determining the impact of the WIAT intervention, it was, therefore, important to consider contact with the woods as an important interaction term in the model because being in the intervention or control groups does not imply contact with the woods.

With respect to the EQ-5D data collection, both the old and new versions of the EQ-5D questionnaire were used. The 3L version was used in wave one while the new 5L version was used in waves two and three. The duration between waves was 12 months.

The required sample size for the wider WIAT study was calculated at 2,100 respondents at each of the three waves, which makes 1,050 respondents per intervention or control group. This sample size was determined by a power calculation based on a study by Stigsdotter et al. (2010), to be able to detect gender differences in the effect of the intervention (Silveirinha de Oliveira et al., 2013). An external market research company known as Progressive Partnership Ltd was used to collect all the data at the three waves.

There was a total of 5,460 observations for both the intervention and control group as shown in Table 6-1 below:

Wave	**Intervention**	**Control**	**Total**
1 (baseline)	1,061	1,056	2,117
2	740	932	1,672
3	816	855	1,671
Total	2,617	2,843	5,460

Table 6-1: Total number of observations for the wider WIAT study

Of these 5,460 observations, 1,361 observations formed the unbalanced panel which has been described earlier in this chapter.

6.2.2 Dealing with missing data

The extent of missing data across all variables in the wider WIAT study was minimal and ranged from 0% to 4.4% (Elizalde, 2016). The EQ-5D data had only 17 missing responses for all the three waves (0.3%). Despite having very few missing data, the whole WIAT data-set was multiply-imputed by Elizalde (2016) using a multivariate imputation by chained equations regression (MICE) approach (Azur et al., 2011). This approach fills in the missing values multiple times based on observed values to account for the statistical uncertainty in the imputation. In the case of the WIAT data-set, 10 sets of simulated values to complete the missing values were used (Elizalde, 2016).

6.2.3 The impact of the WIAT intervention

The use of both the 3L and 5L versions of the EQ-5D questionnaire was potentially problematic in terms of analysis because it is a known fact that the 3L and the 5L versions of EQ-5D questionnaire are two different instruments which result in different profile vectors, hence their utilities might not be directly comparable (Janssen et al., 2008; Hernández-Alava and Pudney, 2016). This being the case, it was important to ensure that the utilities calculated from responses of the EQ-5D 5L index profiles in wave two and three were consistent with the responses from the EQ-5D 3L index profiles in wave one or vice versa.

Using the EQ-5D health states for the EQ-5D 3L version for wave one, utilities were derived using the most commonly used approach of using the predetermined utility value sets for each health profile based on the UK population (Dolan, 1997). In case of reported deaths, a utility value of zero is normally assigned (Morris et al., 2016). However, there were no reported deaths in the WIAT study. Value sets for utilities for the new EQ-5D 5L version are only available for the representative English population based on their preferences (Devlin et al., 2016). These are, however, not yet available for the wider UK population. It remains unclear if the preferences of the English population are consistent with those of the wider UK population to use these value sets to the Welsh or Scottish population, for example.

Interestingly, the 5L health states provide 3125 indices which are distributed on a scale of (-0.594, 1) (with 55555 = -0.594; 11111 = 1) (van Hout et al., 2012) and the 3L version provides 243 indices which are distributed on the same scale of (-0.594, 1) (with state 33333 = -0.594; state 11111=1(Dolan, 1997)). These distribution similarities provide comparability of the two versions of the EQ-5D although their means and median are different.

Based on the distribution similarities and in the absence of the utility value sets for the 5L version for the wider UK population, the EuroQol group that developed the EQ-5D tool recommend the crosswalk approach, also known as mapping, to derive utilities for the 5L from the 3L EQ-5D version (van Hout et al., 2012). This approach is based on a response mapping that estimated the relationship between the responses to the EQ-5D 3L and 5L descriptive systems. The 3L responses were predicted using frequencies of cross-tabulating the responses on the 3L and 5L and subsequently, the transition probabilities associated with the 3L were applied to their index value to obtain value sets for the 5L (van Hout et al., 2012). Table 6-2 below shows the transition probabilities from the 5L version to the 3L version per dimension:

Transition probabilities: from the 5L system to the 3L system per dimension

5L→3L	Mobility	Self-Care	Activity	Pain	Anxiety
1→1	1.00	1.00	1.00	1.00	1.00
1→2	0	0	0	0	0
1→3	0	0	0	0	0
2→1	0.18	0.17	0.20	0.20	0.21
2→2	0.82	0.83	0.80	0.80	0.79
2→3	0	0	0	0	0
3→1	0	0	0	0	0
3→2	1.00	1.00	1.00	1.00	1.00
3→3	0	0	0	0	0
4→1	0	0	0	0	0
4→2	0.93	0.76	0.67	0.60	0.51
4→3	0.07	0.24	0.33	0.40	0.49
5→1	0	0	0	0	0
5→2	0	0	0	0	0
5→3	1.00	1.00	1.00	1.00	1.00

Table 6-2: Transition probabilities 5L to 3L version. Source: (van Hout et al., 2012).

These transition probabilities are applied in the crosswalk approach to calculate the EQ-5D utilities using Stata. Similarly, there is a Microsoft Excel tool known as the "EQ-5D-5L Crosswalk Index Value Calculator" developed by the EuroQol group which calculates the crosswalk index values for the EQ-5D-5L dimension scores. Using both Stata and Microsoft Excel approaches yielded the same results of utilities.

Having derived the EQ-5D utilities for all the three waves, it was essential to firstly determine the impact of the WIAT study. This could simply be determined by comparing the pre-and post-intervention utilities. However, this could only be valid in the absence of confounding which could mix up the effects of the intervention with other factors unrelated to the intervention (Abadie, 2005; Remler and Van Ryzin, 2010; Pearce and Greenland, 2014).

As a natural experiment, the WIAT study had a control group which could act as a counterfactual to determine the casual effect of the intervention. The counterfactual establishes what would have happened in the intervention group in the absence of any intervention, to be sure about the actual effect (Abadie, 2005; Khandker et al., 2010; Gertler et al., 2011; Pearce and Greenland, 2014; Fricke, 2015). Even though this type of design controls for some confounding and provides the counterfactual through a control group, in reality, respondents in the intervention and control groups are probably non-equivalent at baseline which might still result in individual differences (Morgan et al., 2000; Campbell et al., 2007; Khandker et al., 2010). In this case, the individual differences can be adjusted for at the analysis stage.

<u>Difference-in-differences (DiD) approach</u>

There are many approaches that are used to establish the casual effect of an intervention such as the propensity score matching (PSM); the instrumental variables; and the regression discontinuity approach (Khandker et al., 2010; Kenkel and Suhrcke, 2011; White and Sabarwal, 2014). However, these approaches require additional statistical techniques to construct a comparison group to act as a counterfactual.

The control group, however, is already existent in the WIAT study. When this is the case, many studies recommend the use of a regression-based model known as the difference-in-difference (DiD) approach (Shadish et al., 2002; Krabbe and Weijnen, 2003; Richardson and Manca, 2004; Manca et al., 2005; Angrist and Pischke, 2008; Craig et al., 2011; Villa, 2012; Wooldridge, 2012; Pearce and Greenland, 2014; Schäfer et al., 2014; Gasparrini and Lopez Bernal, 2015; Ryan et al., 2015). This approach was, therefore, considered suitable to estimate the impact of the WIAT intervention given its design. This DiD approach identifies the casual effect of an intervention by contrasting the change in outcomes pre- and post-intervention, for the intervention and control groups (Griffin et al., 2010; O'Neill et al., 2016). Its implementation requires an intervention group; a control group; then a before and after period of an intervention, which were all present in the WIAT study (Bertrand et al., 2004; Khandker et al., 2010; Silveirinha de Oliveira et al., 2013; Botosaru and Gutierrez, 2015; Grabich et al., 2015; Ryan et al., 2015).

The DiD approach establishes the counterfactual through the control group with the following underlying assumptions: first, that the average change in outcome of the intervention group equals the average change in outcome of the control group in the absence of the intervention. This is also known as the ‘parallel’ trend assumption; and second, that the effects caused by time trends such as unexpected and unpredicted events between the control and the intervention group when without intervention do not vary with time. This is referred to as the ‘common shock’ assumption (Angrist and Pischke, 2008; Mora Villarrubia and Reggio, 2012; Lin and Hsu, 2014).

To implement the DiD approach through a regression model, a multilevel model was used because of the nesting in the WIAT data. This means that the WIAT data was hierarchical with three levels. Some respondents were present in all the three waves; some were in wave one and two only and others were in wave one and three only. This clearly implies a possible lack of independence of individual responses in the three waves, making the standard ordinary least squares (OLS) regression unsuitable (Albright and Marinova, 2010; Hamilton, 2012).

As noted earlier in the chapter, the multilevel model included an interaction term related to contact with the woods to determine the impact of the intervention on HRQoL, which can only happen when there is contact with the woods (Hartig et al., 2014; WHO, 2016). The following two self-reported questions in the WIAT main survey which asked whether respondents had visited local woodlands, parks or green spaces in the last year or two, formed the basis of this interaction variable and the positive responses implied that the respondents might have benefited from contact with nature:

B3. Have you visited <u>these local woodlands</u> in the <u>last year</u>?
SHOW CARD A

SINGLE CODE	CODE	ROUTE
Yes	1	Go to B4
No	2	Go to B10

D1. Have you visited local parks or green spaces in the <u>last 12 months</u>?
SHOW CARD A

SINGLE CODE	Code
Yes	1 – Go to Q D2
No	2 – Go to Q E1

In this case, the impact of the intervention on the HRQoL was determined by contrasting the interaction between type of site, wave and contact with woods with the interaction between type of site and wave. The following specification of the model was used:

$$eq5d_{it} = \beta_0 + \beta_1 Type\_site_i + \sum_{t=2}^{T} \beta_t Wave_t + \beta_2 contact_i + \delta_t\, Type\_site_i * Wave_t * contact_i + \sum_{c=i}^{N} \beta_3 X_{it} + \varepsilon_{it} \quad (8)$$

Where $eq5d_{it}$ is the outcome of interest for respondent *i* at a given wave. The individuals were observed in a pre-intervention *t=0* (wave one) and post intervention *t=1* (wave two) and *t=2* (wave three). Between these periods, the intervention group was exposed to the intervention. This is denoted as $Type\_site_i$=*1* if exposed to intervention and $Type\_site_i$=*0* for the control group. Being in contact with nature is denoted as $contact_i = 1$ if in contact with nature and $contact_i = 0$ if not in contact with nature. β_0 is the constant, β_1 represents the group effect of being in an intervention, β_2 represents the group effect of being in contact with nature, β_t represents time effect, $Wave_t$ is a dummy variable for wave two and three, δ_t from the interaction between type of site , wave and contact with nature in a given wave, c represents individuals, β_3 represents the value of individual characteristics effect, X_{it} are individual characteristics of the respondents and ε_{it} is the error term for unobservable factors.

Of interest in this DiD regression model is mainly the δ_t. Thus, the interaction between the type of site, wave and contact with the woods dummy variables. The three-way interaction in this case has several different interpretations. In this case, it could be seen as representing how type of site modifies the wave and contact with nature interaction, or how contact with nature modifies the type of sites and wave interaction, or as a piece of how contact with nature and type of site are jointly modified by the time effect (wave), or several other ways of putting this all together. So, the Difference in Difference (DiD) could not just be the co-efficient δ_t on the three-way interaction term. The best way to interpret δ_t in this study is to consider that a three-way interaction means that there is a two-way interaction of type of site (dummy variable of being in an intervention or control group) and wave (dummy variable of being in wave 1, 2 or 3) that varies across levels of being in contact with nature (dummy variable of being exposed to nature or not). The contrast approach in this case is used to tests for a two-way interaction effect (Type of site#Wave) at each level of contact with nature. This establishes the impact of the intervention in wave two for the physical intervention and wave three for both the physical and social intervention. The analysis of data for the multilevel model was done in Stata software version 12.1 (StataCorp, 2013) using '*mi estimate: xtmixed*' a multilevel regression Stata command for the imputed data.

The *mi estimate:* command runs estimations on each imputation separately. This means that 10 estimates were run for the WIAT data and the results of all estimations are then combined and displayed as output. The operator with two vertical lines towards the end of the command (||) indicates the beginning of the random effects specification without which results in a standard fixed effects regression model. The random effects output was used to calculate the interclass correlation coefficient which test the strength of the correlation among respondents in one nest and across the nests.

This whole statistical analysis using multilevel modelling allowed the determination of the change in HRQoL in the intervention group relative to the control group resulting from the WIAT intervention between wave one and two after the physical intervention; and wave one and three after both the physical and social intervention. After establishing the effect of the WIAT intervention on HRQoL, the next stage was to conduct a cost-utility analysis. The next section discusses how this cost-utility analysis was done in the context of the WIAT study.

6.2.4 Cost-utility analysis of the WIAT intervention

In general, it is important to address the question of whether an intervention is good value for money given the opportunity cost of the investment related to its delivery in the face of limited budgets. For this reason, a cost-utility analysis was carried out for the WIAT study using the cost of resource use estimated in chapter five and the utilities representing health-related quality of life (HRQoL) derived in this chapter from the EQ-5D descriptive system responses. The time horizon for the analysis was two years reflecting the time from wave one to wave three when the physical intervention and both the physical and social interventions were given. The next section returns briefly to costing to discuss how the mean cost of the WIAT intervention per individual was calculated.

Cost of resource use

The cost input required in the cost-utility analysis framework for the WIAT study is the incremental mean cost of the physical intervention and both the physical and social interventions per individual between the intervention and control group. For the WIAT study, there was no cost incurred in the control group in any of the waves, hence, the cost component only relates to the intervention group. This mean cost was estimated by dividing the total costs of the intervention consisting of the physical intervention and both the physical and social interventions by the WIAT study population. The study population was defined to include settlements of 20,472 people around intervention sites as previously explained in chapter two (FCS, 2011b). The mean cost for the WIAT physical intervention was, therefore, estimated at £7.68 and for both the physical and social intervention was estimated at £11.80.

<u>Health-related quality of life (HRQoL)</u>

As regards the HRQoL, the expected utilities for the intervention and control group were estimated using the multi-level regression-based predictive approach. Compared with the standard approach, this predictive approach has the advantage of providing the average HRQoL utilities of the population adjusted for baseline characteristics (Manca et al., 2005; Nixon and Thompson, 2005; Härkänen et al., 2013; Briggs et al., 2016). The predictive approach used the Stata *mimrgns* command for the imputed data to estimate expected utility.

The expected health effects (utilities) are from the representative sample of the WIAT study population surveyed with the EQ-5D tool, and in this case, those who were in contact with nature. Table 6 3 below shows the expected HRQoL utility scores with 95% CIs at wave one (baseline), wave two and three for both the intervention and control groups for the unbalanced panel data which was used for the base case analysis:

Intervention group

	Utility	Standard Error	P value	95% CI	
Wave 1	0.788	0.018	0.000	0.753	0.824
Wave 2	0.829	0.025	0.000	0.780	0.879
Wave 3	0.822	0.023	0.000	0.777	0.868

Control group

	Utility	Standard Error	P value	95% CI	
Wave1	0.809	0.016	0.000	0.778	0.841
Wave2	0.783	0.020	0.000	0.744	0.824
Wave3	0.843	0.024	0.000	0.800	0.890

Table 6-3: Expected utility scores for the intervention and control group

The cost-utility analysis involved comparing the incremental mean cost with the incremental expected utilities for the intervention group relative to the control group. Then, the individual-specific quality-adjusted life years (QALYs) were calculated using the commonly used area under the curve approach as shown in (9) (Richardson and Manca, 2004; Manca et al., 2005; Briggs et al., 2016). This approach calculates QALYs as the product of the time difference, in this case waves, and the average of the two measurements of individual utilities (Matthews et al., 1990). Thus, for WIAT study expected utilities U_1, U_2 and U_3 at times W_1, W_2 and W_3, the formula for calculating QALYs for either the intervention or control group becomes:

$$(W_2 - W_1)\,(U_1\ +\ U_2)/2 + (W_3 - W_2)\,(U_2\ +\ U_3)/2 \qquad (9)$$

Where U_1, U_2 and U_3 are expected utilities for wave one, two and three, respectively and W_1, W_2 and W_3 represent wave one, two and three, respectively. The area under the curve approach assumes linear interpolation in the change in utility scores between time intervals (Manca et al., 2005).

Both the costs and QALYs were discounted to account for the differential timings (McIntosh et al., 2010; Morris et al., 2012). The rate employed for discounting was 3.5% as per NICE guidelines (NICE, 2013). The cost-effectiveness analysis was conducted from the societal perspective because the cost and outcomes of the WIAT intervention would affect the society in general. For example, the cost for delivering the intervention implies other benefits to the society forgone. On the other hand, the outcomes of the intervention would accrue directly or indirectly to anyone in society.

The main outcome of the cost-utility analysis was the incremental cost-effectiveness ratio (ICER) expressed as the cost per QALY gained and net monetary benefit (NMB) as expressed in (5) and (6), respectively, in chapter two (NICE, 2013).

Sensitivity analysis

It is generally acknowledged that both the mean cost and expected utility parameters used in cost-utility analysis model are not known with certainty. For this reason, it is recommended that uncertainty surrounding these parameters should be quantified to help the decision-making process of whether to adopt the intervention (Briggs, 2000; Baio and Dawid, 2011; NICE, 2013; Wolowacz et al., 2016). The concept of uncertainty and approaches of dealing with it in economic evaluations are discussed in chapter two. The uncertainty surrounding the cost and effect for the WIAT study was explored using a commonly used approach of probabilistic sensitivity analysis (PSA) also known as parametric bootstrapping (Briggs et al., 2006; Gray et al., 2010). The PSA was undertaken using 5 000 Monte Carlo simulations which is a process of repeatedly creating random data, to sample from the probability distribution assigned to the cost and utility parameters through bootstrapping. Briggs et al. (2006) recommend any number above 1,000 times to be acceptable for the estimation of 95% confidence intervals from the bootstrapped replicates of data using approaches such as the percentile method. This method uses the lower and upper percentile (0.025 and 0.975), respectively, from the simulations to disperse uncertainty in the parameters.

The probability distributions assigned to the parameters are defined to reflect the nature of the data as discussed in chapter two. For example, a Gamma distribution was used to model uncertainty in the cost parameter because costs are constrained between zero and positive, and Gamma distribution has the same property, hence suitable (Briggs et al., 2006; Edlin et al., 2015). The alternative probability distribution for cost parameter is the LogNormal which is often employed in regression analyses and results in the same outcome with the Gamma distribution when applied to a sufficiently large sample (Edlin et al., 2015). On the other hand, Beta distribution was used to model the effect of the intervention because utility is bounded between zero and one (Briggs et al., 2006; Edlin et al., 2015). The Beta distribution can be used confidently when the expected value is close to 1 and the variance is small (Edlin et al., 2015). It can be noted that the WIAT utility data have these qualities as depicted in Table 6.3 above.

To compute the probability distributions, standard errors are an important input. The standard errors for the cost parameter were computed from the estimate (standard deviation) of how individual observations of cost data varied in the costing model using the following formula:

$$SE = \frac{SD}{\sqrt{N}} \tag{10}$$

Where *SE* is the standard error, *SD* is the standard deviation and *N* is the cost data observations. Whereas the standard errors for utility were generated from the predictive margins of the multilevel regression as shown in Table 6.3.

It was assumed that the main parameters of the CUA model were uncorrelated to each other. Correlation, in this case, means that the information that determines the value of one parameter, partially determines the value of the other parameter (Edlin et al., 2015). However, correlation is likely to be present when the model involves transitions between states, especially if there are elements of severity within the model. For example, a severe ill health would imply higher costs and lower health-related quality of life. When this is the case, correlations are usually incorporated into a PSA using Cholesky decomposition, a mathematical technique, that considers the impact of one variable on the next variable from the random draws in the PSA (Briggs et al., 2006; Edlin et al., 2015).

All the bootstrapping was performed in Microsoft Excel and implemented using a Microsoft Excel macro. The results of all simulations were combined to give an overall incremental cost-effectiveness ratio (ICER) results.

The results of bootstrapped pairs of incremental mean cost and incremental QALYs were presented using the cost-effectiveness (CE) plane and cost-effectiveness acceptability curve (CEAC). The CE plane depicted the point in the quadrant where each bootstrapped pair of the incremental cost and incremental QALY is positioned on the vertical and horizontal axis representing incremental mean costs and incremental QALYs, respectively. On the other hand, the CEAC showed the probability of the WIAT intervention being cost-effective at given ranges of willingness to pay values compared with the option of doing nothing. The percentile approach was used to estimate a good approximation of the 95% confidence interval by using the 2.5th and the 97.5th percentiles of the bootstrap samples. The next section presents the results of the analyses in detail.

6.3 Results

Table 6-4 below shows the descriptive analysis of the socio-demographic characteristics of the respondents in all the three waves of the WIAT study:

	Intervention (*n*=3165)		**Control** (*n*=3152)	
	Number	%	Number	%
Age range				
16-24	281	9%	237	8%
25-34	526	17%	522	17%
35-44	461	15%	484	15%
45-54	588	19%	580	18%
55-64	438	14%	489	16%
65-74	478	15%	512	16%
75+	390	12%	318	10%
Missing	3	0.1%	10	0.3%
Gender				
Female	1912	60%	1900	60%
Male	1253	40%	1252	40%
Social class				
High managerial	97	3%	140	4%
Supervisory/clerical/junior managerial	577	18%	613	19%
Skilled manual worker	580	18%	587	19%
Semi-unskilled manual worker	698	22%	726	23%
Pensioner/casual/unemployed	1144	36%	1036	33%
Missing	69	2%	50	2%
Highest qualification				
No qualification	1372	43%	970	31%
Level 1	884	28%	1205	38%
Level 2	492	16%	405	13%
Leve 3	223	7%	314	10%
Level 4	192	6%	252	8%
Missing	2	0.1%	6	0.2%
Disability				
No	2729	86%	2786	88%
Yes	426	13%	351	11%
Missing	10	0.3%	15	0.5%
Perceived income				
Living comfortably	773	24%	938	30%
Coping	1611	51%	1593	51%

Finding it difficult	527	17%	411	13%
Finding it very difficult	181	6%	110	3%
Missing	73	2%	100	3%
Children under 16				
No	2287	72%	2263	72%
Yes	876	28%	883	28%
Missing	2	0.1%	6	0.2%
Car ownership				
No	1392	44%	1093	35%
Yes	1773	56%	2059	65%
Smoking				
Never smoked	1177	37%	1635	52%
Smoked in the past	809	26%	620	20%
Currently smoke	1151	36%	879	28%
Missing	28	0.9%	18	0.6%

Table 6-4: Characteristics of the respondents in the intervention and control groups

6.3.1 The impact of the WIAT intervention

Two models were estimated with an inclusion of an interaction term for contact with nature: the unbalanced panel analysis in which respondents were in at least two waves including the first wave (n=1,361) and the cross-sectional analysis for the three waves (n=5,460). The unbalanced panel analysis, in this case, was used for base case analysis. It has the advantage of following up the same respondents in at least two waves from wave one. Therefore, it can provide the true effect of the intervention at both individual and population level. However, the disadvantage of the unbalanced panel analysis, is the small sample size. In contrast, the cross-sectional analysis has the advantage of using a larger sample size and capable of providing results at population level. Table 6.5 below presents the unadjusted results of the DiD approach using the multilevel regression models on the imputed data:

Unbalanced analysis

EQ-5D Utility	Coef.	Std. Err.	P value	95% CI Lower	Upper
Intervention	-0.058	0.031	0.065	-0.119	0.003
Wave					
Wave 2	-0.010	0.032	0.762	-0.073	0.054
Wave 3	-0.034	0.030	0.256	-0.092	0.025
Type_site#Wave					
Intervention#wave 2	0.000	0.053	0.999	-0.104	0.104
Intervention#wave 3	-0.013	0.047	0.786	-0.106	0.080
Nature's visits					
ExpNat	0.077	0.029	0.008	0.020	0.135
Type_site# Nature's visits					
Intervention#ExpNat	0.009	0.043	0.833	-0.076	0.094
Nature's visits #Wave					
ExpNat #wave 2	-0.055	0.046	0.233	-0.146	0.035
ExpNat #wave 3	0.040	0.047	0.394	-0.052	0.133
Type_site# Nature's visits #wave					
Intervention #wave 2# ExpNat	0.120	0.074	0.104	-0.024	0.264
Intervention #wave 3# ExpNat	-0.015	0.070	0.825	-0.153	0.122
Constant	0.797	0.021	0.000	0.755	0.838

Cross-sectional analysis

EQ-5D Utility	Coef.	Std. Err.	P value	95% CI Lower	Upper
Intervention	0.004	0.01	0.77	-0.025	0.034
Wave					
Wave 2	0.029	0.01	0.07	-0.002	0.060
Wave 3	0.003	0.01	0.86	-0.028	0.033
Type_site#Wave					
Intervention#wave 2	-0.009	0.02	0.68	-0.053	0.034
Intervention#wave 3	-0.060	0.02	0.00	-0.103	-0.016
Nature's visits					

ExpNat	0.090	0.01	0.00	0.061	0.120
Type_site# Nature's visits					
Intervention#ExpNat	-0.006	0.02	0.76	-0.048	0.035
Nature's visits #Wave					
ExpNat #wave 2	-0.039	0.02	0.07	-0.081	0.004
ExpNat #wave 3	-0.001	0.02	0.97	-0.045	0.043
Type_site# Nature's visits #wave					
Intervention #wave 2# ExpNat	0.030	0.03	0.35	-0.033	0.092
Intervention #wave 3# ExpNat	0.031	0.03	0.33	-0.032	0.093
Constant	0.811	0.01	0.00	0.790	0.833

Table 6-5: Unadjusted analysis of unbalanced panel and cross-sectional data.

In Table 6.5 above, starting with the unbalanced panel analysis, the first coefficient represents the estimated mean difference in Health-Related Quality of Life (HRQoL) between the intervention and control group prior to the intervention. It is the baseline difference in utilities that existed between the intervention and control groups before any of the WIAT intervention was introduced. There were no statistically significant mean differences between the intervention and control group -0.058, ρ=0.065 (CI -0.119-0.003).

The coefficients for Wave 2 and 3 are the expected mean change in HRQoL from baseline (Wave 1) to Wave 2, after the physical intervention and from baseline to Wave 3 after both the physical and social intervention, resulting from passage of time and unrelated factors to the intervention. This change was as close to zero as possible and statistically insignificant for both waves, from baseline to wave 2 (-0.010, ρ=0.762 CI -0.073-0.054) and from baseline to wave 3 (-0.034, ρ=0.256, CI -0.092-0.025).

Of interest are the coefficients for the interaction between type of site and wave, and type of site, wave and contact with nature. The rest of the coefficients relate to the three-way interaction and a stand-alone interpretation of their coefficients is not intuitive as noted earlier. Given the three-way interaction, the impact of the WIAT intervention can be depicted by contrasting the three-way interaction of type of site, wave and contact with the two-way interaction of type of site and wave. This implies testing whether there is a difference between the three-way and the two-way interaction for those who had contact with nature.

The three-way interaction is statistically insignificant for both the unbalanced panel and the cross-sectional analysis. Table 6-6 below shows the effect of the intervention on contact with the woods after contrasting the three-way interaction of type of site, wave and contact with nature with the two-way interaction of type of site and wave to establish the effect of the intervention on contact with woods.

	Contrast	Standard Error	P value	95% CI	
Wave 2	0.086	0.044	0.050	-0.000	0.173
Wave 3	-0.024	0.045	0.593	-0.112	0.064

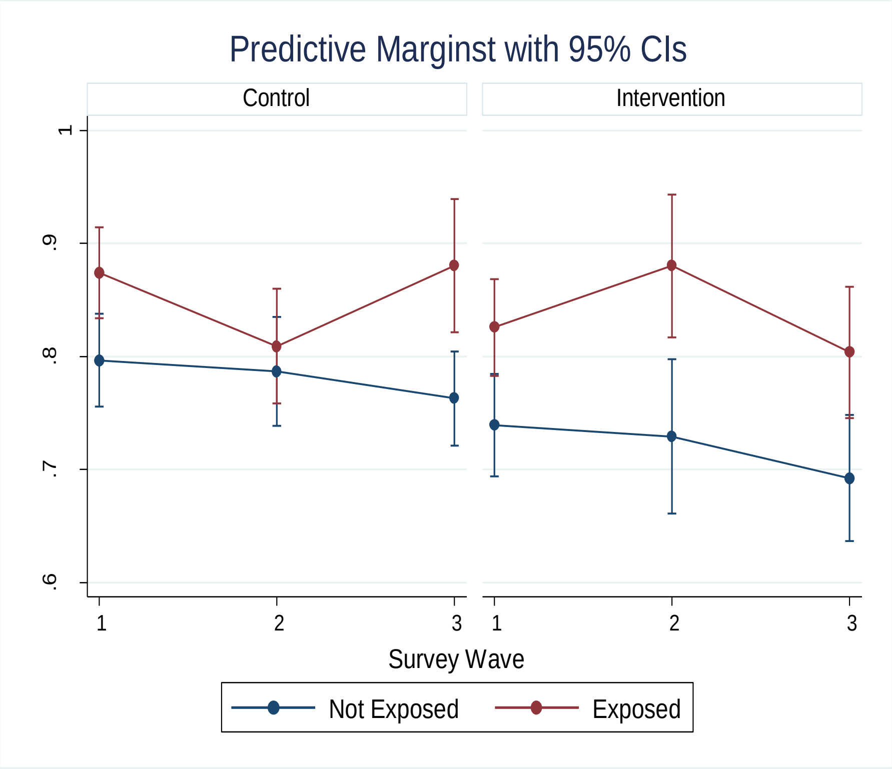
Table 6-6: The effect of the intervention on contact with woods for the unadjusted unbalanced panel analysis.

The results above show that the impact of the WIAT interventions in terms of HRQoL when individuals get contact with nature was statistically insignificant at 0.086, ρ=0.050 (CI -0.000-0.173) for Wave 2 after the physical intervention and in Wave 3, after both the physical and social intervention, the impact was as close to zero as possible, albeit insignificant. The same trend was found for the repeated cross-sectional analysis as shown in Table 6-7 below:

	Contrast	Standard Error	P value	95% CI	
wave 2	0.020	0.022	0.358	-0.023	0.063
wave 3	-0.026	0.022	0.232	-0.069	0.017

Table 6-7: The effect of the intervention on contact with woods for the unadjusted cross-sectional analysis.

These results are presented graphically in Figure 6-1 and Figure 6-2 for the unbalanced panel and cross-sectional analysis, respectively below:



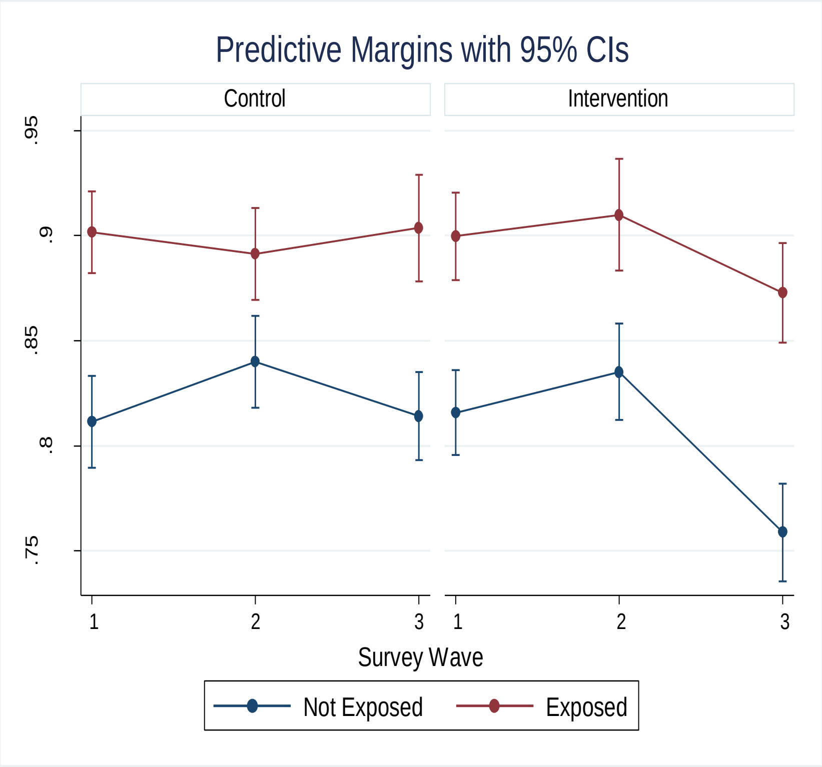


Figure 6-1: Mean change in utility for unbalanced panel analysis

Figure 6-2: Mean change in utility for cross-sectional analysis.

Table 6-8 below presents the results of the model adjusted for age, gender, social class, perceived income, distance band, working status, education, car ownership, life events, smoking, disability, and differences in site pairs for both the unbalanced panel and cross-sectional analysis.

EQ-5D utility score	Model 1: Adjusted unbalanced panel-Type of site* wave* contact: Coef.	Std. Err.	P value	95% CI		Model 2: Adjusted cross-sectional-Type site*wave*contact: Coef.	Std. Err.	P value	95% CI	
Intervention	-0.029	0.026	0.253	-0.080	0.021	0.001	0.012	0.938	-0.022	0.024
Wave 2	-0.029	0.025	0.240	-0.078	0.020	-0.002	0.012	0.871	-0.025	0.021
Wave 3	-0.027	0.023	0.242	-0.071	0.018	-0.005	0.012	0.655	-0.028	0.018
Contact with woods	-0.009	0.023	0.691	-0.054	0.036	0.005	0.011	0.646	-0.017	0.028
Type site*contact with woods										
Intervention*contact with woods	0.008	0.033	0.799	-0.057	0.074	0.005	0.016	0.769	-0.026	0.036
Type site*Wave										
Intervention*wave 2	0.019	0.040	0.635	-0.060	0.098	0.014	0.017	0.411	-0.019	0.047
Intervention*wave 3	0.000	0.036	0.990	-0.071	0.070	-0.016	0.016	0.331	-0.048	0.016
Contact with woods*Wave										
Contact with woods*wave 2	0.004	0.036	0.919	-0.066	0.074	-0.005	0.016	0.751	-0.037	0.027
Contact with woods*wave 3	0.061	0.037	0.097	-0.011	0.132	0.012	0.017	0.458	-0.020	0.045
Type site*contact*wave										
Intervention*wave 2*contact	0.047	0.057	0.403	-0.064	0.158	0.007	0.024	0.774	-0.040	0.053
Intervention*wave 3*contact	0.000	0.054	0.997	-0.105	0.106	0.014	0.024	0.551	-0.032	0.061
Constant	0.866	0.060	0.000	0.748	0.983	0.892	0.024	0.000	0.845	0.939

Table 6-8: Adjusted analysis of unbalanced panel and cross-sectional data

After contrasting the three-way interaction between type of site, wave and contact with woods with the two-way interaction of type of site and wave, the results are presented in Table 6-9 and Table 6-10, respectively below:

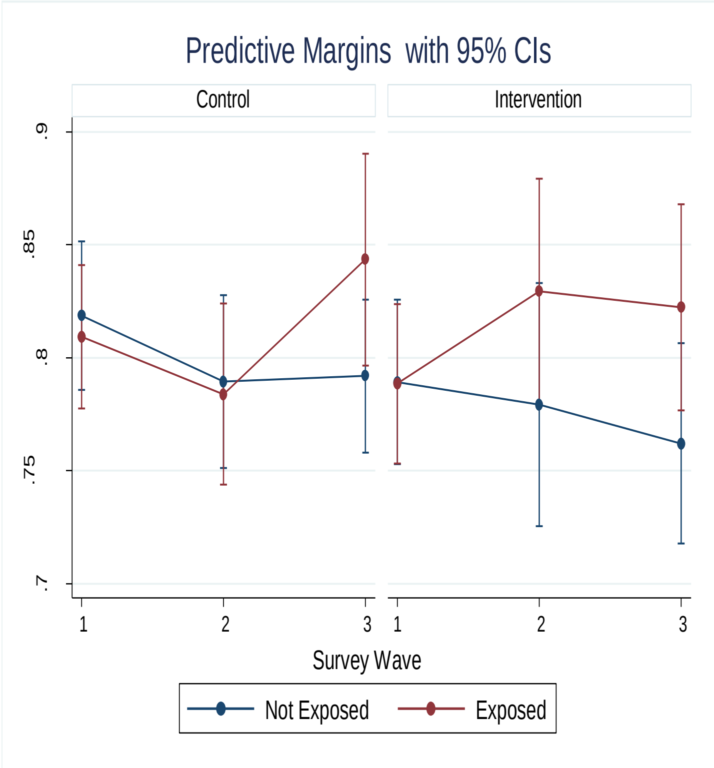
	Contrast	Std.Error	P value	95% CI	
wave 2	0.067	0.039	0.086	-0.009	0.143
wave 3	-0.000	0.040	0.995	-0.078	0.078

Table 6-9: The effect of the intervention on contact with woods for the adjusted unbalanced panel analysis.

	Contrast	Std.Error	P value	95% CI	
wave 2	0.021	0.017	0.227	-0.013	0.054
wave 3	-0.002	0.017	0.916	-0.035	0.032

Table 6-10: The effect of the intervention on contact with woods for the adjusted cross-sectional analysis.

The results in the tables above show that there was again no evidence of any statistically significant impact of the intervention for those individuals in contact with nature in both the unbalanced panel and cross-sectional analysis. Figure 6-3 and Figure 6-4 depict these results graphically:



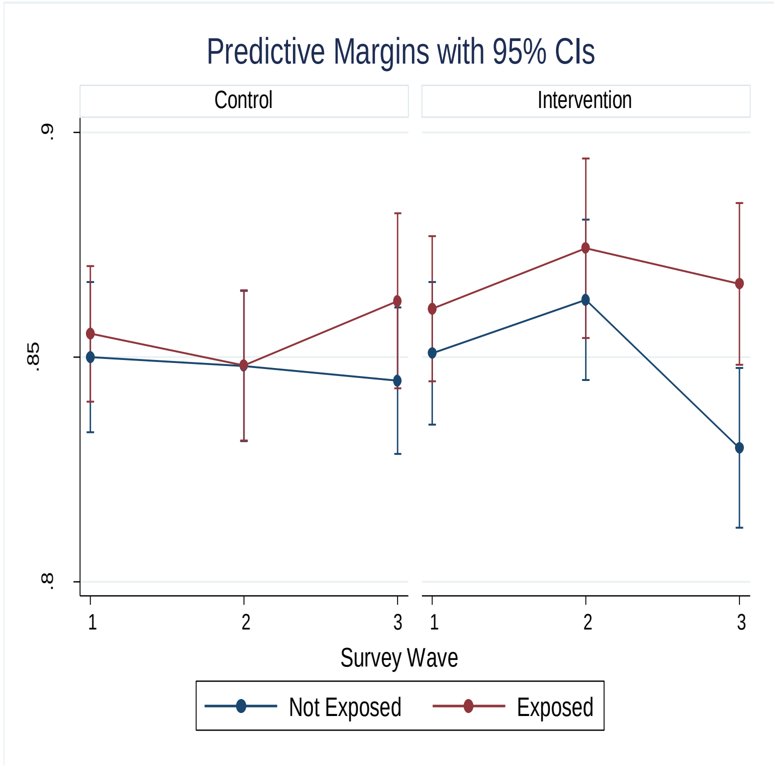


Figure 6-3: Mean utility for the adjusted unbalanced panel analysis

Figure 6-4: Mean utility for the adjusted cross-sectional analysis

Full details of these analyses are presented in Appendix 6. It can be seen from this appendix that the EQ-5D VAS analysis showed similar trends to the EQ-5D descriptive system of statistically insignificant outcomes for both the cross-sectional and unbalanced panel analysis.

As noted earlier, the multilevel regression modelling was used because the WIAT data was nested. Some respondents were in one wave only while others were either in all the three waves or two waves only. The analyses in Appendix 6 show the random effects parameters revealing the degree of variability among respondents in these nests in a form of standard deviations-sd(_cons) and the variability across respondents in the nests, again, in a form of standard deviation-sd(Residual). These measures of variability were used to run interclass correlation coefficient (ICC) tests to examine the degree of correlation within the nests. The formula below was used for both the unbalanced panel and the cross-sectional analysis:

$$\frac{sd(_{cons})^2}{sd(_{cons})^2 + sd(Residual)^2} \tag{11}$$

The ICC for the unbalanced panel was:

$$\frac{0.058^2}{0.058^2 + 0.195^2} = 0.08$$

While that of the cross-sectional analysis was:

$$\frac{0.057^2}{0.057^2 + 0.17^2} = 0.10$$

If the ICC approaches zero, then there is variance at individual level, hence multilevel modelling would perform better. However, if the ICC approaches 1 then there is no variance at individual level, implying that all respondents are not different at each wave. In this case, the ICC is 0.08 for the unbalanced panel analysis and 0.10 for the cross-sectional analysis which indicate that there is variance at individual level and a multilevel regression is plausible. Having established the impact of the WIAT intervention, the next section looks at its cost-effectiveness based on the standard societal willingness to pay per QALY gained from the intervention (NICE, 2013).

6.3.2 Cost-utility analysis of the WIAT intervention

Despite the WIAT intervention showing no statistical meaningful change to the HRQoL in the intervention group relative to the control group, it was possible to make a judgement about whether the WIAT intervention is value for money using a cost-effectiveness analysis through an ICER because the absence of evidence of effect does not mean evidence of absence of effect (Altman and Bland, 1995; Briggs and O'Brien, 2001). Furthermore, economic evaluation is more concerned with the ICER and the exploration of uncertainty around base case results than the significance of hypothesis testing of the effect of an intervention (Briggs and O'Brien, 2001; Gray et al., 2010).

The incremental cost-effectiveness ratio of the WIAT intervention compared with the option of not investing in the intervention was estimated as the difference in mean cost divided by the difference in QALYs using the formula (5) in chapter two. The base case results showed that the WIAT physical intervention was associated with the incremental expected cost of £7.68 and the incremental QALY gain of 0.012. On the other hand, both the physical and social interventions were associated with the incremental expected cost of £11.80 and the incremental QALY gain of 0.023. This translated to an ICER of £641 per QALY gained for the physical intervention and £513 for both physical and social interventions.

The decision on whether the WIAT interventions were good value for money was based on the acceptable willingness to pay (WTP) threshold of between £20,000 and £30,000 that society is willing to sacrifice for each QALY gained from an intervention (NICE, 2013). For the base case results, the net monetary benefit (NMB) for the physical intervention was £231.91 while for both the physical and social interventions, the NMB was £448.68 at WTP of £20,000. At WTP of £30,000, the NMB for the physical intervention and both the physical and social interventions was £351.71and £678.93, respectively.

Sensitivity analysis

Uncertainty around the above base case results was explored using the probabilistic sensitivity analysis (PSA). The results of the PSA are presented using the cost-effectiveness (CE) plane in Figure 6-5 and Figure 6-6 below, for the physical intervention and both the physical and social interventions, respectively. The CE plane depicts the spread of the bootstrapped pairs of the mean cost and mean QALY differences between the intervention and control groups from the 5,000 bootstrap samples of the Monte Carlo simulations.

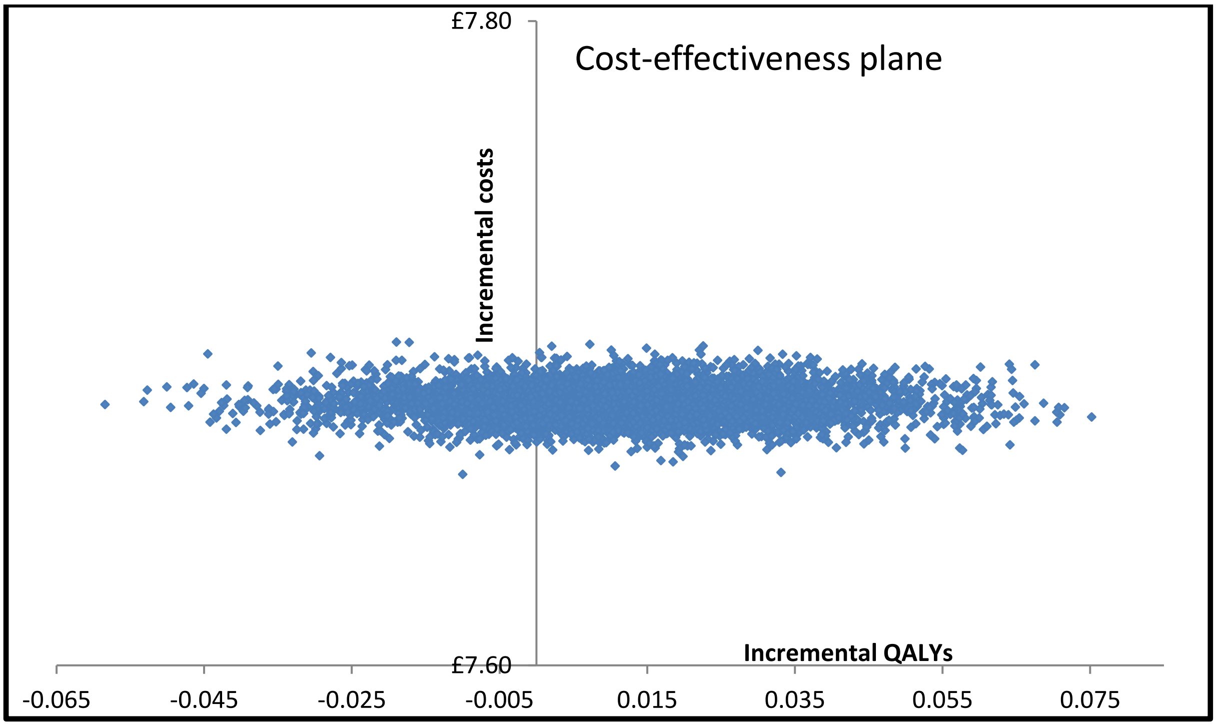


Figure 6-5: Cost-effectiveness plane for the physical intervention for unbalanced panel analysis

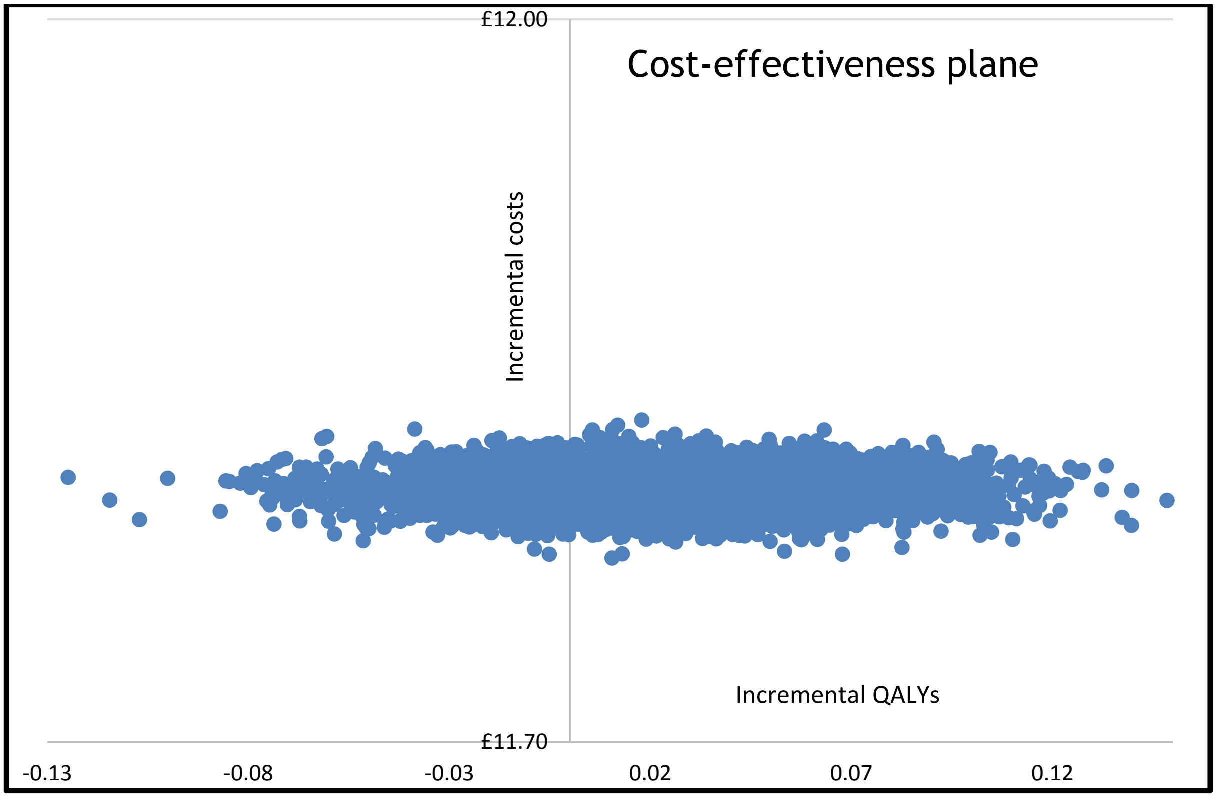


Figure 6-6: Cost-effectiveness plane for both the physical and social intervention for unbalanced panel analysis

According to the 5,000 bootstraps of the incremental mean cost and QALY pairs for the physical intervention in Figure 6-5, the majority of the bootstrapped ICERs are placed on the North-East quadrant of the cost-effectiveness plane while the rest of the bootstrapped ICERS are placed in the North-East quadrant for the physical intervention and both the physical and social interventions, which implies positive cost and positive effect. In this case, a trade-off between cost and effect in terms of QALY needs to be examined by referring to specific thresholds of willingness to pay (WTP) (λ) (Fenwick et al., 2006). This is intuitively and better depicted using the cost-effectiveness acceptability curve (CEAC) as shown in Figure 6-7 below:

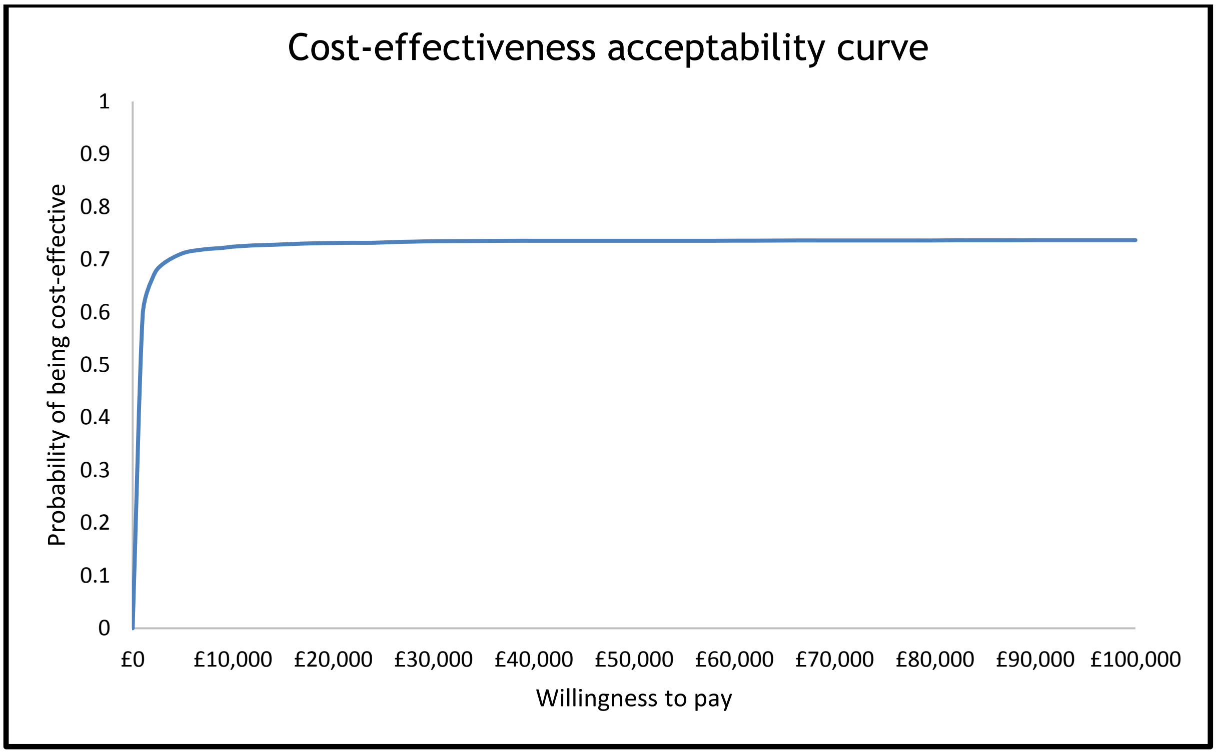


Figure 6-7: Cost-effectiveness acceptability curve for the physical intervention

The cost-effectiveness acceptability curve (CEAC) above shows the percentage of the Monte Carlo simulations in which the physical intervention is cost-effective based on proportions of the bootstrap replications with positive incremental net monetary benefit across a range of willingness to pay values per QALY gained. The physical intervention is about 73% likely to be cost-effective at willingness to pay threshold values of £20,000 and £30,000, respectively.

As regards the CEAC for both the physical and social interventions in Figure 6-8 below, the likelihood of being cost-effective is between 74% and 75% at willingness to pay threshold values of £20,000 and £30,000, respectively.

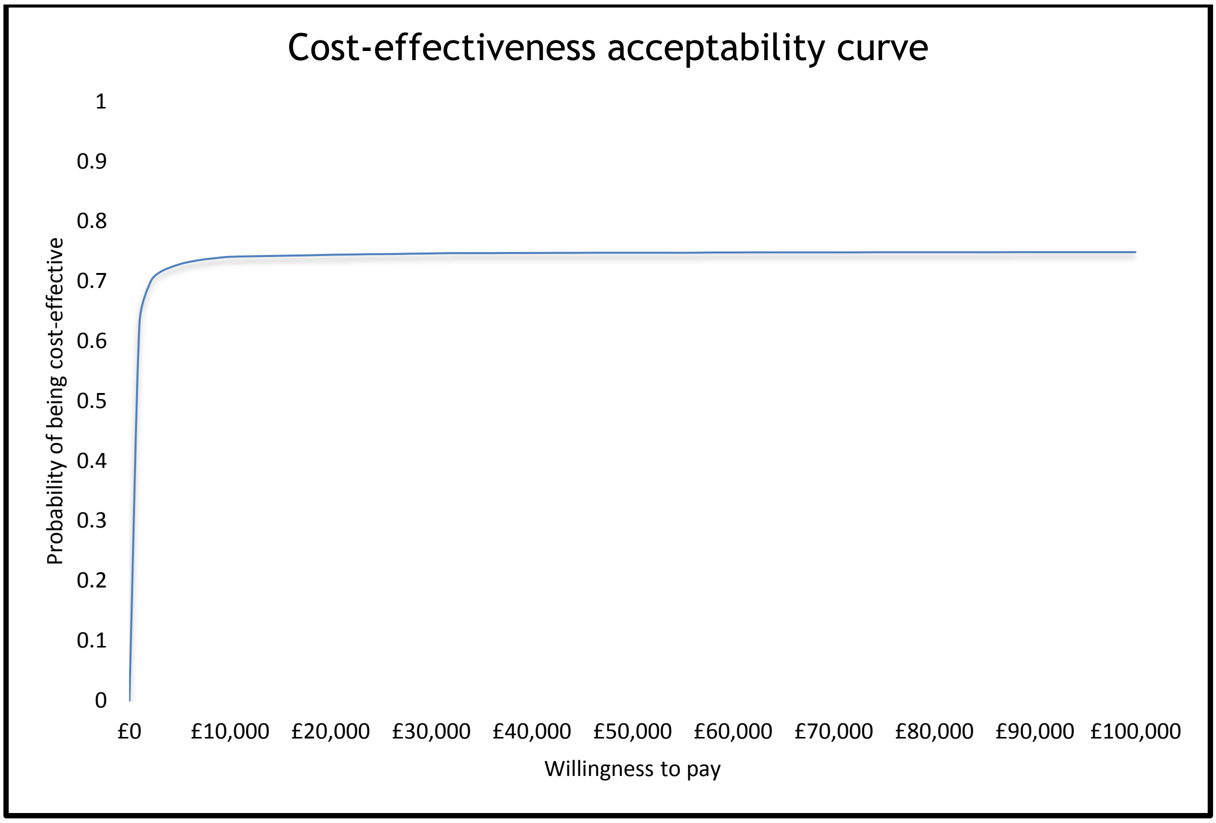


Figure 6-8: Cost-effectiveness acceptability curve for the physical and social interventions.

Table 6-11 below, presents a summary of the bootstrapped results of the cost-effectiveness analysis for the unbalanced panel based on the 5,000 bootstraps for the physical intervention and the combined physical and social intervention. The results show the net monetary benefit at willingness to pay of £20,000 and £30,000 per QALY gained with 95% CI based on bootstrap percentile method.

	Physical intervention	95% confidence interval		Physical and social intervention	95% confidence interval	
Bootstrapped Incremental cost	£7.68	£7.67	£7.69	£11.80	£11.79	£11.82
Bootstrapped incremental QALY	0.012	-0.028	0.051	0.024	-0.049	0.094
Bootstrapped ICER	£627	-£5,757	£5,218	£500	-£3,999	£4,098
Bootstrapped NMB for WTP of £20,000	£238	-£563	£1,019	£127	-£993	£1,869
Bootstrapped NMB for WTP of £30,000	£360	-£841	£1,533	£210	-£1,484	£2,809

Table 6-11: Bootstrapped results of cost-effectiveness analysis for unbalanced panel analysis.

These bootstrapped results above reveal that there was a great deal of uncertainty around the QALY results.

The same analysis was undertaken using the repeated cross-sectional data. The base case results showed an incremental expected cost of £7.68 and an incremental QALY of 0.015, which results in an ICER of £501 for the physical intervention. When both the physical and social interventions are given, the incremental expected cost becomes £11.80 and the incremental QALY is 0.029. This translates to an ICER of £410. Based on the £20,000 and £30,000 WTP thresholds, the NMB is £299 and £452, respectively, for the physical intervention and £563 and £851, respectively for both the physical and social interventions. The CEAC for the physical intervention reveals that the probability of being cost-effective at the WTP of £20,000 and £30,000 is 96%. On the other hand, the probability that both the physical and social intervention is cost-effective is 97% at the WTP of £20,000 and

£30,000, respectively. The bootstrapped results also indicate a great amount of uncertainty around QALYs.

6.4 Discussion

Firstly, this chapter sought to establish the impact of the WIAT intervention on health-related quality of life as measured by the EQ-5D five-dimension descriptive system questionnaire. Secondly, it aimed to determine the cost-effectiveness of the WIAT intervention to help resource allocation decision-making.

As regards the impact of the WIAT intervention, the results of the DiD regression models showed no evidence of any statistically significant change in the HRQoL for those individuals in contact with nature for both the physical intervention and the combined physical and social interventions in the unbalanced panel and cross-sectional analysis, respectively. Despite this outcome, it is essential to note that the use of a difference-in-differences (DiD) regression method helped to deal with attributing the causal impact of the intervention. This is one of the problems in economic evaluations of public health interventions as discussed in chapter four (Weatherly et al., 2009). The DiD approach allowed the adjustment for baseline characteristics differences between the intervention and control groups and other external effects resulting from trends over time in the absence of any intervention.

Turning to the assessment of the cost-effectiveness, given that the WIAT intervention showed an insignificant effect, one option for economic evaluation would have been to undertake a cost-minimization analysis (CMA) as discussed in chapter two. The assumption would have been that the health-related quality of life was the same in the intervention group and control group after giving the physical intervention and both the physical and social interventions. Since the control group did not incur any cost, the alternative option of doing nothing in terms of giving the intervention would have been preferred. However, reliance on the assumption that the effect of the WIAT intervention is the same in both the intervention and control group based on statistical significance tests could be misleading (Briggs and O'Brien, 2001; Gray et al., 2010). When it is found that an intervention has no effect, it does not necessarily mean that the effect is absent (Altman and Bland, 1995; Briggs and O'Brien, 2001). It could just be a case of absence of effect. Hence, CMA would not be helpful in informing resource allocation decisions unless if the study was specifically designed to show the equivalence of either costs or effects (Briggs and O'Brien, 2001), which was not the case with the WIAT study.

With the above discussion in mind and as previously noted, the aim of economic evaluation should be to calculate the incremental cost-effectiveness ratio and explore uncertainty surrounding the parameter estimates of cost and effect rather than rely on the level of significance of the effect (Briggs and O'Brien, 2001; Gray et al., 2010).

The base case results of the cost-effectiveness analysis of the physical intervention and the combined physical and social intervention showed that the WIAT intervention was cost-effective for both the unbalanced panel and cross-sectional analysis based on the accepted willingness to pay thresholds of £20,000, and £30,000 (NICE, 2013). The NMB was higher in the cross-sectional analysis compared with the unbalanced panel analysis for both the physical intervention and the combined physical and social interventions. This could be explained by the differences in the samples. In general, there was a great deal of uncertainty around the QALY results. Therefore, the results should be interpreted with caution.

The results on the effect of the WIAT intervention may be somewhat limited by the following: firstly, given that there was no meaningful impact in terms of health-related quality of life on the intervention group relative to the control group, questions could also be asked if at all the EQ-5D questionnaire was a good measure of the health outcomes for this type of intervention which is related to mental well-being. However, even the Perceived Stress Scale (PSS) and the Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS) used to measure improvements in mental well-being hardly showed any positive effect resulting from the intervention. Hence, it cannot be concluded that the EQ-5D tool was insensitive in this case. It is, however, acknowledged that the EQ-5D questionnaire has been shown to be problematic to use in people with complex or severe mental health related problems and evidence as to whether it is fit for purpose in mental health related studies is mixed (Brazier, 2010; Luyten et al., 2016). This may be attributed to the incapacity of respondents with severe mental health problems to complete the questionnaire.

Secondly, since the WIAT study took the form of a natural experiment and natural experiments provide a counterfactual through a control group which results in a robust causal estimate of the intervention effect, questions arise whether a single natural experiment is enough to provide sufficient evidence on which to base a decision on the cost-effectiveness of an intervention (Sculpher et al., 2006).

Thirdly, another limitation relates to generalizing the results of one particular intervention to other settings (Remler and Van Ryzin, 2010). As the effect of the intervention is only determined from the group that receives the intervention compared with that that does not receive it, it cannot be concluded with certainty that the same effect may be identified elsewhere or will continue in the same manner in the intervention group compared with the group that will never receive the intervention at all.

Fourthly, the shorter time horizon of two years for conducting an economic evaluation is another concern because access to the woodlands will continue to be available to individuals after the completion of the WIAT study. This implies longevity of effect which will need follow-up. Outcomes of most public health interventions may generally take a long time;

continue to exist in the foreseeable future or could be inter-generational (Remler and Van Ryzin, 2010; Park, 2014). In addition, there will, arguably, be maintenance costs and social intervention costs to increase awareness about the positive benefits of woodland on individuals for the WIAT study. However, following up on these could be costly and sometimes not feasible, hence considered as one of the challenges to this study and to economic evaluation of public health interventions in general (Weatherly et al., 2009).

Fifthly, the WIAT study did not record seasonality in terms of when the data was collected. This is a limitation which could have important implications on the results of this study. Woodland visits are hugely impacted by different seasons of the year.

Lastly, the DiD approach used to determine the effect of the WIAT intervention is known to have some limitations. This approach uses the parallel trends and the common shocks assumptions as discussed earlier in this chapter (Angrist and Pischke, 2008; Mills and Patterson, 2011). These assumptions may not be plausible in some settings (O'Neill et al., 2016). For example, the 'parallel' trends assumption may be problematic in that some unobserved confounders may have time-varying effect on the outcome (Dimick and Ryan, 2014). Secondly, the 'common shocks' assumption becomes a challenge in reality as it is difficult to find a control group which meets this assumption in its entirety (Dimick and Ryan, 2014). However, despite these drawbacks, natural experiments and the DID approach offer a fairly good evidence of causation (Remler and Van Ryzin, 2010).

As regards the cost-effectiveness results, they should also be interpreted with caution in a broader sense. As previously stated in chapter five, the costing of resource use for the delivery of the WIAT intervention was not able to capture all relevant activities such as time of the members of the community spent in the implementation of the social intervention. Furthermore, a very small number of QALYs gained from an intervention would generally imply to mean that the intervention had a ‘small effect’ which would likely not be the case especially that this benefit goes out to a large number of people (Phillips et al., 2011).

Another note of caution relates to the failure of the cost-effectiveness analysis to value outcomes that go beyond health. Given that the WIAT intervention has broad outcomes consisting of health and non-health, the cost-effectiveness analysis only provides a partial valuation for only the health-related outcomes of the intervention. For this reason, the results of the cost-effectiveness analysis should not be considered in isolation. Other measures for the non-health related outcomes can be used to compliment the cost-effectiveness analysis through a cost-consequences analysis to give a full picture of the overall effect of the intervention. These broad outcomes can then be combined on the same monetary scale, for example, using the novel integrated approach proposed in this thesis as demonstrated in chapter eight.

In general, despite all the above limitations, these analyses would offer decision-makers with a basis on which to make judgements as to whether to adopt the WIAT intervention compared with the option of doing nothing. The primary benefit of the cost-effectiveness analysis is particularly capable of providing policy or decision-makers with a common yardstick on which to make judgements about the worthiness of an intervention compared with alternative interventions within the health sector.

6.5 Conclusion

The aim of this chapter was to establish the impact of the WIAT intervention and conduct a cost-utility analysis based on utilities derived from the EQ-5D descriptive system. The results on the impact of the intervention showed that there was no evidence to support that there was any meaningful change in health-related quality of life resulting from contact with nature in the intervention group relative to the control group. However, when the incremental mean cost of the WIAT intervention was weighed against the expected incremental QALYs resulting from the intervention, the WIAT intervention turned out to be cost effective based on the acceptable societal willingness to pay thresholds per QALY gained. There was huge uncertainty around the base case results as revealed by the PSA.

The next chapter discusses the assessment and valuation of the non-health related outcomes of the WIAT study. The stated preference discrete choice experiment (SPDCE) technique of indirectly eliciting WTP values was used to value the specific changes or improvements which could be attributed to the WIAT intervention. This was done through mapping the WIAT main study questionnaire items that were considered to measure the non-health benefits to the attributes and levels of the SPDCE.

Chapter 7: The valuation of the non-health outcomes of the WIAT study

7.1 Introduction

This chapter addresses the third objective of the thesis: to value the identified non-health outcomes of the WIAT intervention which are examples of the outcomes of a public health intervention. These include: the enhanced environment which would result in the woodlands being more accessible, more attractive, safe to use and well maintained; the behavioural and perceptual outcomes such as increased visits to woodlands, and taking greater pleasure in the views of the woods; and the social support for environmental use including increased awareness of local woodlands, community engagement and social activities.

Given the broad outcomes of the WIAT intervention, which consist of health and non-health related outcomes, a more appropriate approach to its economic evaluation would be a cost-benefit analysis (CBA) (McIntosh et al., 2010). A CBA is broader in focus as it attaches monetary values to the outcomes of an intervention through an assessment of individuals' willingness to pay (WTP) using preference elicitation methods such as the revealed preference (RP) or the stated preference method (SP) (O'Brien and Viramontes, 1993; Lancsar and Louviere, 2008; McIntosh et al., 2010). However, CBA is rarely used in standard economic evaluations of healthcare due to lack of acceptability of assigning monetary values to health outcomes (McIntosh et al., 2010). This is considered as unethical and favouring only those who can afford to pay.

Recent methodological developments have seen improvements in preference elicitation methods for willingness to pay. The SP method using discrete choice experiment (DCE) has become the preferred approach because it indirectly elicits individuals' WTP values as opposed to direct SP approaches like contingent valuation (CV) (Pearce et al., 2002; Bridges et al., 2011; Fujiwara and Campbell, 2011). In this chapter, the stated preference discrete choice experiment (SPDCE) is used to value the non-health outcomes of the WIAT intervention.

The chapter begins with a brief discussion on the SPDCE approach and how it has been used to assess and value the non-health outcomes of the WIAT intervention followed by the presentation of the results. The SPDCE approach is discussed in detail in chapter two. The chapter then proceeds to discuss the mapping of the WIAT main study questionnaire items that were considered to measure the non-health outcomes to the attributes and levels of the SPDCE. This would allow the calculation of the incremental changes or improvements in the attributes and levels resulting from the intervention. Following this, the willingness to pay estimates from the SPDCE are applied to these incremental changes or improvements in the attributes and levels to estimate their value. Finally, a discussion and conclusion wraps up the chapter.

7.2 Stated preference discrete choice experiment (SPDCE)

To recall, the SPDCE approach uses a specially constructed questionnaire to indirectly elicit WTP values which can be used as input in an economic evaluation (McIntosh et al., 2010; Ryan et al., 2012a; Clark et al., 2014). The respondents are presented with alternatives with the attributes but with varying levels and are asked to make a choice between these alternatives. When a cost attribute is included in a SPDCE, it is possible to indirectly elicit the willingness to pay estimates through an assessment of the trade-offs of the attributes and levels using the marginal rate of substitution (MRS) as shown later in this chapter (Ryan et al., 2008b; McIntosh et al., 2010; Ryan et al., 2012a). The key advantage of the SPDCE approach to indirectly elicit willingness to pay values is that it is sensitive enough to pick up changes caused by the variations of attributes and levels (Evers et al., 2015). The sum of willingness to pay for relevant changes in the levels of attributes affected is the value of the outcome being evaluated.

The notion of making a choice and trading-off of the attributes and levels in the SPDCE preference elicitation task is appealing because it implies opportunity cost, a key concept used to determine the value of a good in economics (Briggs, 2016). As explained in chapter two, the SPDCE approach is based on two concepts: first, that the value of a non-market priced good can be determined from its attributes rather than its consumption per se; and second, that individuals choose goods which give them the highest level of satisfaction (utility) (Lancaster, 1966; McFadden, 1974). These concepts are generally referred to as the theory of value and random utility theory (RUT), respectively. When these theories are considered together, it is possible to estimate the value of a non-market priced good using logistic regression (Kjær, 2005; Hanley et al., 2006; Ryan et al., 2008a; Mentzakis et al., 2011; Londoño and Ando, 2013). The SPDCE approach, arguably, appears to improve on the limitations of directly eliciting willingness to pay values from respondents for use in healthcare economic evaluation (McIntosh, 2006; Green and Gerard, 2009).

However, even when the SPDCE is used, questions remain as to how to incorporate the SPDCE WTP values into an economic evaluation (Tinelli et al., 2016). The SPDCE in this chapter is part of the integrated approach proposed by this thesis which is particularly argued to be suitable for a public health intervention because it considers both the health and non-health outcomes on the same monetary scale using the net monetary benefit (NMB) framework, as demonstrated in chapter eight. The next section discusses how the SPDCE was carried out to value the non-health related outcomes of the WIAT intervention.

7.3 Methods

The most important aspect of the SPDCE is the design process. As explained in chapter two, designing a SPDCE involves five key stages: first, is the identification of attributes; second, is the assignment of levels to the attributes; third, is the development of an experimental design which defines the choice alternatives that would be presented to respondents; fourth, is the development and administration of questionnaires to collect data; and fifth, is the data input, analysis, and interpretation of responses from the survey. Details of what is involved at each stage have been presented in chapter two. In the section that follows, these five steps are discussed in relation to the valuation of the non-health related outcomes of the WIAT study.

7.3.1 Attribute identification and level assignment

Several steps were undertaken to identify relevant attributes and assign levels for the SPDCE of the WIAT intervention. These were: 1) establishing what the SPDCE aimed to value; 2) observational visits to the intervention and control sites before and after the intervention to understand the characteristics of the woodlands; 3) reviewing literature on predictors of woodland use; 4) reviewing the wider WIAT main study questionnaire; 5) mapping the wider WIAT main study of questionnaire items which were considered to measure the non-health outcomes of interest to the SPDCE attributes and levels; 6) discussing and consulting with experts to ensure that all relevant attributes and levels were included and that their framing was appropriate in order to reduce cases of non-attribute attendance; 7) and piloting the draft SPDCE questionnaire to check if it made sense to respondents before the main survey was undertaken. The pilot survey also asked respondents if there were other attributes that they felt could have been included in the SPDCE but were left out. Below are details of the whole process:

1. Aim of the SPDCE

Firstly, it was important to understand the purpose of the SPDCE to be able to map the WIAT main study questionnaire items to the relevant attributes and levels. This required an understanding of the woodland characteristics before and after the intervention. Furthermore, it was essential to understand the conceptual framework of the impacts of the WIAT intervention which depicts the health and non-health outcomes of the WIAT intervention as presented in chapter three. The purpose of the SPDCE, therefore, was to value the non-health related outcomes.

2. Observational visits

It was also necessary to make pre-and post-intervention observational visits to the intervention and control sites of the WIAT study to assess and understand the characteristics of the woodlands. Appendix 3 to Appendix 5 show how the woodlands were, before and after the intervention in both the intervention and control sites.

3. Review of literature

Another important stage of the attribute identification and assignment of levels for this SPDCE was the review of literature to find out predictors of woodland use. This included reviewing literature on baseline studies that informed the design of the WIAT study (Ward Thompson et al., 2004; Ward Thompson et al., 2005; Ward Thompson et al., 2007). It was identified that the theoretical framework of David Canter's 'Theory of Place' (Canter, 1977) was used to identify predictors of woodland use and inform the baseline studies of the WIAT study and subsequently, the development of the WIAT main study questionnaire (Ward Thompson et al., 2004; Ward Thompson et al., 2007).

David Canter's 'Theory of Place' posits that individuals relate to a place because of the influence and interaction of three broad attributes: the physical attributes relating to form and space; the functional attributes as regards behaviours or activities that individuals engage in; and the psychological attributes which include perceptions or conceptions they have about a place (Canter, 1977; Knez, 2005; Ward Thompson et al., 2005; Bell and Ward Thompson, 2008). This individual's connectedness with a particular place is also known as place attachment and includes a combination of affect, emotions, behaviours and actions related to that particular place (Knez, 2005). Bell and Ward Thompson (2008) recommend that when exploring the benefits of a place such as a woodland to the lives of individuals, it is important to consider all the three elements of Canter (1977)'s theory of place: physical attributes; functional attributes; and psychological attributes; and the interaction between them. For this reason, it was important to identify the SPDCE attributes and their associated levels based on this theory. The basis was that the broad sources of influence or satisfaction (utility) for woodland use would border around the three aspects of Canter (1977)'s theory of place.

4. Review of the WIAT main study questionnaire

Another necessary step was to review the wider WIAT main study questionnaire shown in Appendix 1 to identify the self-reported survey questions that measured the non-health related outcomes to be able to link them to the SPDCE attributes and levels retrospectively. As can be seen from Appendix 1, the WIAT main study questionnaire has nine parts (A-I). Part A sought information about the location and gender of the respondents while Part B to H comprised various measurement tools which captured the outcomes of the intervention. Part I was about the socioeconomic characteristics of the respondent.

Four attributes based on Canter (1977)'s theory of place were considered to emerge from the WIAT main study questionnaire items B4; D2; B11; B29; and H1. The identified attributes were:

1. The woodland environmental support which was defined as one which allowed individuals to do the things they wanted to do, either on their own or with others (such as exercise, relaxing, enjoying wildlife) and makes it easy and enjoyable to do them.
2. The time that it takes to walk from home to the woodland.
3. The quality of the woodland environment which include cleanliness; the condition of paths and entrances; the naturalness of its appearance; the views of plants and wildlife.
4. The opportunities for social activities that the woodland offers individuals such as meeting people, community events, guidance on how to use the woodland and about what is going on there.

Figure 7-1 below presents Canter (1977)'s theory of place model that has been adapted to locate the four identified attributes of the SPDCE:

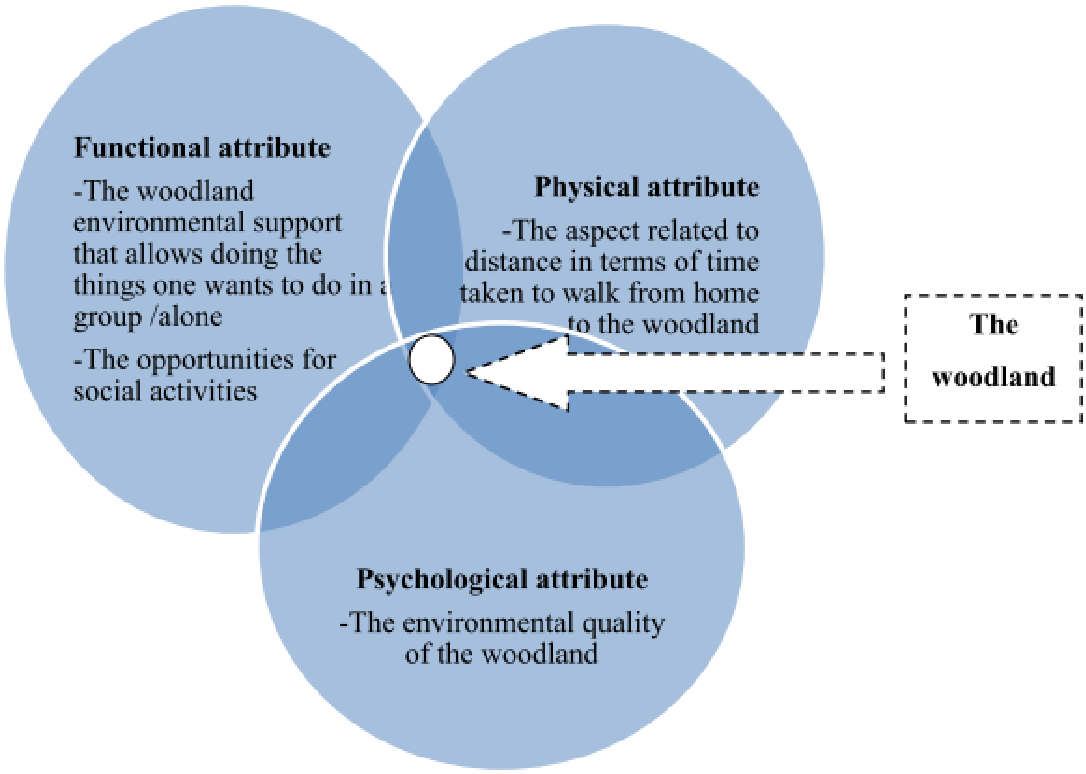


Figure 7-1: Canter (1977)'s theory of place model adapted to locate the attributes of the SPDCE.

The process of identifying attributes and assigning levels of the SPDCE includes the need to consider important aspects such as the maximum number of attributes and levels, and attribute-level ranges. These directly impact on the responses of the SPDCE and can negatively affect the results of the SPDCE as explained in chapter two. In the case of this SPDCE, the process was complex because of the pre-specified nature of the WIAT main study questionnaire whose design did not have the economic evaluation using the SPDCE in mind. Given this limitation, assumptions had to be made to get the 'best-fit' between mapping the existing WIAT questionnaire to the SPDCE, as discussed later in this chapter. This would enable the assessment of the incremental changes or improvements in the attributes and levels resulting from the intervention for both the unbalanced panel analysis as a base case and the cross-sectional analysis.

Then, it would be possible to value the non-health outcomes of the WIAT intervention using the societal willingness to pay estimates obtained from the SPDCE. The ideal situation would have been to identify the attributes and levels of the SPDCE in advance or to have prior knowledge of them and incorporate them in the main study questionnaire alongside the QALY framework. The next section discusses how the mapping of the WIAT main study questionnaire items that were considered to measure the non-health outcomes to the attributes and levels of the SPDCE.

5. <u>Mapping of the WIAT questionnaire to the SPDCE</u>

The process of mapping the attributes and levels of the SPDCE to the identified WIAT main study questionnaire items was pragmatic given that the WIAT main study design did not envisage that a SPDCE would be incorporated at a later stage. For this reason, there was need to make some assumptions to be able to link the WIAT main study questionnaire to the SPDCE attributes and levels. As will be seen later in this chapter, the assumptions made may be somewhat be problematic because they may not reflect how individuals use woodlands. However, it was necessary to make them to be able to develop the integrated approach proposed in this thesis.

Having said that, of interest in the mapping process were the WIAT main study questionnaire items B4; D2; B11; B29; and H1. These were considered to measure the non-health outcomes of the WIAT intervention as described by its conceptual model in chapter three.

Then, the responses from the questionnaire item B4 and D2 shown below were mapped to the SPDCE assigned levels of the attribute related to woodland environmental support for activities "*the woodland environmental support which we define as one which allows you to do the things you want to do, either on your own or with others (such as exercise, relaxing, enjoying wildlife) and makes it easy and enjoyable to do them*" which are: "*No support, Some support, A lot of support*". There are eight specific activities in B4 and D2 including "*other*" that can be pursued in the woodlands. If someone responded they visited the woodland but did not pursue any activity, they were mapped to "*No support*". It is acknowledged that this mapping is problematic because if someone does not pursue any activity in the woodlands, it does not imply lack of woodland support for activities. It may simply be for other reasons like personal choice or culture. The mapping of the attribute-levels "some support" and "A lot of support" was also problematic. The feasible approach was to map the responses to "other" activities to the attribute level "*Some support*" and the responses to the rest of the specified activities were mapped to "*A lot of support*". This is also a limitation. Questions may be asked if this assumption reflects a realistic assessment of how individuals use the woods.

B4. What kinds of activities do you pursue when visiting these local woodlands? SPONTANEOUS

CODE ALL THAT APPLY	Code
Go for a walk	1
Walk the dog	1
Go out with my family	1
Exercise or sport	1
Relax	1
Look at plants or wildlife	1
Participate in an event	1
Other (Please specify)____	1

D2. What kinds of activities do you pursue when visiting local parks or green spaces? SPONTANEOUS

CODE ALL THAT APPLY	Code
Go for a walk	1
Walk the dog	1
Go out with my family	1
Exercise or sport	1
Relax	1
Look at plants or wildlife	1
Participate in an event	1
Other (specify)____ _	1

Questionnaire item B11 was mapped to the attribute "*the time that it takes to walk from home to the woodlands*".

B11. How long would it take you to walk to these local woodlands?

READ OUT, SINGLE CODE	**Code**
_______________ minutes	
Cannot walk (If respondents cannot walk go to B11.1)	0

Given that intervention was expected to make accessible areas of woodlands that were previously inaccessible, this mapping intended to measure changes or improvements in terms of distance reduced to accessible local woodlands because of the intervention.

The attribute related to quality of the woodland environment "*the quality of the woodland which include cleanliness; the condition of paths and entrances; the naturalness of its appearance; the views of plants and wildlife*", was linked to questionnaire item B29 as shown below:

B29. Overall, what do you think about the quality of these local woodlands? SHOW CARD

Very good	Good	Neutral	Poor	Very poor	Do not know what my local woodlands are like
1	2	3	4	5	-98

It was felt that the responses to this questionnaire item considered and summed up the elements that make up the overall quality of the woodland environment. The responses to questionnaire item B29 "*Good*" and "*Very good*" were mapped to the attribute level "*Good quality*; while "*Neutral" and "Do not know*" were mapped to the attribute level *"Average quality*" while "*Poor",* and *"Very poor*" was mapped to the attribute level "*Poor quality*". It is important to bear in mind that this mapping also has weaknesses. For example, "neutral" and "do not know" responses may not necessarily indicate "average quality" of the woodland.

The fourth attribute related to "*the social opportunities that the woodland offers you, connecting with your community through events and or meeting people, and the availability of information such as leaflets and guidance on how to use the woodland, and about what is going on there"*. Questionnaire item H1 of the WIAT main study shown below, was considered to capture, in a broader sense, the social opportunities resulting from the intervention.

PART H
Social Cohesion / Social Capital

Please say how much you agree or disagree with the following statements:

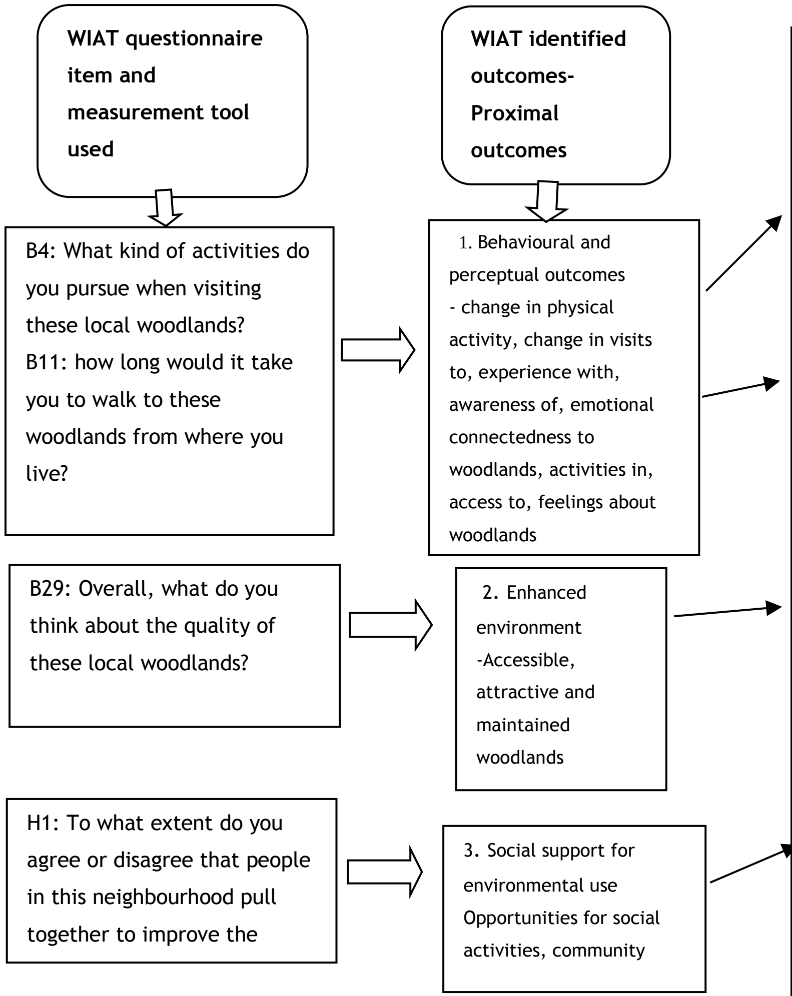
H1. To what extent do you agree or disagree that people in this neighbourhood pull together to improve the neighbourhood?

SHOW CARD P

Definitely agree	Tend to agree	Tend to disagree	Definitely disagree	SPONTANEOUS ONLY: Nothing needs improving	Do not know
1	2	3	4	5	-98

Its responses were collapsed to three levels to allow for mapping to the SPDCE attribute levels. In this case, "*definitely agree*" was mapped to "*Many opportunities*"; and "*tend to agree*" was mapped to "*Some opportunities*"; while "*tend to disagree*", "*definitely disagree*", "*nothing needs improving*" and "*do not know*" were mapped to "*No opportunities*". Again, this mapping has limitations but was considered to be a reasonable approach in the context of the WIAT study design constraint.

Figure 7-2 below presents a summary of the mapping exercise. It shows the WIAT questionnaire items from the WIAT main study that were considered to measure the non-health outcomes and the mapping to the identified attributes and levels of the SPDCE. The last column in the figure presents the reason for the choice of the attribute-levels used in the SPDCE:



Identified SPDCE attribute	Level	Remark/Justification for the level
The woodland environmental support which we define as one which allows you to do the things you want to do, either on your own or with others and makes it easy and enjoyable to do them.	No support, some support, A lot of support	In most cases, free usage and aesthetic characteristics of woodlands provide conducive environment for activities at individual or group level (Greenspace Scotland, 2008; Herzele and Vries, 2011; Groenewegen et al., 2012)
Time that it takes to walk from home to the woodlands	5mins 15mins, 50mins	The baseline is the European Environment Agency (EEA) recommendation of woodlands to be within a 15 minute walk distance from homes (Barbosa et al., 2007). If distance from home was used, then English Nature (EN) UK recommends access to a green space of at least 2 hectares within 300 metres from homes (Barbosa et al., 2007; Schipperijn et al., 2010)which translates to no more than 5mins walk (Balfour and Allen, 2014)
The quality of the woodland environment	Poor quality, Average quality, Good quality	Nuisance in the woodlands such as litter graffiti, signs of vandalism create feelings of fear and deter people from visiting the woods (Ward Thompson et al., 2005)
The opportunities for social activities that the woodland offers you such as meeting people, community events	No opportunities, some opportunities, Many opportunities	There is evidence that green spaces result in cohesive communities (Herzele and Vries, 2011; Lee and Maheswaran, 2011; Arnberger and Eder, 2012; Xiaolu and Md Masud, 2012)
The cost for access to the woods if you lived in a country where you had to pay for access	£0, £15, £50	Hypothetical cost determined from WTP studies in Scotland (Bateman, 1996; Edwards et al., 2009)

Figure 7-2: Mapping of the WIAT questionnaire items to the SPDCE attributes and levels.

An additional cost attribute was included in the SPDCE as shown in Figure 7-2 above, as the payment vehicle to enable the indirect estimation of the WTP for the attributes and levels as mentioned earlier (Hanley et al., 1998; Morrison et al., 2000; Kjær, 2005; Hoyos, 2010; McIntosh, 2010; Aravena et al., 2014).

While the inclusion of a payment vehicle is critical in the estimation of WTP values, its choice is not without problems. In most cases, it is associated with negative utilities which can lead to protest responses in the SPDCE survey (Kjær, 2005). This being an environmental related SPDCE, the most common payment vehicle used in literature is a tax or levy payment, while other environmental related SPDCEs have used donations as a payment vehicle (Kjær, 2005; Gyrd-Hansen, 2013; Vecchiato and Tempesta, 2013). All these payment vehicles are associated with limitations. Tax or levy payments have caused equity concerns and donations have been associated with 'free-riding' or a 'purchase of moral satisfaction' not reflecting the actual value of a good (Kahneman and Knetsch, 1992; Kjær, 2005).

This supports the view that care should be exercised when deciding on the appropriate payment vehicle to minimise or avoid protest responses from respondents. Another aspect of the payment vehicle which has caused much debate in SPDCE studies is the payment duration, whether it should be weekly, monthly, or yearly (Kjær, 2005). There are also other considerable problems which arise from the payment vehicle such as: protest bidding where respondents may be unwilling to pay any amount beyond a certain threshold for some improvements in attributes; or respondents may simply choose an alternative which appears to be cheaper irrespective of the gain on other attributes (Kjær, 2005; Gyrd-Hansen and Skjoldborg, 2008).

One approach to dealing with these problems has been to choose the payment vehicle that is realistic, context and case specific, use common practice or conservative payment duration (Kjær, 2005; Can and Alp, 2012).

As regards this SPDCE, an entrance fee to woodlands would have been a suitable payment vehicle but it has a limitation of only considering direct use while ignoring non-use value (Bateman et al., 2002; Pearce et al., 2002). Non-use value is that which people assign to a public good even if they never have and never will use it (Bateman et al., 2002). For this reason, a nominal fixed annual payment vehicle was proposed for this study in a form of an annual subscription per household to account for both direct use and non-use values. This yearly subscription per household may be problematic as it may not reflect how individuals access woodlands in Scotland where access is normally free. To mitigate this problem, respondents were asked to assume they lived in a country where they had to pay an annual subscription to access woodlands.

To determine the annual payment vehicle amount, it was important to establish a per visit willingness to pay to access woodlands and the average number of yearly visits per household to woodlands in Scotland, since the SPDCE was administered to the Scottish population. However, due to scarcity of Scottish studies on willingness to pay estimates for woodland access, two English studies were used (Bateman, 1996; Edwards et al., 2009). Bateman (1996) estimated the average number of visits to woodlands per year per household to be 15 and elicited a woodland per-visit measure of WTP of £0.82. A payment vehicle amount for this study was, therefore, determined from Bateman (1996) study which translated to approximately 15 x £0.82 per year per household which equals £12.29. Another study in Scotland revealed a WTP of £1.03 per visit per person including accompanying children to non-Forestry Commission woodlands in 2007/08 prices which translated to approximately £1.17 per visit per person including accompanying children (about £17.55 per annum per person assuming that each person makes on average 15 visits to the woodlands) (Edwards et al., 2009).

In the case of this SPDCE, the hypothetical cost for access of woodland per visit per annum per household has been assumed to range between £0 and £50 as a fixed annual subscription regardless of use or non-use.

It is essential to note that the availability of a time attribute as a continuous variable in this study provided an additional analysis to estimate the willingness to give up time in minutes to walk from home to access the woodland with improvements in each attribute.

6. Discussions and consultations with experts

The whole undertaking explained above required much time and benefited from numerous discussions and consultations with different groups of people, experts in the field of health and environment, health economics, and landscaping to ensure that relevant attributes and levels were included and that the framing was appropriate.

7. Piloting

Three pilot surveys were undertaken before the final SPDCE design and survey. The feedback from the pilot surveys was used to refine and reframe the attributes and levels of the SPDCE. For example, responses to follow-up questions of the first pilot survey suggested that the framing of two the attributes and their levels was unclear and not well defined. The first attribute related to quality of the woodland which was initially phrased as: *"your thoughts on the quality of the woodlands"*. The second attribute related to the opportunities for social activities that the woodland offers which was presented as *"the social opportunities that the woodland offers…"* Furthermore, the word *"choose"* was changed to *"prefer"* in the choice question *"which woodland would you choose?"* The word *"prefer"* was considered to be more suitable in this case as it implies identifying the trade-offs between attribute levels as opposed to *"choose"* which would commonly mean *"demand"*. In addition, other suggestions from consultations with experts recommended the provision of more information to the respondents in the initial context setting of the questionnaire about the choice tasks at hand and to bold important terms and phrases to highlight the salient attributes and associated levels as shown in the SPDCE questionnaire in Appendix 8. These changes would help identify important differences between the attributes and clarify the context of the choice tasks. In turn, they could facilitate the trading-off of attribute levels between the alternatives of the woodland presented.

There was also a recommendation to include the consistency and reliability tests in the SPDCE. The consistency test includes a choice set which is clearly dominant, theoretically, on attribute levels. This is used to check the rationality in choice decision-making to ensure that respondents understand the concept of the SPDCE when expressing their preferences. As regards the reliability check, it is simply the re-insertion of one choice set from the original design somewhere later in the questionnaire. This checks the degree of replicability of measurement over time and over different respondents. Having considered the attributes and their associated levels, the next stage was to construct an experimental design for the SPDCE. The next section discusses the experimental design process.

7.3.2 Construction of an experimental design

An experimental design is typically a matrix of values that represents attribute-levels and is used to map attributes and their associated levels into sets of alternatives which respondents choose from in a SPDCE survey. More details on experimental design are presented in chapter two. The question that arises is how best to allocate the attribute levels in a matrix. For the SPDCE in this chapter, three steps were undertaken to construct its experimental design: first, was the decision on the type of coding to use to assign the values of levels in the matrix location in a systematic manner while obeying some pre-determined statistical dimensions; second, was the clear specification of the SPDCE model to use; and third, was the determination of the type of experimental design to use.

As regards the coding of levels, of interest to this SPDCE was the effects coding, also known as orthogonal coding. This type of coding offers the possibility of testing that the attributes are not correlated with each other in the design (Hensher et al., 2015; Ryan et al., 2012b). It should be noted that the attribute-levels that were continuous (time and cost) were not effects coded but used their actual values instead as recommended (Bech and Gyrd-Hansen, 2005). Table 7-1 below shows the effects coding for the attribute levels of this SPDCE:

Identified attributes	**Effects level coding**
1. The woodland environmental support which we define as one which allows you to do the things you want to do, either on your own or with others (such as exercise, relaxing, enjoying wildlife) and makes it easy and enjoyable to do them	-1, 0, 1
2. The time that it takes to walk from home to the woodland	5mins,15mins, 50mins
3. The quality of the woodland environment which include cleanliness; the condition of paths and entrances; the naturalness of its appearance; the views of plants and wildlife	-1, 0, 1
4. The opportunities for social activities that the woodland offers you such as meeting people, community events, guidance on how to use the woodland and about what is going on there	-1, 0, 1
5. Cost for access to the woodlands, if you imagine you lived in a country where you had to pay for access to it in a form of an annual subscription	£0, £15, £50

Table 7-1: Effects coding of attribute levels.

After coding the attribute levels, it was important to understand the specific choice problem that the experimental design was required for by way of clearly specifying the model. This entailed making some important considerations. Firstly, the number of alternatives required for the SPDCE was decided. The configuration of the attribute-levels, clearly resulted in two alternatives: woodland A which was good; and woodland B which was better in terms of attribute-levels. To reflect real choice decision-making situation, an opt-out alternative of choosing 'none of these' of the two types of woodlands was included. It was also important to determine whether the alternatives should be labelled or unlabelled. Two generic (unlabelled) alternatives were preferred with each alternative sharing the same generic parameters (attributes): woodland A; woodland B; and 'none of these' option. The unlabelled alternatives allowed the estimation of WTP through marginal rate of substitution (MRS) which is explained later in the chapter.

In Table 7-1 above, the level ranges for the continuous attributes "time" and "cost" are 0, 15, 50 and £0, £15, and £50, respectively. These level ranges were carefully chosen to achieve a balance of them not being too narrow or too wide. As explained in chapter two, it has been found that wider ranges result in smaller standard errors, hence statistically preferable than narrow ranges, although in some cases, too wide ranges can be problematic as they can lead to dominant choice alternatives to govern the SPDCE (Kjær, 2005; Rose and Bliemer, 2008; Rose and Bliemer, 2009; Choicemetrics, 2014).

All the above considerations resulted in the model specification (2) described in chapter two, showing the probability that a respondent would choose a particular woodland configuration of attributes and levels. It is important to note that the utility function for the option 'none of these' is zero. The implication of this is that it is not useful in the model specification but useful during the analysis and does not affect the experimental design in any way even when imposed later in the design (Street and Burgess, 2007; Ryan et al., 2008b).

Another consideration is that of degrees of freedom as shown in (3) in chapter two. This resulted in six choices sets as minimum using the rule of thumb formula (3) as below:

$$S \geq \frac{6}{(2-1)} = 6$$

Where S is the choice situations. It is essential to note that, in this case, the unique observation is two and not three (thus either making a choice or not). The consideration for the degrees of freedom was looked at together with another consideration of attribute-level balance (each attribute level should appear an equal number of times for each attribute) in the matrix (Rose and Bliemer, 2009). In this study, there were three levels for each attribute which yielded a minimum of three choice sets to satisfy the attribute level balance property. However, since the attribute-level balance consideration is looked at in combination with that of the degrees of freedom then this study required a minimum of 6 choice sets.

In general, generating a design that takes into consideration all of the above can be complex (Johnson et al., 2013). However, special SPDCE software design programs are commonly used (Carlsson, 2011). This study used Ngene, a commercial software (version 1.1.2) to generate an initial experimental design for the pilot surveys and the final design for the study (Choicemetrics, 2014). The limitation of Ngene software is that it is strictly used for experimental designs and cannot be used for analysis. Recently, a Stata module known as '*Dcreate*' has been developed for use in constructing experimental designs for SPDCEs (Hole, 2015). Due to time constraint, it was not possible to compare the experimental designs constructed using Ngene and *Dcreate* applications in this study. However, since Stata is readily available to many researchers, and can also be used for SPDCE analysis, it can be speculated that *Dcreate* would be widely adopted in SPDCEs experimental designs because of these advantages.

Having decided on all these considerations, a decision had to be made on the type of the experimental design to use. The options were either to use a full factorial or a fractional factorial design. The difference between these designs was discussed in chapter two. A full factorial design includes all possible combinations of attribute-levels such that including all five attributes, each with three levels, could have resulted in a full factorial design of 243 possible combinations of levels. This could have resulted in many choice-sets to be presented to respondents. As this was practically not feasible, this SPDCE followed the current practice of using a statistically reduced design in terms of combinations of levels but still capable of estimating the main effect in the SPDCE model, which is known as a fractional factorial design (Louviere et al., 2000; Rose and Bliemer, 2009; Choicemetrics, 2014; Pfarr et al., 2014).

A fractional factorial design should have the good qualities discussed in chapter two, which include: orthogonality, which is zero correlation between attribute levels of choice alternatives to allow for independent determination of each attribute's influence on observed choices; attribute level balance which requires that all levels of each attribute should appear with equal frequency across profiles in order to obtain information about each attribute without prejudice on one another; minimal overlap of levels which means that the probability of repeated attribute-level within a choice set is minimized in order to provide maximum information about respondents trade-offs; and utility balance which means that the alternatives in choice sets should be close in utility space for respondents in order for them to have equal chances of being chosen. It was also noted previously in chapter two that, in many cases, it is impossible to create an experimental design that satisfies all these four qualities at once because some of them may conflict with each other (Huber and Zwerina, 1996).

A fractional factorial experimental design for the SPDCE of this study followed a recommended two-staged process: first, an initial design was constructed based on the principle of orthogonality, a purely statistical specification that ensures that attributes in the design are not correlated with each other (Rose and Bliemer, 2009; Can and Alp, 2012; Domínguez-Torreiro, 2014). The initial assumption of an orthogonal design is zero prior information about the strength and or direction of individual preferences (Bliemer et al., 2008; Rose and Bliemer, 2009; Hensher et al., 2015). This design was used for pilot surveys whose results (coefficients) were used as prior information in the second stage.

The second stage involved constructing another design which is known as an optimal or efficient design which uses the prior information obtained in the first stage. This optimal or efficient design does not aim at achieving orthogonality but seeks to optimise the statistical efficiency of the SPDCE model in terms of reducing the standard errors of parameter estimates (Scarpa and Rose, 2008). The reduction of the standard errors in the estimated SPDCE parameters results from the prior information used. This, in turn, results in a small sample size requirement, hence cost saving (Louviere et al., 2008). After constructing an experimental design, the next step was to develop and administer the SPDCE questionnaire. The next section looks at how this was done.

7.3.3 Questionnaire development and administration

The output of the experimental design was framed into a questionnaire instrument to be presented to a sample of respondents. A total of 18 choice sets were generated from Ngene software instead of the minimum six choice sets from the rule of thumb formula of degrees of freedom. According to the developers of Ngene, the reason for this large number of choice sets is that the software could not generate a design with a small number of six choice sets which satisfied the qualities of a good design

(Huber and Zwerina, 1996), particularly orthogonality, hence the next available orthogonal design had 18 choice sets (Rose and Bliemer, 2009).

The actual SPDCE questionnaire

As can be seen from Appendix 8, at the start of SPDCE questionnaire, respondents were introduced to the subject of research and to the researchers. This was followed by an explanation of the context of the survey, the attributes, a provision of an example of the choice task, an emphasis on the importance of participation and confidentiality. Respondents were told where to direct queries and their time commitment on the survey. Some guidance on how to proceed answering the choices questions was also given. These aspects of a questionnaire are important in any good survey (Bennett, 1999).

Next, respondents were told to imagine a situation in which they had access to either woodland A or B that differed from each other in their attribute-levels. Thinking about the woodland attributes given to them, they were asked to choose which woodland they preferred as ideal. To mimic the real choice decision making situation, they also had an option of choosing none of the woodland options.

These attributes had levels, except for the ‘none of these’ option, which were in turn varied across their ranges to define each alternative. The respondents were then asked to choose among the three alternatives. The process proceeded in an iterative way for the 18 choice sets in order to build up a set of trade-off preferences for each respondent (Burgess et al., 2012; De Ayala et al., 2012; Vecchiato and Tempesta, 2013).

After the choice questions, respondents were asked further questions on whether there were other attributes of woodlands that they considered important but were not included. They were also asked the level of difficulty of the choice tasks on a five-point scale of very easy, easy, ok, difficult and very difficult. The questionnaire went further to ask respondents' socio-economic characteristics.

Prior to data collection for the pilot and final surveys, ethical approval was sought from the University of Glasgow ethics committee for this SPDCE. Ethics clearance was necessary because the survey involved respondents' time. The approval was granted under project reference number 200140011 as shown in Appendix 7.

Another important consideration before collecting data was the sampling frame and the size of the sample to be used for pilot and final surveys. Since the study sought to value the attributes and levels of woodlands, the appropriate main sampling frame for eliciting preferences was the general public rather than the WIAT study population because woodlands are public goods funded by the tax payer which implies other benefits to the society forgone. Furthermore, the WIAT study targeted deprived communities, hence their WTP estimates would not be representative.

As regards the sample size, there is no general consensus or guidance on the appropriate sample size for a SPDCE as previously discussed in chapter two. There is also no definitive statistical formula to calculate the appropriate sample size partly because of many complexities relating to the level of difficulty of questionnaires, question format, number of alternatives, the number of attributes and levels, the required precision of the results, the expected variability of choices made, and any proposed subgroup analysis for the SPDCE (Marshall et al., 2010). Previous studies have shown that sample sizes of 40-100 respondents may be sufficient for reliable statistical analysis (de Bekker-Grob et al., 2013). Orme (2006) proposes a total of 300 respondents for robust quantitative research and a minimum of 200 per group for subgroup analysis (Marshall et al., 2010; Rose and Bliemer, 2013).

While a large sample size may provide robust results and give the statistical power of a SPDCE, practically, large sample sizes are costly and difficult to obtain and a poor experimental design may further compromise the ability to retrieve meaningful statistical parameter estimates (Rose and Bliemer, 2009), even with a large sample. A small sample size, on the other hand, may lessen the reliability of the parameter estimates.

Efficient designs have the potential benefit of reducing confidence intervals of parameter estimates in a SPDCE model hence permitting the use of reduced sample sizes (Kerr and Sharp, 2009). The argument put forward for use of small sample size when an efficient design has been obtained is that efficient designs result in larger decreases in standard errors than those obtained when larger sample size is used (Rose and Bliemer, 2009). However, the drive to use statistically efficient designs may have severe unintended consequences where respondents focus on some attribute-levels while ignoring others, a behaviour known as heuristics, which result in non-attribute attendance (Flynn et al., 2016), as noted chapter two.

The minimum sample size for this study was calculated using the rule of thumb formula (4) discussed in chapter two for calculating SPDCE samples. This resulted in the minimum sample size of 25.

$$N \geq 500 \times \frac{3}{3 \times 20} = 25$$

This is so far the best guidance in the absence of empirical evidence on SPDCE sample size in healthcare (Marshall et al., 2010). The justification for this sample size is that it would yield minimum number of observations enough to estimate a robust model (Hensher et al., 2015). In this case, with the 20 choice tasks for the final design including the consistency and reliability tests, the observations would be 500 (25 x 20 choice sets). The final survey for the SPDCE used (n= 510) respondents to have robust results, well above the estimate from the rule of thumb by an order of magnitude. However, in general, relying on rules of thumb is problematic because such rules cannot be strictly accurate and reliable (de Bekker-Grob et al., 2015).

Following the development of the SPDCE questionnaire, three pilot surveys were undertaken. The results of these pilot surveys helped to improve the final SDPCE design, and were critical in gauging its feasibility.

The first one used an orthogonal design known as the optimal orthogonal in the differences (OOD) design based on Street et al. (2005) . This is a special type of an orthogonal design which has absolutely no attribute level overlaps. It maximizes the differences in the attribute-levels across alternatives hence does not allow attribute-level overlaps thereby forcing the trading-off of all attributes in the choice set. This pilot survey was conducted online with a sample of 60 adult members of the general public (aged 16 and above) recruited across Scotland. These respondents were members of the panel of a market research company, ResearchNow. These members are compensated through a reward point system. When a given threshold is reached, the accumulated points can be redeemed as gift vouchers or charity donations. ResearchNow used a targeted invitation strategy to ensure that the sample was representative. This resulted in having a sample comprised of 50% males and 50% females. Other than using ResearchNow, another option was to personally administer the questionnaire using pen and paper. As noted earlier, a SPDCE survey involves presenting individuals with many choice sets to complete. To obtain completed responses from pen and paper survey could have been challenging and slow. Using the survey company and online questionnaire resulted in quick complete responses, and provided respondents with the opportunity to think through their choices without being influenced by the interviewer.

The absence of attribute-level overlaps in the OOD design meant that respondents were presented with an extremely difficult choice making decision task (Johnson et al., 2013). Each attribute in a choice set was set to be different across alternatives which potentially promoted a particularly dominant attribute to govern the experiment and resulted in non-attribute attendance (Lagarde, 2013; Choicemetrics, 2014; Flynn et al., 2016).

When the SPDCE pilot survey was analysed in Stata 13 (StataCorp, 2013), the SPDCE model did not converge which indicated that there was a problem (Hensher et al., 2015). A diagnosis of different attribute-level combinations was done through tabulation in Stata to trace the problem. The results revealed that respondents were not making trade-offs of the attributes and levels in their choice making decisions. The possible problems pointed to the OOD experimental design used which did not allow any attribute-level overlap, hence, respondents were forced to trade-off all the attributes in the choice set (Kragt and Bennett, 2012; Flynn et al., 2016; Flynn, 2016). This might have been particularly difficult.

The quest for a highly statistically efficient design by using the OOD design might have negatively affected the responses to the SPDCE by promoting some heuristics. This experience is consistent with the recent findings of Flynn et al. (2016) and Flynn (2016). In their studies, it was found that the majority of respondents did not trade across attributes in designs with no attribute-level overlaps such as the OOD designs based on Street et al. (2005).

The collapse of this pilot study led to a rethink of the SPDCE experimental design. The first consideration was to change the design. In addition to changing the design, it was thought that the survey questionnaire would benefit if reliability and consistency tests were included. The second pilot incorporated all these changes including the use of an ordinary orthogonal fractional design. Bolding of all important terms and phrases throughout the SPDCE survey was also done to highlight salient attributes and their associated levels to facilitate the understanding of the context of the SPDCE.

The second pilot model had run as expected and used a convenience sample (n=23) from the general public in Glasgow, Scotland. Its results were used as priors in the creation of an efficient experimental design for the third pilot survey (Rose and Bliemer, 2009). The third pilot used an efficient SPDCE design with prior information (coefficients from the second pilot), and aimed to test the feasibility of the SPDCE survey prior to launching the final survey. It was mainly administered randomly to colleagues in Public Health and Health Economics and Health Technology Assessment (HEHTA) group at the University of Glasgow. There were 20 respondents. This sample size was deemed to be adequate to test the final survey. The analysis of this pilot survey showed that the final survey was feasible.

Then, the final survey (n=510) was conducted online by a market research company which was used in the first pilot survey, called ResearchNow. Questions may be asked about the unresolved problems of representativeness of online surveys and that they require respondents to be computer literate, which may be a limitation (Pearce et al., 2002; Kjær, 2005; Shah et al., 2015). However, it was considered that the advantages of an online SPDCE survey outweighed the disadvantages. For example, some advantages of online surveys for SPDCEs include: offering flexibility to respondents in terms of response time; providing independent treatment of each choice set presented to respondents at each click of the button so that each choice set is not compared to any other set in the survey; and being relatively quick, hence cost saving (Pearce et al., 2002; Hensher et al., 2015).

The survey was administered to a representative sample of members of the general public (aged 16 and above) across Scotland. Like the first pilot, the respondents were panel members who are compensated through a reward point system. When a given threshold is reached, the accumulated points can be redeemed as gift vouchers or charity donations.

In the final survey, the same respondent was presented with 20 choice sets to complete at a point in time in one go. There were 510 completed responses to the final SPDCE questionnaire which yielded 20,400 observations, with the exclusion of a "*none of these*" option. No follow-ups through reminders were required since the survey was targeted such that there no refusals to complete the SPDCE questionnaire or cases of objecting the use of a cost attribute which would imply paying for access to woodland. Once the SPDCE data was collected, the next stage was to input the data in a form that it could be analysed. The following section discusses how this was done.

7.3.4 Data input and analysis

The set-up of data was done in advance using a dummy dataset in Microsoft Excel, then transferred to Stata 13 (StataCorp, 2013) to ensure the feasibility of analysing the actual SPDCE data (Ryan et al., 2008b). All data were set-up as a panel so that each row of the dataset represented one alternative for one respondent (Long and Freese, 2014).

Effects coding was used for coding categorical or qualitative attribute-levels while continuous variables assumed their actual values in the model (Bech and Gyrd-Hansen, 2005; Mercer and Snook, 2005; Bridges et al., 2011; Hensher et al., 2015; Ryan et al., 2012b). An additional variable representing a dependent choice outcome was created to signify the choice decision made for each alternative in a choice set. Then, the SPDCE data were analysed in Stata 13 (StataCorp, 2013) using the recommended multinomial logit (MNL) model, also commonly known as the conditional logit (clogit) model as a starting point (Hauber et al., 2016). This model has the capability to capture of the dynamics of repeated choices since it assumes that the unobserved factors are independent over time in repeated choices situations (Train, 2009; Vojáček and Pecáková, 2010; Long and Freese, 2014). Furthermore, it is easy to use and interpret such that it is considered as a "workhorse" of SPDCEs (Hensher and Greene, 2003; Kjær, 2005; Rose and Bliemer, 2009; Long and Freese, 2014; Hensher et al., 2015).

The socioeconomic characteristics of the respondents of the SPDCE were not added into the regression model directly because they normally do not vary when making a particular choice in the SPDCE questionnaire (Ryan et al., 2008b). However, these could be added to the model later as interaction terms with the attributes to enable the understanding of whether responses vary with the given characteristics (Ryan et al., 2012b). This was not done in this SPDCE because the aim was to estimate WTP values. Further analysis of the SPDCE could also involve supporting questions to the SPDCE survey which would help to clarify or explain some decision strategies used by respondents when making choices (Pearce et al., 2002; Kjær, 2005; Krupnick and Adamowicz, 2007; Carlsson, 2011; Kreye et al., 2014; Hensher et al., 2015). The next section presents the results of the final SPDCE survey.

7.4 Results

From a sample of 510 respondents, only 32% chose a "*None of these*" option (32%) which implied that only 68% of the responses were used for analysis. About 88% of the respondents passed the reliability test while 84% passed the consistency test and only 3% failed both tests, which implied that the responses were rational and consistent. Hence, all responses were included in the analysis. Furthermore, because of the negligible number of those who failed both tests (3%), dropping them made no significant difference to the results. The recommendation is that those who fail these tests should not be removed from the sample as they may have valid reasons for doing so and random utility models (RUM) models are robust to errors made by individuals when making preferences (Ryan et al., 2012b).

There was a follow-up question asking if there were other woodland attributes that respondents considered important but that were not included. None of the respondents identified any new attribute. This means that all the identified attributes were relevant as revealed in the results later in the section. The choice tasks were rated as very easy, easy, and ok by 92% of respondents which implies that the SPDCE choice tasks and survey in general, was realistic and not complex.

The characteristics of the sample for the final SPDCE are presented in Table 7-2 below. The mean age of the respondents was 47. A representative sample of the Scottish population was used (n=510) which resulted in 52% males and 48% females. 79% were the white Scottish population while other ethnicities comprised of 21%.

Category	n (%)
Gender	
1. Male	264 (52%)
2. Female	246 (48%)
Average age (years)	47
1. 16-24	58 (11%)
2. 25-34	83 (16%)
3. 35-44	89 (17%)
4. 45-54	85 (17%)
5. 55-64	89 (17%)
6. 65+	106 (21%)
Occupational status	
1. Working part-time	75 (15%)
2. Working full-time	210 (41%)
3. Not working	110 (22%)
4. Student	36 (7%)
5. Other	79 (15%)
Level of education	
1. Secondary school	89 (17%)
2. Vocational/trade/college	75 (15%)
3. Higher/A levels	97 (19%)

	4. University	242 (47%)
	5. other	7 (1%)
Ethnicity		
	1. White Scottish	403 (79%)
	2. Other European	73 (14%)
	3. Mixed	8 (2%)
	4. Indian /Indian Scottish or British	3 (1%)
	5. Other	12 (2%)
	6. Do not want to state ethnicity	11(2%)
Household income in the last 12 months		
	1. < £3,900	12 (2%)
	2. £4,000-£19,999	112 (22%)
	3. £20,000-£31,999	108 (21%)
	4. £32,000- £55, 999	153(30%)
	5. £56,000+	71 (14%)
	6. Prefer not to say	56 (11%)
Children under 16 years in the household		
	1. Yes	136 (27%)
	2. No	374 (73%)
Dog ownership		
	1. Yes	118 (23%)
	2. No	392 (77%)
Disability		
	1. Yes	51 (10%)
	2. No	452 (89%)
	3. Prefer not to say	7 (1%)

Table 7-2: Socio-economic characteristics of the respondents

A mixed (MXL) logit also known as a random effects or random parameter logit (RPL) model (Hauber et al., 2016) was used to analyse the final SPDCE data in Stata 13 (StataCorp, 2013). The model assumes that the probability of making a choice from alternatives depends on the attributes of the alternatives and individual- specific variations in preferences and also controls for the within and across variability of respondents (Train, 2009). Table 7-3 below presents the results of the main SPDCE:

Attribute	Coef	SE	95% CI	Odds ratio	SE	95%CI	P value
The supportive woodland environment that allows you to do enjoyable activities easily (Base level-No support)							
Some support	0.84	0.043	0.76-0.93	2.32	0.10	2.14-2.53	<0.001
A lot of support	1.03	0.043	0.95-1.12	2.80	0.12	2.58-3.05	<0.001
the time it takes to walk from home to the woodland	-0.02	0.001	-0.03--0.02	0.98	0.00	0.97-0.98	<0.001
The quality of the woodland environment (Base level-Poor quality)							
average quality	1.04	0.044	0.95-1.13	2.83	0.12	2.60-3.08	<0.001
Good quality	1.25	0.044	1.17-1.34	3.50	0.15	3.21-3.82	<0.001
The opportunities for social activities (Base-level-No opportunities)							
some opportunities	0.31	0.045	0.22-0.39	1.36	0.06	1.24-1.48	<0.001
Many opportunities	0.43	0.042	0.35-0.51	1.53	0.06	1.41-1.66	<0.001
The cost for access to the woodland	-0.04	0.001	-0.038--0.35	0.96	0.00	0.96-0.97	<0.001

Table 7-3: SPDCE results from the random parameter logit (RPL) model showing coefficients and odds ratios for the attributes and levels.

The results of the final SPDCE in Table 7-3 above show that all the coefficients of the attribute-levels were statistically highly significant

($\rho<0.001$). All the coefficients passed the theoretical validity with the expected sign. This means that all the identified attributes have an influence on individuals when making choices on which woodlands to visit. Quality of the woodland environment was most influential in woodland choice decision with the following odds ratios: average environmental quality (OR 2.83, 95%CI: 2.60 to 3.08); good environmental quality (OR 3.50, 95%CI: 3.21 to 3.82). This was followed by the supportive woodland environment attribute: some support (OR 2.32, 95%CI: 2.14 to 2.53); a lot of support (OR 2.80, 95%CI: 2.58 to 3.05) while the attribute relating to the opportunities for social activities was preferred least: some opportunities (OR 1.36, 95%CI: 1.24 to 1.48); and many opportunities (OR 1.53, 95%CI 1.41 to 1.66).

Furthermore, the attribute coefficients related to the cost for access to woodlands and the time it takes to walk from home to the woodland had negative association with the choice of woodland decision. Thus, the cost for access (OR 0.96, 95%CI: 0.96 to 0.97) representing a 4% decrease in woodland visits for any unit (one pound) increase in a yearly subscription of the cost for access to woodlands. The same trend would happen for any minute increase for the time it takes to walk from home to the woodlands (OR 0.98, 95%CI: 0.97 to 0.98) which represents a 2% reduction of woodland visits.

Further estimates were made regarding the trade-offs that individuals made between the attributes of the woodlands. This allowed the calculation of their willingness to pay for access in a form of a yearly subscription to access woodland with improvements in a given attribute; and the time in minutes that individuals are willing to walk from their homes to access a woodland with a given attribute. This was calculated using the coefficients of the random parameter logit model which were statistically significant as the ratio of the attribute of interest divided by the negative of the coefficient on the cost attribute and the time attribute respectively. The results are shown in Table 7-4 below:

Attribute	Coef	WTP for access as an annual subscription $-(\frac{\beta attribute}{\beta cost})$	95% CI	Willingness to give up time (minutes) to walk from home to woodlands $-(\frac{\beta attribute}{\beta time\ attribute})$	95% CI
The supportive woodland environment that allows you to do enjoyable activities easily (Base level-No support)					
Some support	0.844	£ 23.18	£20.61 -£25.88	36 minutes	31-40mins
A lot of support	1.031	£ 28.32	£25.63-£31.19	44 minutes	39-49mins
the time it takes to walk from home to the woodland	-0.024	-£ 0.65	-£0.58 - -£0.71		
The quality of the woodland environment (Base level-Poor quality)					
average quality	1.040	£ 28.58	£25.83-£31.50	44 minutes	39-50mins
Good quality	1.254	£ 34.45	£31.53-£37.57	53 minutes	48-60mins
The opportunities for social activities (Base-level-No opportunities)					
some opportunities	0.305	£ 8.39	£5.94-£10.89	13 minutes	9-17mins
Many opportunities	0.428	£ 11.76	£9.43-£14.14	18 minutes	14-22mins
The cost for access to the woodland	-0.036			-2 minutes	-1 - -2mins

Table 7-4: WTP and willingness to give up time (minutes) to access a woodland

The results above reveal that individuals are willing to pay more to access woodland with good environmental quality followed by good environmental support and many social opportunities, in that order. They are not prepared to pay anything to access a woodland that is far away from their homes as revealed by the negative WTP of -£0.65, (95%CI: -£0.58--£0.71). Thus, they are willing to pay as follows: 1) £28.58, (95% CI: £25.83-£31.50) as an annual subscription to access the woodland with an average environmental quality and £34.45, (95% CI: £31.53-£37.57) for the woodland with good environmental quality. This comes up to £5.87 for the additional improvement in environmental quality from average to good. Then, 2) they are willing to pay £23.18, (95% CI: £20.61-£25.88) for woodland with some environmental support and £28.32, (95% CI: £25.63-£31.19) for woodland with a lot of environmental support. This makes £5.15 for the additional improvement from some environmental support to a lot of environmental support. Individuals are further willing to pay £8.39, (95% CI: £5.94-£10.89) for the woodland that offers some social opportunities and £11.76, (95% CI: £9.43-£14.14) for that which offers many social opportunities. This means that they are willing to pay £3.37 for that additional improvement from some social opportunities to many social opportunities.

In terms of their willingness to give up time in minutes to walk from home to a woodland, it was found that on average individuals were willing to give up 53 minutes, (95% CI: 48-60 minutes) of their time to walk from home to the woodland which has good environmental quality, and 44 minutes, (95% CI: 39-50 minutes) for a woodland with average environmental quality; they were prepared to sacrifice 44 minutes, (95% CI: 39-49 minutes) and 36 minutes, (95% CI: 31-40 minutes) of walking time to access a woodland that offers a lot of environmental support and some environmental support, respectively; and were ready to spend 18 minutes, (95% CI: 14-22 minutes) walking to access the woodland that offers many social opportunities, and 13 minutes, (95% CI: 9-17 minutes) for a woodland that provides some social opportunities. Lastly, it was also shown that they were not willing to give up their time to walk to a woodland that was needed paying for access as shown by the negative willingness to give up time in minutes (-2, 95% CI: -1- -2 minutes).

Immediately after the SPDCE questions were follow-up questions. The next section presents the findings from these follow-up questions. The follow-up questions aimed to obtain insights into additional attributes of importance which were not originally included in the SPDCE survey design (Ryan et al., 2008a; Coast et al., 2012; Pfarr et al., 2014); and also sought to check if the survey was well-framed to make sure that it was clear and well understood to respondents.

In the final survey, 80% of the respondents felt that all important attributes were included in the SPDCE while 20% considered that other attributes could have been included in the SPDCE. These suggestions included attributes such as woodlands that are free of dog litter, or woodlands with wardens to conduct spot-checks to ensure that dogs were on lead; woodlands with dog-free areas; and woodlands with park rangers to ensure the safety and well-being of users against anti-social behaviour. As regards the complexity of the choice tasks that were presented to respondents for the final SPDCE survey, 51% said the SPDCE was very easy or easy. Cumulatively, 91% considered the choice tasks to be very easy, easy, or ok while the remaining 9% felt that they were either difficult or very difficult to complete as shown in Table 7-5:

How easy/difficult	Frequency	Percent	Cum
Very easy	100	19.61	19.61
Easy	164	32.16	51.76
OK	201	39.41	91.18
Difficult	38	7.45	98.63
Very difficult	7	1.37	100

Table 7-5: The level of SPDCE complexity

The results above suggest that the choice tasks were generally acceptable. Further follow-up question aimed to establish the time that individuals were willing to sacrifice to access the woodland, the respondents were asked to indicate the time (in minutes) that they are willing to walk from their homes to the woodland. It was found that they were willing to walk, on average, for 24 minutes (95% CI: 23-25 minutes) from home to access a woodland as show in Table 7-6 below:

Willingness to walk (minutes)

N	Mean	Std dev	Std Error	95% CI	Min	Max
510	23.98	13.5	0.6	22.8-25.2	0	60

Table 7-6: Willingness to walk (minutes)

The smallest amount of time in minutes that individuals were prepared to sacrifice to walk from home to a woodland was zero minutes (not prepared to walk) while the highest amount of time was 60 minutes (1hour). These results are within the range of the attribute-levels related to time (5mins, 15mins, 50mins) that were used in the study.

A follow-up question on willingness to pay to access the woodlands revealed that 27% of the respondents did not want to pay to access woodlands. Further to this question, respondents were asked to state the reasons why they would protest paying for access to woodlands. Content analysis, a technique used to systematically classify open-ended responses to survey questions (McIntosh et al., 2010), was used to identify broad themes of the responses to the follow-up question. Three broad themes were identified from the responses: 1) natural endowment; 2) affordability; and 3) non-use.

The first reason for not being prepared to pay for access to woodlands was that woodlands were a natural endowment, therefore, free for everyone to use and that they should be left to self-manage themselves naturally. Second reason was that of affordability in the current economic downtime.

Others suggested that an indirect cost would be preferable such as a donation or money could be raised through the provision of amenities such as cafes, and shops which could be funded from taxes. The third reason was that of non-use. These are people who do not use woodlands either out of lack of interest, or do not use the woodlands enough to warrant paying for access to them. A cost attribute is clearly problematic as it might not reflect how individuals access woodlands in Scotland where access is normally free. As a way of mitigating this problem, respondents were asked to assume they lived in a country where they had to pay to access woodlands and were assured that the cost attribute would not, in any way, affect the way they normally use woodlands.

Having presented the results of the SPDCE, the next section discusses the valuation of the incremental changes or improvements in the attributes and levels that measured the non-health benefits resulting from the intervention.

7.5 The valuation of the non-health outcomes of the WIAT intervention

The unbalanced panel data of the WIAT study was used to calculate the incremental changes or improvements in the attributes and levels resulting from the intervention for the base case analysis. A method that uses summary measures, commonly known as an area under the curve was used (Matthews et al., 1990; Manca et al., 2005). This method is commonly used in the calculation of quality- adjusted life years (QALYs). In this case, this approach considered respondents' responses to construct a single number which reflected their overall response curve. Then, the area under the response curve was the change or improvement in attributes or levels that can be attributed to the intervention at an individual level using the unbalanced panel data. This was followed by the use of the cross-sectional data to calculate the incremental changes or improvements at population level.

A summary measure approach used the product of the time difference and the average of the proportion of responses for two measurements (Matthews et al., 1990). Thus, for proportion measurements X_1, X_2 and X_3, which represent the proportion of responses at wave one, two and three (T_1, T_2, and T_3), the formula becomes:

$$(T_2 - T_1)\,(X_1 + X_2)/2 + (T_3 - T_2)\,(X_2 + X_3)/2 \qquad (12)$$

Table 7-7 below presents the calculation of the incremental changes or improvements in the attributes and levels resulting from the intervention for the unbalanced panel analysis:

B4-What kind of activities do you pursue when using these local woodlands?					**Change/improvement in environmental support for activities**	
Site	Attribute level	Wave1	Wave2	Wave3		
		% of responses	% of responses	% of responses	Some support	A lot of support
Control	Some support	0.01	0.08	0.06	-0.11	0.11
	A lot of support	0.99	0.92	0.94		
Intervention	Some support	0	0	0.01		
	A lot of support	1.0	1.0	0.99		
B11-How long would it take you to walk to these local woodlands?					**Change/improvement in access to woodlands in terms of time (mins)**	
Site	Attribute	Wave1	Wave2	Wave3		
		Responses in mins	Responses in mins	Responses in mins		
Control	Time in minutes-continuous	10.34	4.06	11.08	4.69	
Intervention	Time in minutes-continuous	13.6	8.32	8.68		
B29-Overall, what do you think about quality of these local woodlands?					**Change/improvements quality of woodland environment**	
Site	Attribute level	Wave1	Wave2	Wave3		
		% of responses	% of responses	% of responses	Average	Good
Control	Average	0.31	0.32	0.26	0.14	0.15
	Good	0.44	0.44	0.57		
Intervention	Average	0.46	0.34	0.34		
	Good	0.38	0.61	0.59		

H1-To what extent do you agree or disagree that people in this neighbourhood pull together to improve the neighbourhood?					Change/improvements in social cohesion	
Site	Attribute level	Wave1	Wave2	Wave3		
		% of responses	% of responses	% of responses	Some opportunities	Many opportunities
Control	Some opportunities	0.56	0.69	0.51	-0.21	0.14
	Many opportunities	0.12	0.08	0.14		
Intervention	Some opportunities	0.6	0.49	0.48		
	Many opportunities	0.12	0.2	0.18		

Table 7-7: Incremental changes or improvements in the attributes and levels.

As argued by McIntosh (2006), there is no reason why the willingness to pay values from the society obtained from the SPDCE for the best configuration of attributes and level, cannot be used to estimate the total value of the incremental changes in attribute and attribute-levels resulting from the intervention. Following this argument, the WTP values from the SPDCE for this study were used to calculate the total willingness to pay for the identified non-health benefits of the WIAT intervention represented by the incremental changes or improvements in the attributes and levels of the good woodland as shown in Table 7-8 below using the unbalanced panel analysis:

WTP estimates for attribute and attribute-levels from SPDCE			95% CI	Incremental changes or improvements in attributes/levels		Value
Some support	wtps1	£23.18	£20.61-£25.88	Δs1%	-0.11	-£2.55
A lot of support	wtps2	£28.32	£25.63-£31.19	Δs2%	0.11	£3.12
Time	wtpt	-£ 0.65	-£0.58- -£0.71	Δt (mins)	4.69	-£3.05
Average	wtq1	£28.58	£25.83-£31.50	Δq1%	0.14	£4.00
Good	wtpq2	£34.45	£31.53-£37.57	Δq2%	0.15	£5.17
Some opportunities	wtpo1	£ 8.39	£5.94-£10.89	Δo1%	-0.21	-£1.76
Many opportunities	wtpo2	£11.76	£9.43-£14.14	Δo2%	0.14	£1.65

Table 7-8: Total WTP for the identified non-health benefits for unbalanced panel analysis

The total value of the identified non-health benefits for the configuration of a good woodland as a result of the WIAT intervention was calculated using the following formula:

$$Total\ value\ = \Delta s2\ \times wtps2\ +\ \Delta t\ \times wtpt\ + \Delta q2\ \times wtpq2 + \Delta o1\ \times wtpo1 + \Delta o2\ \times wtpo2 \quad (13)$$

Where Δs2 is the incremental change or improvement in attribute-level ‘a lot of support’; wtps2 is the societal willingness to pay for the attribute-level ‘a lot of support; Δt is the incremental change or improvement in attribute ‘time’; wtpt is the willingness to pay for an additional increase in time (minutes) to walk from home to the woodland; Δq2 is the incremental change or improvement in attribute-level ‘good quality’; wtpq2 is the societal willingness to pay for the attribute-level ‘good quality’; Δo2 is the incremental change or improvement in attribute-level ‘many opportunities’; and wtpo2 is the societal willingness to pay for the attribute-level ‘many opportunities’.

$$\begin{aligned} &\mathit{Total\ value\ of\ a\ good\ woodland} \\ &\quad = 0.11 \times £28.32 + 4.69 \times -£0.64 + 0.15 \times £34.45 \\ &\quad + 0.14 \times £11.76 = £6.89 \end{aligned}$$

It is essential to note that formula (13) above is part of the formula (7) for the integrated approach presented in chapter four. The results above suggest that there were positive changes or improvements in the following attribute-levels “a lot of support”; “Average environmental quality”; “good environmental quality”; and “many opportunities” because of the intervention at individual-level, thus, when the same individuals who were present in at least two waves including the first wave were followed up. Overall, the total willingness to pay for the incremental changes or improvements in the attributes and levels of the WIAT intervention in the two-year time horizon is £6.89. This represents the amount that individuals are prepared to secure the changes or improvements in the attributes and levels which is a proxy to the identified non-health related benefits of the WIAT intervention.

When the same analysis was done using cross-sectional data, the results in Table 7-9 reveal that, overall, there was negative change or improvement in the attribute-levels “good environmental quality” because of the intervention at population-level.

WTP estimates for attribute and attribute-levels from SPDCE			**95% CI**	**Incremental changes or improvements in attributes/levels**		**Value**
Some support	wtps1	£23.18	£20.61-£25.88	Δs1%	-0.04	-£0.93

A lot of support	wtps2	£28.32	£25.63-£31.19	Δs2%	0.04	£1.13
Time	wtpt	-£ 0.65	-£0.58- -£0.71	Δt (mins)	3.78	-£2.46
Average	wtq1	£28.58	£25.83-£31.50	Δq1%	0.26	£7.43
Good	wtpq2	£34.45	£31.53-£37.57	Δq2%	-0.10	-£3.45
Some opportunities	wtpo1	£ 8.39	£5.94-£10.89	Δo1%	-0.12	-£1.01
Many opportunities	wtpo2	£11.76	£9.43-£14.14	Δo2%	0.08	£0.94

Table 7-9: Total WTP for the identified non-health benefits for cross-sectional analysis.

The total value of the willingness to pay at population-level for the good woodland is £-3.79

$Total\ value\ of\ a\ good\ woodland\ =\ 0.04\ \times\ £28.32\ +\ 3.78\ \times -£0.64\ +\ -0.1\ \times\ £34.45 +\ 0.08\ \times\ £11.76$ =£-3.79.

It can be noted that the total WTP value at an individual-level is negative. This is not surprising because the samples are different. In this case, the unbalanced panel is perhaps a more powerful form of analysis despite the small sample size given that the same individuals who were present in at least two waves including the first wave were followed up.

7.6 Discussion

This chapter has demonstrated a stepwise process of undertaking a SPDCE to value the identified attributes and levels of woodlands. The challenges of carrying out a SPDCE have been highlighted such as: the choice of an experimental design which is very much dependent on the researcher; sample size calculation which has no clear guidance and is dependent on the rules of thumb; and the type of model for analysis which is dependent on the assumptions made. The chapter has shown the importance of engaging in discussions and consultations with different groups of people, experts and conducting pilot surveys during the process of identifying attributes and levels and constructing experimental designs for the SPDCE.

The results of the SPDCE suggest that the Scottish population values good woodland environmental quality highly. This includes cleanliness, good paths and entrances, good naturalness in appearance with views of varied plants and wildlife. These results are consistent with, and confirm quantitatively the findings of previous qualitative studies discussed in chapter three. For example, a study by Sugiyama et al. (2015) which sought the association between public open space attributes and recreational walking revealed that investing in a single high environmental quality park may be more effective in promoting health behaviours such as walking than providing many parks with an average environmental quality. A recent SPDCE in Berlin, Germany also found that environmental quality in terms of cleanliness and maintenance mattered most to individuals at any time when they visited parks (Bertram et al., 2017).

In addition to good woodland environmental quality, individuals in Scotland prefer woodlands that offer a lot of environmental support that allows them to do the things they want to do, either on their own or with others such as exercise, relaxing, enjoying wildlife and makes it easy and enjoyable to do them. This is followed by the preference for woodlands that provide many social opportunities such as meeting people, community events, and the provision of guidance on how to use the woodland and about what is going on in the woods at a given time. Generally, the Scottish population would not want to pay for access to woodlands and would not visit woodlands that are further away from home. Their willingness to pay to access woodlands is dependent on whether woodlands offer good environmental quality, a lot of environmental support, and many social opportunities, in that order. They are also willing to sacrifice their time measured in minutes to walk from home to access woodlands which offer good, and average environmental quality, a lot of environmental support, some environmental support, many, and some social opportunities, in that order. In addition, they are not prepared to sacrifice any of their time to walk to woodlands that require paying for access.

The results of the SPDCE have implications for policy. For example, consideration of the individual preferences for woodland attributes revealed from the SPDCE would help inform policy makers on prioritisation if woodlands are used as a public health intervention. Specifically, when considering investing in improving certain attributes of woodlands to increase access as a goal of a public health intervention to improve health and reducing inequalities at population-level.

As regards WTP estimates and willingness to give up time to walk from home to the woodlands, policy makers can use these estimates to make comparisons and rankings of desirability of woodland attributes and can be informed about how much people value attributes of woodlands.

While all this is very important for policy, the WTP values obtained from the SPDCE in this chapter was subsequently used in an economic evaluation to value the identified non-health outcomes of the WIAT intervention through the incremental or improvements in the attributes and levels of the woodlands resulting from the intervention.

The analysis of the non-health outcomes shows that the WIAT intervention was beneficial at least in the two-year period of analysis for both the unbalanced panel analysis even though the monetary benefits were minimal. This can be attributed to the short time horizon in which the economic evaluation was undertaken. As noted previously, interventions which aim to improve the society’s well-being take a long time to manifest (Tchouaket and Brousselle, 2013; Mays and Mamaril, 2015). There is, perhaps, need for a long-term follow up.

This analysis, however, has limitations. For example, the estimates for the payment vehicle used in the SPDCE were based on two English studies (Bateman, 1996; Edwards et al., 2009) which may not directly be transferable to another setting like Scotland. In addition, the assumption of a yearly subscription to access woodlands might not be realistic.

The other major limitation relates to the wider WIAT study which was not conceptualized and designed with the SPDCE in mind. As such, the mapping of the WIAT main study questionnaire items considered to measure the non-health outcomes to the attributes and levels of the SPDCE was pragmatic, hence problematic in terms of the assumptions made which might not be realistic of how woodlands are used. This could have implications on the results, hence, they should be interpreted with caution. However, subject to some of the assumptions underlying the mapping, the SPDCE can considered robust in terms of its WTP estimates.

Furthermore, the initial data collection design was a repeat cross-sectional at three waves (wave one-baseline, wave two after the physical intervention, and wave three after the social intervention). However, in the end, it turned out that some respondents were in one wave only, while others were in both wave one and two; wave one and three; and wave two and three, and some were in all the three waves. This meant that there were two types of analyses that could be carried out: unbalanced panel and cross-sectional analysis. Using the unbalanced panel analysis means that the same individuals were followed up in at least two waves from wave one. This would provide a true reflection of the changes or improvements in attributes or levels resulting from the intervention at an individual-level. It is because of this reason that the unbalanced panel analysis was used for base case analysis. However, the small sample size for the unbalanced panel data could be problematic to provide robust results. In contrast, using the cross-sectional analysis had the advantage of a big sample size but implied that there were at least different respondents at each wave.

7.7 Conclusion

While there is currently an increased use of SPDCEs in healthcare (de Bekker-Grob et al., 2012; Clark et al., 2014), it is scarce to have studies which demonstrate how the willingness to pay estimates from SPDCEs could be used in an economic evaluation (Tinelli et al., 2016). This chapter has demonstrated how the societal WTP values obtained from the SPDCE can be used to value the incremental changes or improvements in the attributes and levels resulting from an intervention.

The next chapter presents the integrated approach proposed by this thesis. First, it presents, in a disaggregated format of a cost-consequences analysis (CCA), the results of the health outcomes of the WIAT intervention from the extra-welfarist approach of CUA and the non-health outcomes of the intervention valued through of the welfarist approach of the SPDCE. These results are, then, combined to complement or add value to each other in a net monetary benefit (NMB) framework as depicted in the conceptual model of the integrated approach in chapter four. This way, the NMB framework monetizes the QALYs which allows the combination of the non-health outcomes of the intervention.

The cost associated with the delivery of the WIAT intervention is then subtracted to arrive at the NMB for both the health and non-health outcomes. The advantage of the NMB is the transparency in comparing multiple approaches and the possibility to consider multiple willingness-to-pay thresholds or values (Donaldson et al., 2011) as discussed in chapter eight.

Chapter 8: The integrated approach

8.1 Introduction

This chapter presents the integrated approach proposed by this thesis. It addresses the final objective of the thesis, concerned with the development of a broader economic evaluative space for a public health intervention.

The integrated approach combines the results of the cost -utility analysis (CUA) for the health outcomes, with the stated preference discrete choice experiment (SPDCE) results for the identified non-health outcomes of the WIAT study. This is done in a cost-benefit analysis (CBA) framework using the net benefit approach, specifically the net monetary benefit (NMB) specification.

CBA is rarely used in standard economic evaluations of healthcare due to lack of acceptability of assigning monetary values to health outcomes (McIntosh et al., 2010). This is considered as unethical and favouring only those who can afford to pay. However, the integrated approach uses the CBA through the stated preference discrete choice experiment (SPDCE) which indirectly elicit willingness to pay value without directly attempting to attach any monetary value to outcomes. Furthermore, this integrated approach uses the NMB framework in an innovative to combine the CUA and the SPDCE results on the same monetary scale. This is possible through using willingness to pay values from the SPDCE to value the incremental changes or improvements in the attribute and levels which have been considered to measure the non-health outcomes of the intervention. The total WTP value for the attributes and levels reflect the amount that individuals are willing to pay to secure the changes or improvements in those attributes or attribute-levels of woodlands which is a proxy to the value of the non-health outcomes. For the health outcomes, the willingness to pay values used is the threshold of between £20,000 and

£30,000 per unit of health gained in terms of QALYs. This range of willingness to pay for health outcomes is the societal acceptable range in the UK as recommended by the National Institute for Health and Care Excellence (NICE, 2013). The cost associated with the delivery of the WIAT study is then subtracted, to arrive at the NMB of both health and non-health outcomes.

The overall structure of this chapter is as follow: first, it discusses the concept of the net benefit framework with focus on the NMB framework. Following this, the results of the cost of resource use are presented and compared with both the CUA and the SPDCE results in a disaggregated format of a cost-consequences analysis (CCA). Then, the presentation of the proposed broader economic evaluative space for a public health intervention, termed as the integrated approach which combines the cost-utility analysis results with the SPDCE results through the NMB framework follows. A general discussion on the integrated approach is presented followed by a conclusion.

8.1.1 The net benefit framework

The net benefit framework is key to the proposed integrated approach and it is essential to recall the discussion on cost-utility analysis in chapter two. In a cost-utility analysis, the decision rules for considering the adoption of an intervention are expressed in a form of the incremental cost-effectiveness ratio (ICER) which is the ratio of the expected cost difference (incremental costs) over the ratio of the expected health outcomes difference (incremental effect) between the intervention and control groups as shown in (5) in chapter two. Then, the ICER decision rule becomes that if the ICER of an intervention $(\frac{\Delta cost}{\Delta effect})$ is less than the maximum WTP (λ), then it is worthwhile, otherwise it is not worth undertaking (Gray et al., 2010; Glick et al., 2015). This is represented as: $\frac{\Delta cost}{\Delta effect} < \lambda$, where λ is the willingness to pay (WTP).

It was also discussed in chapter two that the ICER can result in four possible outcomes which are normally depicted using a graph known as a cost-effectiveness (CE) plane. The y-axis is the incremental cost whereas the x-axis is the incremental effect. The ICER outcomes can fall into the four quadrants (North West-NW; South West-SW; North East-NE; and South East-SE) of the CE plane as shown in chapter two. If the ICER falls in the SE quadrant where costs are negative and effects are positive effects, the intervention is considered to be dominant, thus, it achieves better outcomes at lower cost. If the ICER falls in the NW quadrant where the costs are positive and the effects are negative, the intervention is said to be dominated, thus, it achieves poorer outcomes at higher cost. However, if the ICER falls in the NE quadrant with positive costs and positive effects or in the SW quadrant with negative costs and negative effects, it becomes problematic to interpret the ICER. The positive ICERs can belong to either the NE or the SW quadrant and the negative ICERs can be for the NW or SE quadrants such that the ICERs per se are not informative about the cost-effectiveness of an intervention without additional information (Briggs et al., 1997; Briggs and Fenn, 1998; Zethraeus et al., 2003; Fenwick et al., 2004; NICE, 2013). In this case, the trade-off between costs and effects needs to be examined by comparing to specific thresholds of WTP (λ) (Fenwick et al., 2006). The intervention would be considered as cost-effective if the ICER is lower than the WTP threshold $\frac{\Delta C}{\Delta E} < \lambda$ for ICERs in the NE quadrant and higher than the WTP threshold $\frac{\Delta C}{\Delta E} > \lambda$ for ICER in the SW quadrant. In this case, the four quadrants of CE plane are interpreted using a dichotomy. Thus, any intervention falling above the maximum WTP (λ) is not acceptable.

The ICERs present another problem related to the statistical analysis. When the effect size is zero which when dividing the incremental cost by the incremental effect results to infinity ($\frac{\Delta C}{0}$) (Elliott and Payne, 2005; Gray et al., 2010; Glick et al., 2015).

The solution to the above ICER problems is normally the use of the net benefit approach (Stinnett and Mullahy, 1998) which is a composite measure with cost-effectiveness and cost-benefit analysis elements in it and can be derived by rearranging the cost-effectiveness decision rule of $\frac{\Delta cost}{\Delta effect} < \lambda$ such that when the ICER of an intervention ($\frac{\Delta cost}{\Delta effect}$) is less than the maximum WTP (λ), then it is worthwhile, otherwise it is not worth undertaking (Gray et al., 2010; Glick et al., 2015). The cost-effectiveness part of the net health benefit framework becomes a linear expression after the rearrangement of the cost-effectiveness decision rule as shown below:

$$net\ health\ benefit\ (NHB) = \Delta effect\ - \ \Delta cost/\lambda \qquad (14)$$

As can be noted from the above expression, the construction of the NHB is on the health outcome scale. Similarly, the cost-benefit part after the rearrangement of the cost-effectiveness decision rule becomes a linear expression constructed on the cost scale known as the net monetary benefit (NMB), previously discussed (6) in chapter two and shown again as below:

$$net\ monetary\ benefit\ (NMB) = \Delta effect\ x\ \lambda - \ \Delta cost$$

In both expressions, $\Delta Effect$ is the incremental effect and ***Δcost*** is the incremental cost. The decision rule for both expressions then becomes that when the NHB or NMB is greater than zero, the intervention is cost-effective while if it is less than zero, then the intervention is not cost-effective (Morris et al., 2012; Edlin et al., 2015).

Of interest in this chapter is the latter cost-benefit analysis expression of the cost-effectiveness decision rule in terms of the NMB framework. This framework is preferred compared with the NHB framework partly because it allows the expression of effectiveness on the monetary scale through monetization of QALY. Furthermore, the NHB has the potential drawback of being undefined when the willingness to pay (λ) is zero ($NHB = \Delta effect\ - \ \Delta cost/0$) (Glick et al., 2015).

Taking advantage of the ability of the NMB framework to monetize the QALY using the willingness to pay threshold values and the monetary valuation of the non-health outcomes using the SPDCE, this chapter explores using this framework to combine the CUA and the SPDCE approaches on the same monetary scale in the proposed integrated approach. The next section discusses this approach in detail.

8.2 Methods

The integrated approach involved using the results of the cost-utility analysis (CUA) in chapter six and those of the SPDCE in chapter seven. The health-related outcomes of the WIAT intervention in chapter six were measured using the standard EQ-5D questionnaire and valued in terms of QALYs. The EQ-5D questionnaire cannot capture the outcomes of an intervention that go beyond health, hence it is considered inadequate for the non-health outcomes. Given that the WIAT intervention has outcomes that are non-health related, a SPDCE which was mapped to the wider WIAT study questionnaire items that were considered to be responsible for the non-health outcomes. The WTP estimates from the SPDCE which represent the value that individuals attach to the identified attributes and levels were used to value the incremental changes or improvements in the attributes and levels. This results in an estimate of individuals' WTP to secure the changes or improvements or willingness to accept compensation for being worse in terms of the changes to the attributes and levels because of the intervention.

The cost of resource use for delivery of the WIAT study discussed in chapter five is associated with all the outcomes of the WIAT study (health and non-health related). This cost of resource use together with the health and non-health outcomes resulting from the intervention are then presented in a balance sheet format through a cost-consequences analysis (CCA) approach. The CCA allows the cost of the intervention to be assessed against both the QALYs from the traditional CUA and the WTP values from the SPDCE applied to the incremental changes or improvements in the attributes and levels linked to the non-health outcomes. This way, the decision-makers are offered with an array of outcomes to choose from for varied decision contexts.

Then the disaggregated health and non-health outcomes are combined to complement or add value to each other on the same monetary scale as the cost, using the NMB framework (Stinnett and Mullahy, 1998; Drummond et al., 2015; Edlin et al., 2015). As explained earlier, this is possible through rescaling the incremental health effects between the intervention and control group into a monetary value using the cost-effectiveness WTP threshold of between £20,000 and £30,000 (NICE, 2013) as a value for each unit of the health effect gained. The same approach is applied to the non-health effects resulting from the intervention by rescaling the incremental changes or improvements in the attributes and level into monetary values using the societal WTP from the SPDCE. The incremental costs are, then, subtracted from the combined value of the monetized health gains and the monetary value of the non-health outcomes resulting in a NMB framework for both the identified health and non-health outcomes of the WIAT study. The results of the integrated approach are presented in the next section.

8.3 Results

The results for both the health and non-health outcomes of the WIAT study are firstly presented in a form of a cost-consequences analysis (CCA). A single summary ratio of incremental cost provides information on the cost differential between the intervention and control group for all the outcomes gained as a result of the WIAT study. This cost relates to the resources used for the implementation of the WIAT study. As discussed in chapter five, the control group had zero cost. An incremental effect in terms of QALYs and incremental change or improvement in the attributes and levels ratio summary represent the health and non-health gains from the WIAT intervention respectively.

To be able to combine the CUA results for the unbalanced panel analysis, and those of the SPDCE that are presented in the CCA, in a single metric, a net monetary benefit (NMB) framework was used as follows:

When λ is £20,000, the specification results in the following for the physical intervention:

$$NMB = 0.012\ x\ £20{,}000 - £7.68 = £232.32$$

While for both the physical and social intervention, it results in the following:

$$NMB = 0.024\ x\ £20{,}000 - £11.80 = £468.20$$

When λ is £30,000, the specification results in the following for the physical intervention:

$$NMB = 0.012\ x\ £30{,}000 - £7.68 = £352.32$$

While for both the physical and social intervention, it results in the following:

$$NMB = 0.024\ x\ £30{,}000 - £11.80 = £708.20$$

When the results of the CUA and the SPDCE using formula (13) for the configuration of a good woodland are combined in the integrated approach (7) presented in chapter four, the NMB specification for the unbalanced panel for the physical intervention, when WTP (λ) for the health outcomes is £20,000, becomes:

$$\begin{aligned} NMB \ = \ 0.012 \ \times \ £20{,}000 + \ 0.11 \ \ \times \ £28.32 \ + \ 4.69 \ \times -£0.64 \\ + \ 0.15 \ \ \times \ £34.45 + \ 0.14 \ \ \times \ £11.76 - £7.68 = £239.21 \end{aligned}$$

Whereas for both physical and social intervention, the results become:

$$NMB = 0.024 \times £20{,}000 + 0.11 \times £28.32 + 4.69 \times -£0.64 + 0.15 \times £34.45 + 0.14 \times £11.76 - £11.80 = £475.09$$

Whereas when λ is £30,000, the NMB results for the physical intervention and both the physical and social intervention become £359.21 and £715.09, respectively. Overall, in both cases, the NMB is greater than zero, implying that the intervention is cost-beneficial in terms of both health and non-health benefits.

The same analysis was done for the WIAT cross-sectional data. The NMB for the physical intervention and both physical and social intervention at WTP of £20,000 for the health outcomes was £288.53 and £564.41. At WTP of £30,000, the NMB became £438.53 and £854.41, respectively.

Table 8-1 below presents the results summary of the costs consequences analysis and the integrated approach of the WIAT study for the unbalanced panel analysis.

Cost-consequences analysis and the integrated approach for the WIAT study

Cost of resource use	Incremental cost	Bootstrapped	95% CI			λ(£20000) × Δeffect	λ(£30000) × Δeffect	Integrated NMB (£20000)	Integrated NMB (£30000)	
Physical intervention	£7.68	£7.68	£7.67	£7.69						
Physical & social intervention	£11.80	£11.80	£11.79	£11.82						
Consequences	Incremental QALY									
QALYs-unbalanced panel physical intervention	0.012	0.012	-0.028	0.051		£232.32	£352.32	£239.21	£359.21	
QALYs-unbalanced panel both physical & social interv	0.023	0.024	-0.049	0.094		£468.20	£708.20	£475.09	£715.09	
ICER-physical interv	£641	£627	-£5,757	£5,218						
ICER-physical & social interv	£513	£500	-£3,999	£4,098						
Non-health outcomes for a configuration of a good woodland	Incremental change/improvement in attribute/level				WTP	95% CI		λ (SPDCE) × Δeffect		
A lot of support (percentage)	0.11				£28.32	£25.63	£31.19	£3.12		
Time (minutes)	4.69				-£0.65	-£0.58	-£0.71	-£3.05		
Good quality (percentage)	0.15				£34.45	£31.53	£37.57	£5.17		
Many opportunities (percentage)	0.14				£11.76	£9.43	£14.14	£1.65		
Total								**£6.89**		

Table 8-1: Cost-consequences analysis and integrated approach for the WIAT study.

8.4 Discussion

As discussed in chapter two, the cost utility analysis approach is associated with the extra-welfarist viewpoint. The argument of extra-welfarism is that the information on which to base judgement about the results or output of a healthcare economic evaluation should be broader than individual utilities and should be based on the extent it contributes to 'health, itself as the ultimate objective (Gyrd-Hansen, 2005). On the other hand, the SPDCE is linked to the welfarist viewpoint which considers that the output of healthcare should be judged using information on utilities gained by individuals in a society and their overall welfare is the sum of these individual utilities (Birch and Gafni, 1996; Gyrd-Hansen, 2005; Brouwer, 2009; Buchanan and Wordsworth, 2015). The limitations of extra-welfarism point to its narrow focus on health as the only objective that matters in an economic evaluation while welfarism has the drawback of restricting itself to individual utilities (Coast, 2009; Morris et al., 2012; Hurley, 2014).

Given that public health interventions have multiple or varying outcomes like those of the WIAT study, it becomes problematic to establish a particular economic evaluation viewpoint of the decision context between the welfarism and extra-welfarism. For this reason, as argued by Buchanan and Wordsworth (2015), there is need to widen the evidence base by applying a theoretically sound viewpoint.

This thesis proposes the use of the integrated approach demonstrated above which combines the extra-welfarist CUA approach with the welfarist SPDCE approach in a cost-benefit analysis framework using the NMB framework. This approach reflects the conservative view of attempting to make differing two viewpoints complement or add value to each other when conducting an economic evaluation of a public health intervention, rather than trying to establish superiority or similarities between the two viewpoints. This was made possible by initially using the cost-consequences analysis (CCA) to present both the cost and the relevant outcomes, given that the multiple and varied outcomes could not be summarized using a composite measure. This way, the CUA addressed the question of achieving the maximum health outcome given the willingness to pay threshold by the decision-maker while the SPDCE revealed the value that is attached to identified non-health benefits as measured by individuals' WTP to secure the changes or improvements in the attributes and levels. Overall, the integrated approach showed that the WIAT study was value for money in a broader cost-benefit analysis framework. The NMB approach is akin to cost-benefit analysis (CBA) in that the costs and outcomes are all measured in money terms. The decision rule is the same as that of CBA. If the NMB is positive, then the intervention is preferred while the negative NMB implies that the intervention is not cost-effective hence should not be adopted.

The use of the NMB in the integrated approach is innovative as it provides a new broader conceptualization and operational approach capable of considering both health and non-health outcomes in economic evaluations of public health interventions. It maintains the use of the conventional CUA approach to value the health-related outcomes in terms of QALYs and uses the methodologically accepted SPDCE which indirectly elicits WTP values from the society. These approaches are combined on the same monetary metric to present the overall monetary value of the broad outcomes of the intervention.

While this chapter has demonstrated the feasibility of the proposed integrated approach, there are methodological challenges which may have implications on the results. Firstly, as previously discussed in chapter seven, the pragmatic approach to map the SPDCE attributes and levels to the WIAT main study questionnaire items that were considered to measure the non-health outcomes was problematic. The WIAT main study was conceptualized and designed with the SPDCE in mind, hence it was a challenge to link the SPDCE to the main study questionnaire to allow the assessment of the incremental changes or improvements in the attributes or levels resulting from the intervention. This resulted in assumptions that may somewhat be considered as not reflecting reality. As recommended in Wildman et al. (2016) recently, this limitation could be overcome if there is prior knowledge or already agree set of attributes. These attributes and levels could be incorporated in a questionnaire. This would allow the questionnaire to capture the right information linked to the SPDCE which would enable the calculation of incremental changes or improvements in attributes and levels at given time points alongside the QALY framework.

Secondly, the NMB framework in the proposed integrated approach requires the use of WTP value (λ) for the health-related outcomes to be known or estimated in order to monetize the incremental effects and bring them on the same monetary scale as costs (Edlin et al., 2015). This could be a problem, especially in cases where there is no decision on the value or where a credible WTP value does not exist (Edlin et al., 2015). In the event that the WTP values are not readily available in some contexts, Glick et al. (2015) advise to use any estimate of policy-relevant values of willingness to pay. Currently, in the UK, the willingness to pay value (λ) itself is the subject of debate and questions still remain as to how it was arrived at, whether it is an appropriate range and how the ideal WTP can be estimated (McCabe et al., 2008; Donaldson et al., 2011; NICE, 2013; Claxton et al., 2015).

The integrated approach could have benefited from the PSA to explore uncertainty around the non-health outcomes. However, the problem arose in determining the standard errors for improvements or changes in the attributes due to the pragmatic mapping approach that was adopted. Future studies could benefit from PSA if the SPDCE is developed alongside a wider study to measure the non-health outcomes. Predictive margins approach similar to the one used for HRQoL utilities in this study could be used to calculate standard errors for the PSA. This would also require determining the probability distribution for each non-health outcome. The difficulty in conducting PSA for the integrated approach in this study is a limitation.

Another limitation stems from the cost of resource use used in the proposed integrated approach. As noted in chapter five, the cost measurement approach used for the WIAT study could not capture all relevant costs. For example, the cost of community involvement in delivering the social intervention was not measured and valued because of the complexity of measuring and valuing individuals' time contribution to the intervention. Furthermore, the average cost of the WIAT intervention used in the CUA is based on the population of 20,472 in the intervention sites (FCS, 2011a; FCS, 2015). This population estimate could have implications on the cost-effectiveness of the WIAT project.

8.5 Conclusion

This chapter has presented the results of the WIAT study in a combined form of health and non-health benefits through the proposed integrated approach using the net monetary benefit (NMB) framework. The chapter has demonstrated that the integrated approach is feasible. While the NMB framework is known to solve problems associated with ICERs, this chapter has shown that it can also be used to combine the CUA approach that values the health outcomes and the SPDCE approach that captures and values the non-health outcomes of a public health intervention in a way that the two approaches complement or add value to each other. The necessary requirement for the proposed integrated approach is that the WTP for the health outcomes should be known, certain or can be calculated. This integrated approach has the advantage of allowing the exploration of the overall impact of the intervention across sectors of the economy other than health.

It is essential to note that for the successful implementation of the integrated approach, it was necessary to go into the realms of environmental and transport economics where broader economic evaluative techniques such as the SPDCE are well developed.

This integrated approach is not attempting to resolve or abate the debate between the extra-welfarism and welfarism in economic evaluations but attempts to operationalize an economic evaluation space where the two viewpoints complement or add value to each other. It is expected that this integrated approach would appeal to many researchers and could be developed further.

The next chapter concludes the thesis. It provides a general discussion of the overall thesis and how it has addressed the six research questions set out at the beginning. The chapter proceeds to discuss the strengths and weaknesses of the approaches used; and the implications for research, policy, and resource allocation decision-making. Future research work based on this thesis is also proposed.

Chapter 9: General discussion and conclusion

9.1 Summary of the thesis

This chapter provides a summary of the thesis, presents key results of the empirical analysis of the WIAT study, discusses the contribution of the thesis to the existing body of knowledge, considers the implications for future research and policy, and highlights the overarching strengths and weaknesses of the study. Suggestions for future work are also presented.

To recall, this thesis began by introducing the scope of the study, its objective, and the research questions that it aimed to address. The standard approach to economic evaluations in healthcare has maximization of health-related outcomes as the primary objective given finite healthcare budgets. The commonly used and recommended unit of outcome is the quality-adjusted life year (QALY) (NICE, 2013; Drummond et al., 2015). While this approach is well suited for interventions within healthcare (Morris et al., 2012; NICE, 2013; Drummond et al., 2015), it becomes problematic, inadequate, or unsuitable for valuing outcomes that go beyond health, particularly related to public health (Weatherly et al., 2009; Curtis, 2014; Lawson et al., 2014; Payne and Thompson, 2015). The well-known challenge is how to consider the non-health outcomes in a traditional economic evaluation framework of a QALY (Weatherly et al., 2009; McIntosh et al., 2010; Curtis, 2014; Coast et al., 2015a; Payne and Thompson, 2015).

Currently, there is no clear guidance on how to conduct an economic evaluation that incorporates both health and non-health related outcomes of an intervention (Owen et al., 2011; Payne and

Thompson, 2015). Given this lack of appropriate methodological guidance, this thesis aimed to explore the development of a broader economic evaluative space for a public health intervention with broad outcomes consisting of health and non-health. It took advantage of an existing natural experiment of the Woods In and Around Towns (WIAT) study in Scotland as a case study for empirical analysis (Silveirinha de Oliveira et al., 2013). The WIAT study has a broad array of outcomes (health and non-health related) which are examples of outcomes of some public health interventions.

The standard economic evaluation of cost-utility analysis (CUA) using the standardised EQ-5D questionnaire was used to measure and value the health-related outcomes in QALYs (Dolan, 1997; Edlin et al., 2015). Then, the stated preference discrete choice experiment (SPDCE) was used to elicit the societal willingness to pay (WTP) for the identified attributes and levels of the woodlands. These attributes and levels were mapped to the WIAT main study questionnaire items considered to measure the non-health outcomes. Following this, the WTP values from the SPDCE were applied to value the incremental changes or improvements in the attributes and levels resulting from the WIAT intervention. The results of both the CUA and the SPDCE were, then, presented in a cost-consequences analysis (CCA). Subsequently, these results were combined using the net monetary benefit framework on the same monetary scale, and in a way that they complemented and added value to each other. This was possible through the ability of the NMB framework to monetize the QALY, and effectively transforming the cost-effectiveness decision rule into a cost-benefit analysis (CBA) framework.

The above approach has been termed as “the integrated approach” in this thesis. This thesis has demonstrated that the integrated approach is feasible and provides a broader economic evaluative space for both health and non-health outcomes of a public health intervention. The integrated approach addresses some of the drawbacks that are presented by each of the economic evaluation techniques when used as stand-alone. In particular, it offers a practical solution to the challenges of conducting an economic evaluation of public health interventions as discussed in chapter four. The approach brings together methods of analysis that belong to opposing viewpoints of extra-welfarism and welfarism (the CUA and the SPDCE, respectively).

To be able to develop the integrated approach using the WIAT case study, the following research questions explored:

1. What were the costs of resources involved in the delivery of the WIAT study?
2. How were the health outcomes of the WIAT study measured and valued?
3. Was the WIAT study effective in terms of improving the health-related quality of life (HRQoL)?
4. Was the WIAT study cost-effective in terms of the health outcomes?
5. How could the non-health outcomes of the WIAT study be assessed and valued?
6. How could both the health and non-health related outcomes of the WIAT study be considered in an economic evaluation on a single metric scale?

9.2 Summary of key results

Regarding the results of the empirical analysis of the WIAT case study, both the unbalanced panel and the cross-sectional difference-in-differences (DiD) analysis showed statistically insignificant change in mean health-related quality of life for the individuals in contact with nature in the intervention group relative to the control group for both wave two, after the physical intervention and wave three, after the combined physical and social intervention.

The cost-effectiveness analysis base case results for the unbalanced panel analysis showed that the physical intervention and the combination of the physical and social intervention were value for money. The cross-sectional analysis showed similar results. There was huge uncertainty around the results of both cases.

The results of the SPDCE for the unbalanced panel analysis showed a higher total WTP value for the attributes and level changes or improvements resulting from the intervention compared with the cross-sectional analysis. This was perhaps due to differences in the samples.

Overall, the integrated approach revealed that the WIAT interventions were cost-beneficial in terms of both health and identified non-health outcomes with a positive NMB at both WTP threshold values of £20,000 and £30,000 for the health outcomes and at the WTP estimates from a general Scottish population for the non-health outcomes.

9.3 Contribution to knowledge

The key aspect of this integrated approach is the methodological design of using existing economic evaluation techniques in a novel way, particularly the net monetary benefit (NMB) framework. The NMB framework is traditionally used to resolve the incremental cost-effectiveness ratio problems as discussed in chapter two and eight (Stinnett and Mullahy, 1998; Fenwick and Byford, 2005; Fenwick et al., 2006). The ability of the NMB framework to monetize the QALY in the CUA, effectively incorporates the QALY into a CBA framework (Glick et al., 2015; Wildman et al., 2016) and provides the opportunity to enable the combination of the extra-welfarist CUA results for the health outcomes with the welfarist SPDCE results for the non-health related outcomes. This, in turn, provides a broader economic evaluative space for public health interventions capable of dealing with the broad outcomes on the same monetary scale. This approach would help fill the gap in public health economic evaluations literature in the absence of clear guidelines on how to conduct an economic evaluation of a public health intervention.

The integrated approach proposed in this thesis benefited from numerous discussions with the supervisory team of this thesis and the arguments in McIntosh (2006). McIntosh (2006) argues that developments in cost-effectiveness methodology including the net-benefit framework, could benefit cost-benefit analysis approaches, particularly SPDCEs while still maintaining their theoretical advantage of indirectly eliciting societal WTP values. It essential to note that, recently, Wildman et al. (2016) have proposed a similar approach to this integrated approach in their paper on valuing both health and social care benefits of assisted living technologies (ALTs) in an economic evaluation. However, this thesis has gone further to demonstrate how the integrated approach could be operationalized using the WIAT case study and be applicable to the broad outcomes of a public health intervention.

9.4 Implications for research

The practical implication for research of this thesis is that it has demonstrated that different economic evaluation techniques can be combined to complement and add value to each other. This has been shown in a way that does not establish superiority of one technique over another or trying to make the different techniques look similar in one way or the another as has previously been presented in economic evaluation discussions (Phelps and Mushlin, 1991; Johannesson, 1995; Donaldson, 1998b; Bleichrodt and Quiggin, 1999; Dolan and Edlin, 2002; Hansen et al., 2004; Gyrd-Hansen, 2005; Kenkel, 2006).

9.5 Implications for policy or decision-making

It is suggested that public health interventions could contribute to the well-being of the society at large at lower costs (Kelly et al., 2005; Kelly et al., 2010; Owen et al., 2011). However, this gives rise to related questions such as how much health and non-health outcomes can a public health intervention produce to justify a given cost; should a public health intervention be cost saving rather than cost-effective; and does the intervention do good rather than harm to society. Answers to these questions are important to policy or decision-making because some public health interventions are associated with an opportunity cost, implying that that the money invested in them could as well be allocated to other health activities or across other sectors to benefit the society (Brousselle and Lessard, 2011; Trueman and Anokye, 2013).

It has also been established that public health interventions require a reasonable time frame for their outcomes to manifest (Tchouaket and Brousselle, 2013; Mays and Mamaril, 2015). An important policy or decision-making question then becomes how long is enough for the benefits of a public health intervention to start manifesting themselves. As noted previously, it is also essential to recognize that some public health interventions can have inter-generational outcomes such that they need to be followed up for a long period (Park, 2014). An interesting policy or decision-making question is the feasibility of a long-term follow-up. The WIAT study economic evaluation was conducted only after two years and the question that arises is whether this economic evaluation time horizon was enough. Furthermore, another question is how to capture the maintenance costs and longevity of effect in economic evaluation, given that access to the woodlands would continue even after the end of the WIAT project.

The non-health outcomes of a public health intervention are becoming increasingly important to policy and decision-making because of the role they play in contributing to the well-being of individuals at population-level and at a relatively low cost (Kelly et al., 2010; Trueman and Anokye, 2013). Hence, an economic evaluation of public health interventions can aid resource allocation decision-making.

9.6 Strengths and weaknesses of the integrated approach

9.6.1 Strengths

The proposed integrated approach uses the NMB framework to combine the health and non-health related outcomes in a way that they complement and add value to each other on the same monetary metric. This approach has the strength which stems from the ability of the NMB framework to transform the cost-effectiveness decision rule on a monetary scale which enabled the combination of the CUA with the SPDCE.

The integrated approach has another strength of bringing together varied interests of stakeholders in terms of broad outcomes (health and non-health) of a public health intervention in an economic evaluation while at the same being able to fully appraise the multiple and varied consequences of a public health intervention. The approach further showed that different economic evaluation techniques can complement or add value to each other rather than using any of the techniques with a view of establishing superiority or equivalence of one method over another. This allows the comparison of interventions within healthcare and across other sectors to be made in resource allocation decision-making.

Another strength of this study relates to the valuation of the health-related outcomes. It maintains the use of the conventional approach of cost-utility analysis (CUA) using the EQ-5D questionnaire to measure the health outcomes, while the non-health outcomes are captured and valued through the SPDCE which appeals to researchers because it indirectly elicits willingness to pay (WTP) values from individuals. Its WTP elicitation tasks involve the notion of making a choice between alternatives or trading-off attributes and levels which implies opportunity cost (Briggs, 2016; Wildman et al., 2016).

The conduct of the SPDCE in this thesis presents another strength to the study. For example, a survey company was used which resulted in a high response rate because of the targeted approach used. In addition, the analysis of the SPDCE revealed that the identified attributes seemed to incorporate most of the attributes considered important by individuals. This was revealed from the responses of the follow-up question during the pilot and the main study survey which asked individuals to suggest attributes that they felt should have been included but left out in the SPDCE survey. There was no suggestion of any additional attributes that could be included in the SPDCE.

The SPDCE was conducted online by a survey company and one major advantage of an online SPDCE survey is that it ensures the independent treatment of each choice set presented to respondents at each click of the button. This implies that each choice set is not compared to any other set in the survey compared with the pen and paper survey, and more importantly, an online SPDCE survey is relatively quick, hence cost saving.

To the best knowledge, this is one of the rare studies that has attempted to combine the extra-welfarist approach with the welfarist approach in a manner that the two viewpoints complement and add value to each other. In addition, the proposed integrated approach provides a new conceptualization which provides a practical solution to deal with the broad outcomes of a public health intervention in an economic evaluation.

9.6.2 Weaknesses

Despite the above strengths, the study was complex and presented some challenges. It was not feasible to capture and value all the relevant costs and outcomes resulting from the WIAT study. Only some of the costs related to resource use for the delivery of the WIAT study and health outcomes that were measured by the EQ-5D questionnaire were valued in this thesis, including the identified non-health outcomes through the attributes of the SPDCE.

Costing of resource use

Turning to internal costs, the focus was on the time the Forestry Commission's staff directly spent on the WIAT study. This, however, ignored, for example, such costs related to staffs' time spent on meetings. Regarding the external costs, it was difficult to measure and value community involvement in terms of time dedicated to the delivery of the WIAT social intervention, hence, it was ignored in the costing. Furthermore, top-down approach used to measure the cost of resources used in the delivery of the WIAT study only captured high-level summaries of cost of resource use compared with the bottom-up approach which would have been more detailed. However, the top-down approach was preferred because it was easy and less costly to undertake. Caution is, therefore, required when interpreting the cost estimates because the use of top-down approach implies a trade-off between precision of the cost estimates and ease of implementation. This is a limitation although not very problematic in this thesis because the focus was on the average cost as an input in the economic evaluation.

On the estimation of the average cost of the intervention, it is essential to note that it was calculated based on the study population of 2000 individuals who live within 1km of the WIAT woodlands although any member of the community is eligible. This estimate of the population is conservative and may not be applicable in other settings.

Another weakness of the study comes from the risk of bias arising from recording of resource use in terms of time and activities carried out during the physical and social intervention. Although the recording of the cost activities was done as soon as the activities were undertaken or costs were incurred, inevitably, this retrospective recording was prone to some omissions and bias. Recording the cost activities as soon as they were undertaken helped to minimize recall bias while the probabilistic sensitivity analysis (PSA) was used to propagate uncertainty surrounding the cost estimate.

<u>Valuation of the health outcomes</u>

While there were other tools in the WIAT main study questionnaire for measuring outcomes of the wider WIAT study, the focus of this thesis was on the health outcomes as measured by the EQ-5D questionnaire. The reason was that it allowed an undertaking of an economic evaluation in terms QALYs. However, it was noted previously that two different versions of the EQ-5D questionnaire were used: the old 3L version in wave one and the new 5L version in wave two and three. This was potentially problematic in terms of analysis. The 3L and the 5L versions of EQ-5D questionnaire are two different instruments which result in different profile indices, hence their utilities might not be directly comparable. To ensure that the utilities calculated from responses of the EQ-5D 5L index profiles in wave two and three were consistent with the responses from the EQ-5D 3L index profiles in wave one or vice versa, a cross-walk mapping approach was undertaken based on the distribution similarities of the two versions of the EQ-5D. However, mapping items from one measure to another to estimate utilities for economic evaluations is known to be problematic because of the difference in content coverage of the tools involved (van Hout et al., 2012). However, this weakness was somewhat mitigated in this study because the dimensions and distributions of the 3L and 5L EQ-5D versions are similar.

Another weakness of this study stems from the use of the EQ-5D questionnaire to measure the health-related outcomes of the WIAT study. The WIAT study objectives included improving the mental well-being of individuals, therefore, questions can be asked as to whether this generic EQ-5D questionnaire was a suitable tool. The five dimensions of the EQ-5D are known to be incapable of fully capturing other aspects of health, particularly mental health (Shah et al., 2016). Despite this known weakness, the EQ-5D questionnaire was still used on the strength that it can measure the general health, including mental well-being of individuals, and allows the derivation of utilities for use in an economic evaluation. The Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS) which is well-suited, condition-specific questionnaire for mental well-being, and also the international physical activity questionnaire (IPAQ)-short form for measuring changes in physical activity, were included in the wider WIAT main study for key outcomes. However, these measurement tools were not used for economic evaluation because they do not allow the derivation of utilities which can be used for QALY calculation.

Furthermore, the study design of the WIAT could be another source of weakness, The WIAT study is a natural experiment it is known that generalizing the results of a natural experiment to other settings is problematic (Remler and Van Ryzin, 2010). For example, the results from the unbalanced panel and cross-sectional analysis in this study are different because of different samples. The effect of the intervention is only determined from a particular group that received the intervention compared with the control group. Therefore, it is uncertain if the same effect would prevail or follow the same trend in a different setting.

In addition, the main assumptions of the difference-in-differences (DiD) approach used to establish the impact of the WIAT intervention may be problematic. For example, the parallel trends and the common shocks assumptions that the approach uses, as discussed in chapter six, may be implausible in many settings (Angrist and Pischke, 2008; Mills and Patterson, 2011). The 'parallel' trends assumption may be problematic in that some unobserved confounders may have time-varying impact on the outcome (Dimick and Ryan, 2014). Second, the 'common shocks' assumption may be unrealistic because it would be difficult to find a control group which exactly matches this assumption (Dimick and Ryan, 2014).

Valuation of the non-health outcomes

The SPDCE was used to capture and value the identified non-health outcomes of the WIAT study as described in its conceptual frameworks. However, this conceptual framework does not consider all the potential range of non-health outcomes of an environmental improvement intervention to woodlands which might include improved drainage which might protect the soil, improved ecosystem services such as tourism which might result in increased local spend and improve the economy, increased orientation to nature (ten Brink et al., 2016). This can be considered as a limitation.

Another limitation relates to the use of an online survey for the SPDCE which can be problematic because only individuals who were the computer literate and had access to computers or any computing technology were recruited. However, it was generally considered that the benefits offered by an online survey, as mentioned earlier in this chapter, might outweigh the concerns of sampling an online survey. Furthermore, the SPDCE presented another problem related to scope as it was administered to only respondents from Scotland. This could affect the generalizability and transferability of the willingness to pay estimates because preferences may differ according to differences in a number of factors including setting (Hiligsmann et al., 2014).

Another weakness relates to the SPDCE study design. The SPDCE developed after the WIAT study design, which prompted the mapping of the WIAT main study questionnaire items considered to measure the non-health outcomes to the attributes and levels of the SPDCE retrospectively. This was problematic because of the pragmatic approach undertaken. Some assumptions which were made might not reflect how individuals use the woodlands. These assumptions were discussed in chapter seven. Another question that can be asked is about whether the identified questionnaire items are likely to be affected by the intervention. For example, distance and social interactions, in the case of the WIAT study.

The unbalanced and cross-sectional panel nature of the WIAT data presented another problem. The aim of undertaking the mapping was to measure the incremental changes or improvements in the attributes and levels related to the non-health outcomes resulting from the intervention. However, while the unbalanced panel analysis followed up the same respondents in at least two waves from the baseline (wave one), the sample size was small. The cross-sectional analysis had a relatively big sample size but did not follow-up the same respondents from wave one to three to be able to capture the actual changes or improvements in the attributes and levels resulting from the intervention. Furthermore, the sample size calculation for the WIAT study was powered for the perceived stress scale measure of mental well-being and questions can be asked if at all the sample size used was powered enough to detect meaningful changes or improvements in the attributes and levels caused by the intervention, the way it has been used in this integrated approach.

Furthermore, the economic evaluation of the WIAT study was only undertaken after two years while potentially, access to the woodlands would continue to be available after the end of the study, which implies longevity of effect. The benefits of the WIAT intervention could even be inter-generational (Park, 2014). However, it is known that evaluations related to green spaces fall short of measuring the effect over long periods of time (PHE, 2014). This can be considered as an issue that could impact on the valuations of the outcomes. At the same time, it is acknowledged that, practically, it is very difficult to have longer time lags before an economic evaluation of an intervention is conducted or before follow-ups can be made because of limited budgets and time. Sometimes it is not even feasible to have long periods to conduct follow-up studies. Perhaps a good approach would be to have a design that gathers information on long term improvements of health with an outcome measure that can directly attribute the improvements to the intervention such as reduction in number of visits to the general practitioner, for example.

9.7 Future work

Further to the integrated approach, future studies should consider the scope effects of combining different WTP values elicited from different methods as has been done in this study. Furthermore, it would be interesting to design the SPDCE alongside the wider study survey or to have prior knowledge or already agreed on attributes and levels. That way, the attributes and levels could be incorporated in the wider study questionnaire, thereby avoiding the problems that come with mapping. In addition, more research is required to explore the application of other suggested approaches on how to deal with the challenges of conducting an economic evaluation of a public health intervention. A systematic review would reveal all the approaches being suggested in economic evaluation literature but rarely been used in practice. For example, future work in economic evaluation of public health interventions would benefit from: firstly, using the social objective framework discussed in chapter four, which combines the standard economic evaluation with the social objectives that address health inequalities along with the objective of maximizing health; using a trade-off approach between maximizing health and equity as suggested by Cookson et al. (2009); using the capability approach that moves away from health or utility maximization to evaluating an intervention based on individual's ability to function; using a cost-utility analysis (CUA) with an expanded QALY incorporating an additional non-health related dimension; using a CUA with a multi-sectoral approach where the costs and outcomes of an intervention are captured simultaneously between sectors and adjusted for a single available budget or resource across all sector of the economy; and lastly, using a subjective well-being measure.

In addition to the above, further work regarding SPDCE is needed to deal with hypothetical bias. It is acknowledged that SPDCE are liable to hypothetical bias as they involve a series of hypothetical choice situations with finite alternatives which consist of attributes and levels that are varied. Respondents are then asked to choose the most preferred alternative. The extent of this hypothetical bias is, however, unclear because its assessment requires conducting studies that compare the SPDCE with the observed or revealed preference discrete choice experiment. These types of studies are practically difficult or costly. Future research on hypothetical bias in SPDCE should, therefore, concentrate on ways of mitigating it such as loading the SPDCE with cheap talk and certainty scales which are a communication aimed to obtain credible responses as extras as suggested by Fifer et al. (2014).

Obviously, there will be concerns about the cognitive burden to respondents for increasing the tasks which adds to complexity of the SPDCE. However, recently, it has been found that task complexity in a SPDCE could improve the learning curve of the respondents resulting in response certainty and improved statistical precision of the SPDCE model (more is better than less) (Regier et al., 2014).

9.8 Conclusion

This chapter has presented a general overview of the thesis. It has revisited the research questions that the thesis aimed to address with the aim of developing the proposed integrated approach. The chapter proceeded to present a summary of key results from the empirical analysis of the WIAT case study. The contribution that this thesis makes to existing knowledge has also been presented. Then, implications for research, policy or decision-making were also discussed. This was followed by a discussion on the strenghts and weaknesses the study. Lastly, the findings of this research have led to some suggestions on the areas for future work. This chapter has outlined the areas that would be interesting to research and develop on in future.

In general, the thesis has demonstrated the use of familiar economic evaluation techniques in a novel way, and the untraditional combination of varying economic evaluation viewpoints of welfarism and extra-welfarism. It has further shown that conducting economic evaluations of interventions with broad outcomes consisting of health and non-health is challenging but not impossible. The proposed integrated approach draws upon the strengths of exsiting methods of analysis and techniques such as CUA for the health outcomes and SPDCE for the non-health outcomes, then combining them using a well-known net monetary benefit (NMB) framework as a conversion tool to bring the broad outcomes on a single scale of money. Hence, a broader economic evaluative space.

While the integrated approach has been developed using a case study of the WIAT, it is argued to be particularly suitable for the economic evaluation of a public health intervention.

Appendices

Appendix 1: The WIAT main study questionnaire

WIAT/NIHR Project: woodlands and wellbeing

Main Study

- **CAPI Number** (automatic)
- **Reference Number** (manually entered from sample sheet)
- **Date** (automatic)
- **Time beginning** (automatic)
- **Time end** (automatic)
- **INTERVIEWER NUMBER** (manually entered)
- **LOCATION** [INTERVIEWER TO SELECT FROM LIST]:
  - Dalkeith
  - Glasgow (Pollok)
  - Glenrothes
  - Johnstone
  - Motherwell
  - Paisley (Linwood)
- **DISTANCE BAND** [INTERVIEWER TO SELECT FROM LIST]:
  - 150
  - 300
  - 500
  - 750
  - 1500
- **SAMPLE TYPE**:
  - Main (automatically codes as Main until Spare sample is approved & released by client)
  - Spare
- **Confirm address**:
  - Yes (continue)
  - No (close)

INTRODUCTION:
Good morning/afternoon I am ……….. and I would like to speak to…. *OR……* to participate in a survey. Your household has been chosen at random from among your area's postal addresses [ONLY FOR NEW HOUSEHOLDS]. The survey forms part of a project being undertaken by the University of Edinburgh and its partners to find out what you think about your local environment and your wellbeing.

Please be assured that the information you provide will be treated as entirely confidential and it will not be possible to identify any individual in any published use of the research. The survey is being administered by market research agency Progressive Partnership, the University of Edinburgh and its partners who all abide by the rules and guidelines of the Market Research Society. Named contact: The interview should take no more than 25 minutes. Thank you very much for your time. If respondent says they have done this before explain: As *{respondent name}* participated in the survey last year, we would like to ask them to participate again this year in order to measure any change in opinions and attitudes over time. New household (not participated before): The person who we would like to answer the questions is the adult member (aged 16 or over) of your household who has the next birthday. It is important that the right person answers the questions to ensure that we get an accurate picture of your views. The interview should take no more than 25 minutes. Thank you very much for your time.

- **OUTCOMES** [SINGLE CODE ONLY]:
    - Effective – go to A1a
    - Refusal – go to A1a
    - No reply
    - Named respondent not in at the moment – please call back/rearrange suitable time
    - No contact with selected person
    - Away/in hospital during survey period
    - Selected person senile/incapacitated
    - Inadequate English (not possible to use interpreter)
    - No contact made with a responsible adult
    - Office refusal (telephone/letter)
    - Not traced
    - Derelict/demolished
    - Empty/vacant
    - Business/industrial only (not private)
    - Other (specify) ___________

PART A

A1a. Named Contact:

Firstly, to make sure I am interviewing the correct person, can you confirm that you are *{respondent name, gender, age}*

SINGLE CODE	CODE	ROUTE
Yes	1	Go to A1b
No	2	Ask for correct named contact or go to A1c

If respondent is a Refusal at OUTCOMES, then thank and close

If respondent is an Effective at OUTCOMES, then continue

A1b. Named Contact:

To make sure I am interviewing at the correct address, can you confirm that you live in the current address *{full address and postcode}*

SINGLE CODE	CODE	ROUTE
Yes	1	Go to A2
No	2	Thank you and close

A1c. Named contact:

Reason why it is not the same respondent:

- Named contact not in at second attempt
- Named respondent not available during fieldwork period
- Named respondent no longer living at this address
- Named respondent not known at this address
- Someone else refused participation on the named respondents behalf
- Other reason, please specify ___________

If Named Contact is not around and another adult within the household is happy to take part, please generate new code for the new respondent.

A1d. Named contact:

Do you remember taking part in this survey previously?

- Yes, in 2013 (2 years ago)
- Yes, in 2014 (last year)
- Yes, don't remember the years
- No
- Don't know/remember

A1e. New respondent:

For a new respondent, please establish their relationship to the Named Contact and code below:

- Spouse / partner
- Child
- Parent
- Sibling
- Other family member
- Other, please specify ___________
- Not applicable – named respondent no longer living at the address

A2. Respondent's gender [DO NOT READ OUT, INTERVIEWER TO RECORD]

	CODE
Male	1
Female	2

PART B
Local Woodlands

READ OUT:
We want to ask you about woodlands, by which we mean forests and woodlands with small or large areas of trees, under any ownership, both old and new, and of any type.
ASK ALL
B1. Can you name any woodlands around here?

SINGLE CODE	CODE	ROUTE
Yes	1	Go to Q B1.1
No	2	Go to Q B3

B1.1. If yes, please specify:

B1a: ________________________________

B1b: ________________________________

B1c: ________________________________

B1d: ________________________________

B1e: ________________________________

B2. Which of these have you visited in the <u>last 12 months</u>?

	Yes	No	Can't remember
B1a:			
B1b:			
B1c:			
B1d:			
B1e:			

ASK ALL
Now, thinking <u>about these woodlands</u> *{INSERT NAME}*
Instruction: SHOW MAP
B3. Have you visited <u>these local woodlands</u> in the <u>last year</u>?
SHOW CARD A

SINGLE CODE	CODE	ROUTE
Yes	1	Go to B4
No	2	Go to B10

B4. What kinds of activities do you pursue when visiting these local woodlands?
SPONTANEOUS

CODE ALL THAT APPLY	Code
Go for a walk	1
Walk the dog	1
Go out with my family	1
Exercise or sport	1
Relax	1
Look at plants or wildlife	1
Participate in an event	1
Other (Please specify)________________________________	1

B5. How frequently did you visit these local woodlands <u>last winter</u> i.e. between October and March?
SHOW CARD B

SINGLE CODE	Code
Every day	1
Several times a week	2
Once a week	3
Several times a month	4
About once a month	5
Less often	6
Not at all	7
Do not know	-98

B6. How frequently did you visit these local woodlands <u>last Summer</u> i.e. between April and September?
SHOW CARD B

SINGLE CODE	Code
Every day	1
Several times a week	2
Once a week	3
Several times a month	4
About once a month	5
Less often	6
Not at all	7
Do not know	-98

B7. On average during the last 12 months how long, did you normally spend at these local woodlands?
SHOW CARD C

SINGLE CODE	Code
Up to 15 minutes	1
Over 15minutes – 30 minutes	2
Over 30 minutes – 1 hour	3
Over 1 hour – 2 hours	4
Over 2 hours – 5 hours	5
More than 5 hours	6
Do not know	-98

B8. How do you usually get to these local woodlands?

READ OUT, SINGLE CODE	Code
By foot	1
By car	2
By bicycle	3
By public transport	4
By taxi	5
Other (Please specify)____	6

B9. With whom do you usually go to these woodlands?

READ OUT, CODE ALL THAT APPLY	**Code**
Alone	1
Alone with the dog	1
With others, including family and friends	1

NOTE: if respondents chose more than one code at B9, please ask:
B9. 1. With whom do you usually go to these woodlands most frequently?
Interviewer Note: If respondent always walks their dog with someone else please code as 'With others'

READ OUT, SINGLE CODE	**Code**
Alone	1
Alone with the dog	2
With others, including family and friends	3

ASK ALL
B10. How easy is it to get to these local woodlands from where you live?
SHOW CARD D
SINGLE CODE

Very easy	Easy	Neither easy nor difficult	Difficult	Very Difficult	Do not know
1	2	3	4	5	-98

B11. How long would it take you to walk to these local woodlands?

READ OUT, SINGLE CODE	**Code**
_______________ minutes	
Cannot walk (If respondents cannot walk go to B11.1)	0

B11.1. If you cannot walk how long would it take you to get there by other means?
_______________ minutes → by what means? ___________
Unsure
Thinking about <u>these local</u> woodlands and what they are like, tell us what you think of the following:
Please, score each statement according to your level of agreement.
SHOW CARD E
Interviewer Note: if respondents really do not know, please code as 'neutral'
PLEASE PROBE THOROUGHLY

	TICK START, ROTATE, READ OUT	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
B12	The woodlands are free from litter	1	2	3	4	5
B13	Poor entrances make it difficult to get into the woodlands	1	2	3	4	5
B14	I feel safe in the woodlands	1	2	3	4	5
B15	Poorly maintained paths make it difficult to visit the woodlands	1	2	3	4	5

B16	I feel at peace in the woodlands	1	2	3	4	5
B17	I can pursue healthy activities in the woodlands	1	2	3	4	5
B18	The woodlands provide a place to visit with family and friends	1	2	3	4	5
B19	I can see and enjoy wildlife in the woodlands	1	2	3	4	5
B20	I like the natural appearance of the woodlands	1	2	3	4	5
B21	There is a lack of good facilities in the woodlands	1	2	3	4	5

READ OUT: **We are interested in how you experience [these woodlands *{name here, if known}*. To help us understand your experience, we have provided the following statements for you to respond to.**
Please read/listen to each statement carefully, and then ask yourself:
"How much does this statement apply to my experience of the woodlands?"
To indicate your answer, choose one of the numbers on the scale beside it. A sample of the scale with verbal descriptions for the values is given below. SHOW CARD F

Not at all — Completely

0---- ----1---- ----2---- ----3---- ----4---- ----5---- ----6---- ----7---- ----8---- ----9---- ----10

Interviewer Note: Please push as much as possible for a response but if respondent really cannot give an answer please code as 0

B22	Spending time in the woodlands gives me a break from my day-to-day routine	0- -1- -2- -3- -4- -5- -6- -7- -8- -9- -10
B23	There is much to explore and discover in the woodlands.	0- -1- -2- -3- -4- -5- -6- -7- -8- -9- -10
B24	My attention is drawn to many interesting things when I am in the woodlands	0- -1- -2- -3- -4- -5- -6- -7- -8- -9- -10
B25	The woodlands is a place to get away from the things that usually demand my attention	0- -1- -2- -3- -4- -5- -6- -7- -8- -9- -10

B26. Did you visit any local woodlands near where you lived as a child?
SHOW CARD G

SINGLE CODE	**Code**
Almost every day	1
More than once a week	2

Once a week	3
Several times a month	4
Once a month	5
Several times a year	6
Once a year	7
Less than once a year	8
Never	9

B27. Have you been consulted about your views on local woodlands in the last 12 months?

READ OUT, SINGLE CODE	**Code**
Yes	1
No	2

B28. Recently, have you been involved in any of the following community woodland activities?
SHOW CARD H

CODE ALL THAT APPLY	**Code**
Led walks in woodlands	1
Community events in woodlands	1
Educational activities in woodlands	1
Conservation or woodland management work	1
Other (Please specify):____	1
I have not been involved	1

B29. Overall, what do you think about the quality of these local woodlands? SHOW CARD I

Very good	Good	Neutral	Poor	Very poor	Do not know what my local woodlands are like
1	2	3	4	5	-98

B30. How important are these local woodlands around here in making a difference to your quality of life? SHOW CARD J

Very important	Somewhat important	Neutral	Unimportant	Irrelevant	Do not know
1	2	3	4	5	-98

B31. Are you aware of any changes in these particular woodlands?

READ OUT, SINGLE CODE	**Code**	**Route**
Yes	1	Go to Q B31.1
No	2	Go to B32

B31.1 If yes, how would you rate these changes?

Very negative	Poor	Neutral	Good	Very positive
1	2	3	4	5

B31.2 If yes, how did you become aware of these changes?

READ OUT, SINGLE CODE	**Code**

I heard about the changes from others	1
I read about the changes	2
I saw the changes myself	3

B32. Compared to a year ago, do you think you use these particular woodlands...

READ OUT, SINGLE CODE	**Code**
More	1
Less	2
About the same	3

B33. Have you taken part in an organised activity in the woodlands in the last year?

READ OUT, SINGLE CODE	**Code**	**Route**
Yes	1	B34
No	2	C1

B34. If yes, whom with?

READ OUT, SINGLE CODE	**Code**
Alone	1
Alone with the dog	2
With others, including family and friends	3

B35. If yes, when did you take part in the activity?

READ OUT, SINGLE CODE	**Code**
Summer i.e. between April and September 2014?	1
Winter i.e. between October 2014 and March 2015?	2

PART C

Views

C1. Do you have direct views of the local woodland?

INTERVIEWER NOTE: Please remind respondents that these are the woodlands on the map, if required.

READ OUT, SINGLE CODE	**Code**	**Route**
Yes, good view	1	Go to Q C2
Yes, a partial view	2	Go to Q C2
No	3	Go to Q C3

C2. What do you like, if anything, about these views? SHOW CARD K

CODE ALL THAT APPLY	**Code**
It is interesting (people to watch, seasonal change)	1
It is relaxing and takes my mind off things	1
It is just pleasant to look at	1
Some other reason (please say what) ________________________________	1
There is nothing I like about it	1

C3. When you are walking about your neighbourhood, are you aware of any views to woodlands or green spaces?

READ OUT, SINGLE CODE	**Code**	**Route**

Yes	1	Go to Q C4
Yes, a partial view	2	Go to Q C4
No	3	Go to Q D1

C4. What do you like, if anything, about these views? SHOW CARD K

CODE ALL THAT APPLY	**Code**
It is interesting (people to watch, seasonal change)	1
It is relaxing and takes my mind off things	1
It is just pleasant to look at	1
Some other reason (please say what) ______________________________	1
There is nothing I like about it	1

PART D

Other green spaces

READ OUT: Now thinking about parks or green spaces, other than your local woodlands.
Instruction: The respondent's definition of 'local' is being sought. If the respondent asks what is 'local' please say "10-15mins walk from home".

D1. Have you visited local parks or green spaces in the <u>last 12 months</u>?
SHOW CARD A

SINGLE CODE	**Code**
Yes	1 – Go to Q D2
No	2 – Go to Q E1

D2. What kinds of activities do you pursue when visiting local parks or green spaces?
SPONTANEOUS

CODE ALL THAT APPLY	**Code**
Go for a walk	1
Walk the dog	1
Go out with my family	1
Exercise or sport	1
Relax	1
Look at plants or wildlife	1
Participate in an event	1
Other (specify)____	1

D3. How frequently did you visit local parks or green spaces <u>last winter</u> i.e. between October and March? SHOW CARD B

SINGLE CODE	**Code**
Every day	1
Several times a week	2
Once a week	3
Several times a month	4
About once a month	5
Less often	6
Not at all	7
Do not know	-98

D4. How frequently did you visit local parks or green spaces <u>last Summer</u> i.e. between April and September? SHOW CARD B

SINGLE CODE	**Code**
Every day	1
Several times a week	2
Once a week	3
Several times a month	4
About once a month	5
Less often	6
Not at all	7
Do not know	-98

D5. With whom do you usually go to local parks or green spaces?

READ OUT, CODE ALL THAT APPLY	**Code**
Alone	1
Alone with the dog	1
With others, including family and friends	1

NOTE: if respondents chose more than one code at D5, please ask:

D5.1. With whom do you usually go to local parks or green spaces most frequently?

Interviewer Note: If respondent always walks their dog with someone else please code as 'With others'

READ OUT, SINGLE CODE	**Code**
Alone	1
Alone with the dog	2
With others, including family and friends	3

D6. How do you usually get to local parks or green spaces?

READ OUT, SINGLE CODE	**Code**
By foot	1
By car	2
By bicycle	3
By public transport	4
By taxi	5
Other (Please specify) ________________________________	6

PART E

Neighbourhood

E1. How satisfied are you with your quality of life in this neighbourhood?
SHOW CARD L

Very satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very dissatisfied
1	2	3	4	5

E2. Would you advise a friend to live in this neighbourhood?

SHOW CARD M

Completely	Would consider	Neither would nor would not	Unlikely to consider	Not at all
1	2	3	4	5

E3. How satisfied are you with the quality of the physical environment in this neighbourhood?

SHOW CARD L

Very satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very dissatisfied
1	2	3	4	5

PART F

Your Health and Physical Activity

The next few questions are about how you have been feeling day to day recently, for example, if you feel happy, a bit stressed or are finding things difficult. The reason for these questions is to help the researchers understand if your local environment helps you to feel more or less positive. All of this information is completely confidential so please be as honest as you can.

INTERVIEWER NOTE: PLEASE PROBE AS THOROUGHLY AS YOU CAN FOR ALL FOLLOWING QUESTIONS

F1. Within the last 12 months, has anything happened to you (or your family) which has had an impact on how you feel about day-to-day life (better or worse). It might be a positive or negative life event, for or example, loss of a job, personal illness, arrival of a new baby, or a marriage.

How has this event (s) made you feel?

READ OUT, SINGLE CODE	**Code**
Better than normal	1
Much worse than normal	2
No different than normal	3
Nothing has happened in last 12 months which has impacted my life	4

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way.

SHOW CARD N

In the last month...		Never	Almost Never	Sometimes	Fairly Often	Very Often
F2	**How often have you been upset because of something that happened unexpectedly?**	0	1	2	3	4
F3	**How often have you felt that you were unable to control the important things in your life?**	0	1	2	3	4

F4	**How often have you felt nervous and "stressed"?**	0	1	2	3	4
F5*	**How often have you felt confident about your ability to handle your personal problems?**	0	1	2	3	4
F6*	**How often have you felt that things were going your way?**	0	1	2	3	4
F7	**How often have you found that you could not cope with all the things that you had to do?**	0	1	2	3	4
F8*	**How often have you been able to control irritations in your life?**	0	1	2	3	4
F9*	**How often have you felt that you were on top of things?**	0	1	2	3	4
F10	**How often have you been angered because of things that were outside of your control?**	0	1	2	3	4
F11	**How often have you felt difficulties were piling up so high that you could not overcome them?**	0	1	2	3	4

* PSS scores are obtained by reversing responses (e.g., 0 = 4, 1 = 3, 2 = 2, 3 = 1 & 4 = 0) to the four positively stated items (items F5, F6, F8, & F9) and then summing across all scale items.

F. READ OUT:

We are trying to find out what you think about your health. I will first ask you some simple questions about your health TODAY. I will then ask you to rate your health on a measuring scale. I will explain what to do as I go along, but please interrupt me if you do not understand something or if things are not clear to you. Please also remember that there are no right or wrong answers. We are interested here only in your personal view

EQ-5D-5L DESCRIPTIVE SYSTEM - PAGE 2: INTRODUCTION

First I am going to read out some questions. Each question has a choice of five answers.

Please tell me which answer best describes your health TODAY.

Do not choose more than one answer in each group of questions

(Note to interviewer: it may be necessary to remind the respondent regularly that the timeframe is TODAY. It may also be necessary to repeat the questions verbatim)

F12. MOBILITY **READ OUT: First I'd like to ask you about mobility. Would you say that you have:**		Code
No problems in walking about?	□	1
Slight problems in walking about?	□	2
Moderate problems in walking about?	□	3
Severe problems in walking about?	□	4
You are unable to walk about?	□	5

F13. SELF-CARE **READ OUT**: **Next I'd like to ask you about self-care. Would you say that you have:**		Code
No problems washing or dressing yourself?	☐	1
Slight problems washing or dressing yourself?	☐	2
Moderate problems washing or dressing yourself ?	☐	3
Severe problems washing or dressing yourself?	☐	4
You are unable to wash or dress yourself?	☐	5

F14. USUAL ACTIVITIES **READ OUT: Next I'd like to ask you about usual activities, for example work, study, housework, family or leisure activities. Would you say that you have:**		Code
No problems doing your usual activities?	☐	1
Slight problems doing your usual activities?	☐	2
Moderate problems doing your usual activities?	☐	3
Severe problems doing your usual activities?		4
You are unable to do your usual activities?		5

F15. PAIN / DISCOMFORT **READ OUT: Next I'd like to ask you about pain or discomfort. Would you say that you have:**		Code
No pain or discomfort?	☐	1
Slight pain or discomfort?	☐	2
Moderate pain or discomfort?	☐	3
Severe pain or discomfort?		4
Extreme pain or discomfort?		5

F16. ANXIETY / DEPRESSION **READ OUT: Finally, I'd like to ask you about anxiety or depression. Would you say that you are:**		Code
Not anxious or depressed?	☐	1
Slightly anxious or depressed?	☐	2
Moderately anxious or depressed?	☐	3
Severely anxious or depressed?		4
Extremely anxious or depressed?		5

Interviewer Note: If any respondent cannot or does not want to answer F12 – F16, please code as 'Leave blank' in the screen following the question.

F17. EQ VAS - PAGE 4: INTRODUCTION

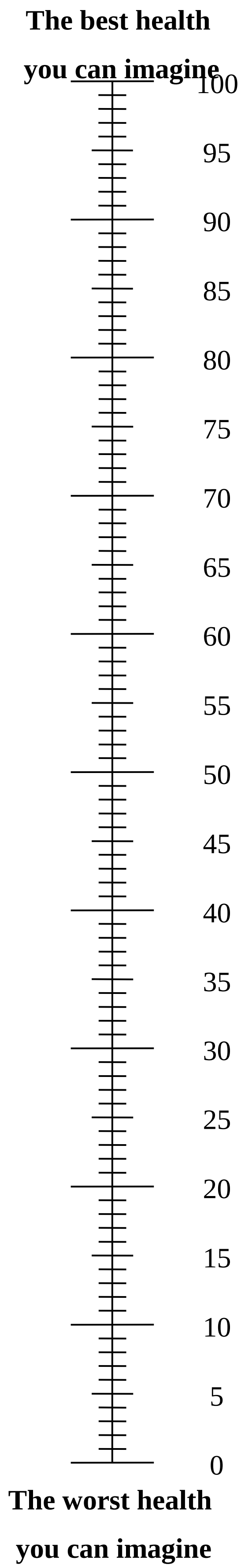
READ OUT: I would now like to ask you to do a rather different task.

To help you say how good or bad your health is, I'd like you to look at the scale, which is similar to a thermometer.

The best health you can imagine is marked 100 (one hundred) at the top of the scale and the worst health you can imagine is marked 0 (zero) at the bottom.

EQ VAS - PAGE 4: TASK
READ OUT: I would now like you to tell me the point on this scale where you would put your health TODAY.

Thank you for taking the time to answer these questions.



F18. How many times have you visited your G.P. during the last month?
NOTE: What is sought with this question is what NORMALLY HAPPENS

READ OUT, SINGLE CODE	**Code**
18.1 GP comes to me (Home visits)	______
18.2. GP visits to the practice	______

F19. Are your day-to-day activities limited because of a health problem or disability which has lasted, or is expected to last, at least 12 months?

READ OUT, SINGLE CODE	**Code**
Yes, limited a lot	1
Yes, limited a little	2
No, not limited at all	3

F20. Do you smoke tobacco at the moment (e.g. cigarettes, pipes, cigars and your own roll-ups)?

READ OUT, SINGLE CODE	**Code**
Never smoked	1
Smoked in the past	2
Currently smoke	3
DO NOT READ OUT Prefer not to say	-99

READ OUT: I am going to ask you about the time you spent being physically active in the <u>last 7 days</u>. Please answer each question even if you do not consider yourself to be an active person. Think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.
READ OUT: Now, think about all the *vigorous* activities which take *hard physical effort* that you did in the last 7 days. Vigorous activities make you breathe much harder than normal and may include heavy lifting, digging, aerobics, or fast bicycling. Think only about those physical activities that you did for at least 10 minutes at a time.
F21. During the <u>last 7 days</u>, on how many days did you do <u>vigorous</u> physical activities?
[Interviewer clarification: Think only about those physical activities that you do for at least 10 minutes at a time**]**
[Interviewer note: If respondent answers zero, refuses or does not know, skip to Question F23]

SINGLE CODE	**Code**	**Route**
______days per week (if 0 go to Q F23)		Go to Q F22
DO NOT READ OUT Don't know/ not sure	-98	Go to Q F23
DO NOT READ OUT Refused	-99	Go to Q F23

F22. How much time did you usually spend doing <u>vigorous</u> physical activities on one of those days?

	Code
__ __ hours per day	
__ __ __ minutes per day	
DO NOT READ OUT Don't know/Not sure	-98

DO NOT READ OUT Refused	-99

[**Interviewer probe**: An average time for one of the days on which you do vigorous activity is being sought. If the respondent can't answer because the pattern of time spent varies widely from day to day, ask: "How much time in total would you spend **over the last 7 days** doing vigorous physical activities?”]

F22.P	Code
__ __ __ hours per week	
__ __ __ minutes per week	
DO NOT READ OUT Don’t know/Not sure	-98
DO NOT READ OUT Refused	-99

READ OUT: Now think about activities which take *moderate physical effort* that you did in the last 7 days. Moderate physical activities make you breathe somewhat harder than normal and may include carrying light loads, bicycling at a regular pace, or doubles tennis. Do not include walking. Again, think about only those physical activities that you did for at least 10 minutes at a time.

F23. During the last 7 days, on how many days did you do moderate physical activities?

Interviewer clarification: Think only about those physical activities that you do for at least 10 minutes at a time

[**Interviewer Note**: *If respondent answers zero,* refuses or does not know, skip to Question F25]

SINGLE CODE	Code	Route
______days per week (if 0 go to Q F25)		Go to Q F24
DO NOT READ OUT Don’t know/ not sure	-98	Go to Q F25
DO NOT READ OUT Refused	-99	Go to Q F25

F24. How much time did you usually spend doing moderate physical activities on one of those days?

[**Interviewer probe**: An average time for one of the days on which you do vigorous activity is being sought]

	Code
__ __ hours per day	
__ __ __ minutes per day	
DO NOT READ OUT Don’t know/Not sure	-98
DO NOT READ OUT Refused	-99

[If the respondent can't answer because the pattern of time spent varies widely from day to day, ask: “What is the total amount of time you spent over the **last 7 days** doing moderate physical activities?]

F24.P	Code
__ __ __ hours per week	
__ __ ____minutes per week	
DO NOT READ OUT Don't know/Not sure	-98
DO NOT READ OUT Refused	-99

READ OUT: Now think about the time you spent walking in the last 7 days. This includes at work and at home, walking to travel from place to place, and any other walking that you might do solely for recreation, sport, exercise, or leisure.

F25. During the last 7 days, on how many days did you walk for at least 10 minutes at a time?

[**Interviewer clarification**: Think only about the walking that you do for at least 10 minutes at a time]

[**Interviewer Note**: *If respondent answers zero,* refuses or does not know, skip to Question F27]

SINGLE CODE	Code	Route
______days per week (if 0 go to Q F27)		Go to Q F26
DO NOT READ OUT Don't know/ not sure	-98	Go to Q F27
DO NOT READ OUT Refused	-99	Go to Q F27

F26. How much time did you usually spend walking on one of those days?

	Code
__ __ hours per day	
__ __ __ minutes per day	
DO NOT READ OUT Don't know/Not sure	-98
DO NOT READ OUT Refused	-99

[**Interviewer probe**: An average time for one of the days on which you walk is being sought. If the respondent can't answer because the pattern of time spent varies widely from day to day, ask: "What is the total amount of time you spent walking over **the last 7 days**?"]

F26.P	Code
__ __ __ hours per week	
__ __ ____ minutes per week	
DO NOT READ OUT Don't know/Not sure	-98
DO NOT READ OUT Refused	-99

READ OUT: Now think about the time you spent sitting on week days during the last 7 days. Include time spent at work, at home, while doing course work, and during leisure

time. This may include time spent sitting at a desk, visiting friends, reading or sitting or lying down to watch television.

F27. During the last 7 days, how much time did you spend sitting on a week day?

[**Interviewer clarification**: Include time spent lying down (awake) as well as sitting]

	Code
__ __ hours per day	
__ __ __ minutes per day	
DO NOT READ OUT Don't know/Not sure	-98
DO NOT READ OUT Refused	-99

[**Interviewer probe**: An average time per day spent sitting is being sought. If the respondent can't answer because the pattern of time spent varies widely from day to day, ask: "What is the total amount of time you spent *sitting* last **Wednesday**?"

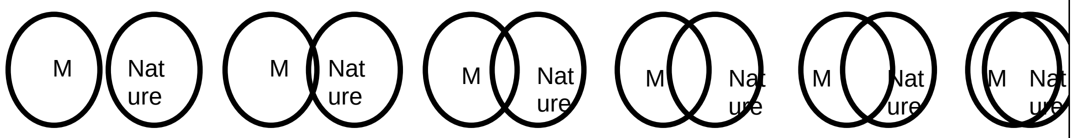
F27.P	Code
__ __ hours on Wednesday	
__ __ __ minutes on Wednesday	
DO NOT READ OUT Don't know/Not sure	-98
DO NOT READ OUT Refused	-99

Please tick the box that best describes your experience of each over the past 2 weeks.
SHOW CARD O

		None of the time	Rarely	Some of the time	Often	All of the time
F28	**I've been feeling optimistic about the future**	1	2	3	4	5
F29	**I've been feeling useful**	1	2	3	4	5
F30	**I've been feeling relaxed**	1	2	3	4	5
F31	**I've been dealing with problems well**	1	2	3	4	5
F32	**I've been thinking clearly**	1	2	3	4	5
F33	**I've been feeling close to other people**	1	2	3	4	5
F34	**I've been able to make up my own mind about things.**	1	2	3	4	5

PART G
Connectedness to Nature Scale (CNS)

G1. Please circle the picture below which best describes your relationship with the natural environment. How interconnected are you with nature?



CODE: 1	2	3	4	5	6	7

PART H
Social Cohesion / Social Capital

Please say how much you agree or disagree with the following statements:
H1. To what extent do you agree or disagree that people in this neighbourhood pull together to improve the neighbourhood?
SHOW CARD P

Definitely agree	Tend to agree	Tend to disagree	Definitely disagree	SPONTANEOUS ONLY: Nothing needs improving	Do not know
1	2	3	4	5	-98

H2. Would you say that …?
SHOW CARD Q

Many of the people in your neighbourhood can be trusted	**Some** can be trusted	A **few** can be trusted	**None** of the people in your neighbourhood can be trusted	SPONTANEOUS ONLY: Just moved here
1	2	3	4	5

H3. How strongly do you feel you belong to your immediate neighbourhood?
SHOW CARD R

Very strongly	Fairly strongly	Not very strongly	Not at all strongly	Do not know
1	2	3	4	-98

H4. I'd like you to think about any groups, clubs or organisations that you've been involved with during the last 12 months. That's anything you've taken part in, supported, or that you've helped in any way, either on your own or with others. Please exclude giving money and anything that was a requirement of your job.
In the **last 12 months** have you given **unpaid** help to any groups, clubs or organisations in any of the ways shown on this card?

SHOW CARD S

CODE ALL THAT APPLY	Code
Raising or handling money/taking part in sponsored events	1
Leading a group/member of a committee	1
Organising or helping to run an activity or event	1
Visiting people	1
Befriending or mentoring people	1
Giving advice/information/counselling	1
Secretarial, admin or clerical work	1
Providing transport/driving	1
Representing	1
Campaigning	1
Other practical help (e.g. helping out at school, shopping)	1
Member of a club (but not actively involved)	1
Any other help	1
None of the above	1

H5. To what extent do you agree or disagree that this local area (within 15/20 minutes walking distance) is a place where people from different backgrounds get on well together?
SHOW CARD P

Definitely agree	Tend to agree	Tend to disagree	Definitely disagree	SPONTANEOUS ONLY- Too few people in the local area	SPONTANEOUS ONLY- All same backgrounds	SPONTANEOUS ONLY- Don't know
1	2	3	4	5	6	-98

H6. To what extent do you agree or disagree that this local area (15/20 minutes walking distance) is a place where residents respect ethnic differences between people?
SHOW CARD P

Definitely agree	Tend to agree	Tend to disagree	Definitely disagree	Do not know
1	2	3	4	-98

H7. What proportion of your friends have similar incomes to you?
SHOW CARD T

all similar	more than a half	about a half	less than a half	SPONTANEOUS ONLY: Don't have any friends	Rather not say
1	2	3	4	5	-99

H8. Do you agree or disagree that you can influence decisions affecting your local area (15-20 minutes walk)?
SHOW CARD P

Definitely agree	Tend to agree	Tend to disagree	Definitely disagree	Do not know
1	2	3	4	-98

H9. How much do you trust the local council?

SHOW CARD U

A lot	A fair amount	Not very much	Not at all	Do not know	Rather not say
1	2	3	4	-98	-99

PART I

Individual Factors

I1. Age	Code
16-24	1
25-34	2
35-44	3
45-54	4
55-64	5
65-74	6
75+	7
Refused	-99

I3. Occupation of Chief Wage Earner

____

____

Refused (-99)

I2. SHOW CARD V, SINGLE CODE
Ethnicity

A. White	
Scottish	1
Other British	2
Irish	3
Gypsy/Traveller	4
Polish	5
Any other White ethnic group, please describe _______________	6
B. Mixed or Multiple ethnic groups	
Any Mixed or Multiple ethnic groups, please describe _______________	7
C. Asian, Asian Scottish or Asian British	
Pakistani, Pakistani Scottish or Pakistani British	8
Indian, Indian Scottish or Indian British	9
Bangladeshi, Bangladeshi Scottish or Bangladeshi British	10
Chinese, Chinese Scottish or Chinese British	11
Any other Asian, *please describe* _______________	12
D. African	
African, African Scottish or African British	13
Any other African, *please describe* _______________	14

I4. DO NOT READ OUT, SINGLE CODE
Social Class

AB (High managerial, administrative or professional: Intermediate managerial, administrative or professional)	1
C1 (Supervisory, clerical and junior managerial, administrative or professional,)	2
C2 (Skilled manual workers)	3
D (Semi and unskilled manual workers)	4
E (State pensioners, casual or lowest grade workers, unemployed with state benefits only)	5

E. Caribbean or Black	
Caribbean, Caribbean Scottish or Caribbean British	15
Black, Black Scottish or Black British	16
Any other Caribbean or Black, *please describe* ____________________	17
Other ethnic group	
Arab, Arab Scottish or Arab British	18
Any other ethnic group, *please describe* ____________________	19
Refused	-99

I5. What is your country of birth?

READ OUT, SINGLE CODE	Code	Route
Scotland	1	Go to Q I7
England	2	Go to Q I7
Wales	3	Go to Q I7
Northern Ireland	4	Go to Q I7
Republic of Ireland	5	Go to Q I6
Elsewhere, please write in the current name of the country: __________	6	Go to Q I6
Refused	-99	

I6. If you were not born in the United Kingdom, when did you most recently arrive to live here? (Do not count short visits away from the UK)

☐☐ month ☐☐☐☐ Year

I7. What is your working status?

SHOW CARD W

SINGLE CODE	Code
Working full-time (30+ hrs per week)	1
Working part-time (less than 30 hrs per week)	2
Self-employed	3
Unemployed	4
Full time student	5
Retired	6
Student	7
Looking after home/ family	8
Permanently sick/disabled	9
Other (Please specify)	10
Refused	-99

I8. Which of these qualifications do you have?

SHOW CARD X

CODE ALL THAT APPLY	Code
O Grade, Standard Grade, Access 3 Cluster, Intermediate 1 or 2, GCSE, CSE, Senior Certificate or equivalent	1

SCE Higher Grade, Higher, Advanced Higher, CSYS, A Level, AS Level, Advanced Senior Certificate or equivalent	1
GSVQ Foundation or Intermediate, SVQ level 1 or 2, SCOTVEC Module, City and Guilds Craft or equivalent	1
GSVQ Advanced, SVQ level 3, ONC, OND, SCOTVEC National Diploma, City and Guilds Advanced Craft or equivalent	1
HNC, HND, SVQ level 4 or equivalent	1
Degree, Postgraduate qualifications, Masters, PhD, SVQ level 5 or equivalent	1
Professional qualifications (for example, teaching, nursing, accountancy)	1
Other school qualifications not already mentioned (including foreign qualifications)	1
Other post-school but pre-Higher Education qualifications not already mentioned (including foreign qualifications)	1
Other Higher Education qualifications not already mentioned (including foreign qualifications)	1
No qualifications	1
Refused	-99

I9. Are you a registered disabled person?

READ OUT, SINGLE CODE	**Code**
Yes (go to questions I 9.1)	1
No	2
Refused	-99

I9.1.If yes please specify,

READ OUT, SINGLE CODE	**Code**
Receiving benefit (Disability Living Allowance (DLA) and Attendance Allowance (AA)	1
Registered for Blue Badge parking permit	2
Other (please specify)	3

I10. And now could you tell me the letter of the group which represents your household total income in the last 12 months, before any deductions for tax, etc.
SHOW CARD Y

SINGLE CODE	Code

Q - Less than £3,999 **per year**/ less than £333 **per month**/ less than £77 **per week**	1
T - £4,000 - £5,999 **per year**/ £333 to 499 **per month**/ £77 to 115 **per week**	2
O - £6,000-£7,999 **per year**/ £500 to 667 **per month**/ £116 to 154 **per week**	3
K - £8,000-£9,999 **per year**/ £668 to 833 **per month**/ £155 to 192 **per week**	4
L - £10,000-£11,999 **per year**/ £834 to 999 **per month**/ £193 to 230 **per week**	5
B - £12,000-£14,999 **per year**/ £1000 to 1250 **per month**/ £231 to 288 **per week**	6
Z - £15,000-£17,999 **per year**/ £1251 to 1500 **per month**/ £289 to 346 **per week**	7
M - £18,000-£19,999 **per year**/ £1501 to 1667 **per month**/ £347 to 385 **per week**	8
F - £20,000-£22,999 **per year**/ 1£668 to 1917 **per month**/ £386 to 442 **per week**	9
J - £23,000-£25,999 **per year**/ £1918 to 2167 **per month**/ £443 to 500 **per week**	10
D - £26,000-£28,999 **per year**/ £2168 to 2417 **per month**/ £501 to 558 **per week**	11
H - £29,000-£31,999 **per year**/ £2168 to 2667 **per month**/ £559 to 615 **per week**	12
A - £32,000-£37,999 **per year**/ £2668 to 3167 **per month**/ £616 to 731 **per week**	13
W - £38,000-£43,999 **per year**/ £3168 to 3667 **per month**/ £732 to 846 **per week**	14
G - £44,000-£49,999 **per year**/ £3668 to 4167 **per month**/ £847 to 962 **per week**	15
N - £50,000-£55,999 **per year**/ £4168 to 4667 **per month**/ £963 to 1077 **per week**	16
E - £56,000 or more **per year**/ £4668 or more **per month**/ £1078 or more **per week**	17
Refused	-99
Do not know [only use this option for older children who do not know their parents income]	-98

I11. Which of the following descriptions comes closest to how you feel about your household's income at present?
SHOW CARD Z

SINGLE CODE	Code
Living comfortably on present income	1
Coping on present income	2
Finding it difficult on present income	3

Finding it very difficult on present income	4
I don't know / prefer not to answer	-99

I12. Are there young children under 16 years living in the household?

READ OUT, SINGLE CODE	**Code**
Yes	1
No	2 - Go to Q I14
Refused	-99

I13. If yes, do you have responsibility for any of this/these child(ren)?

READ OUT, SINGLE CODE	**Code**
Yes	1
No	2

I14. What type of accommodation is this?
SHOW CARD AA

SINGLE CODE	Code
A. A whole house or bungalow	
detached	1
semi-detached	2
terraced (including end-terraced)	3
B. A flat, maisonette or apartment that is	
in a tenement or purpose-built block of flats (including '4-in-a-block')	4
part of a converted or shared house (including bedsits)	5
in a commercial building (for example, in an office building, hotel or over a shop)	6
C. mobile or temporary structure	
a caravan or other mobile or temporary structure'	7

I15 How satisfied are you with your current accommodation?
SHOW CARD L

Very satisfied	Fairly satisfied	Neither satisfied nor dissatisfied	Slightly dissatisfied	Very dissatisfied
1	2	3	4	5

I16. Does your home suffer from any of the following problems?

READ OUT, CODE ALL THAT APPLY	Code
Damp	1
Vibration	1
Cold	1
Dust	1
Mould	1

SINGLE RESPONSE None	1
DO NOT READ OUT Do not know	-98

I17. Do you have a garden?

READ OUT, SINGLE CODE	Code
Yes, private garden	1
Yes – garden shared with others	2
No	3

I18. Do you own a dog?

READ OUT, SINGLE CODE	Code
Yes	1
No	2

I19. Do you have regular access to a car or other motor vehicle?

READ OUT, SINGLE CODE	Code
Yes	1
No	2

I20. How long have you been living in your current neighbourhood?

READ OUT, SINGLE CODE	Code
Less than 1 year	1
1-3 years	2
4-10 years	3
More than 10 years	4

I21. The results of this study are likely to be available in 2016. Would you be interested in receiving some information about the results when they are ready?

READ OUT, SINGLE CODE	Code
Yes	1
No	2

I22. Would you be willing help us out further by taking part in a focus group later in the year or next year?

READ OUT, SINGLE CODE	Code
Yes	1
No	2

I23. If yes to I21 or I22 please ask for appropriate contact details:

- **Name**
- **Telephone Number**
- **E-mail address**

PROGRESSIVE'S QUALITY CONTROL QUESTIONS:

- Name, Age, Occupation
- Full address, including postcode and telephone number
- Key survey questions to be checked
- Append the code by Professor Richard Mitchell (excel file)

Thank, close and classify

Appendix 2: The EQ-5D 5L, source: EuroQol Group (2009)

Under each heading, please tick the **ONE** box that best describes your health **TODAY**

MOBILITY

I have no problems in walking about ❑
I have slight problems in walking about ❑
I have moderate problems in walking about ❑
I have severe problems in walking about ❑
I am unable to walk about ❑

SELF-CARE

I have no problems washing or dressing myself ❑
I have slight problems washing or dressing myself ❑
I have moderate problems washing or dressing myself ❑
I have severe problems washing or dressing myself ❑
I am unable to wash or dress myself ❑

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

I have no problems doing my usual activities ❑
I have slight problems doing my usual activities ❑
I have moderate problems doing my usual activities ❑
I have severe problems doing my usual activities ❑
I am unable to do my usual activities ❑

PAIN / DISCOMFORT

I have no pain or discomfort ❑
I have slight pain or discomfort ❑
I have moderate pain or discomfort ❑
I have severe pain or discomfort ❑
I have extreme pain or discomfort ❑

ANXIETY / DEPRESSION

I am not anxious or depressed ❑
I am slightly anxious or depressed ❑
I am moderately anxious or depressed ❑
I am severely anxious or depressed ❑
I am extremely anxious or depressed ❑

Appendix 3: Haugh Hill/Pollok before and after physical intervention

Haugh Hill/Pollok before physical intervention

Haugh Hill/Pollok after physical intervention

Appendix 4: Linwood before and after physical intervention

Linwood before physical intervention

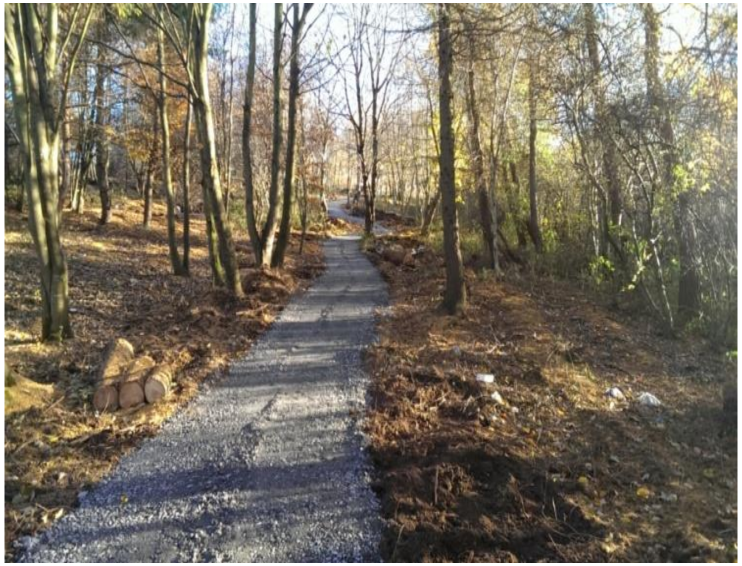
Linwood after physical intervention

Appendix 5:Mayfield before and after physical intervention

Mayfield before physical intervention

Mayfield site after physical intervention

Appendix 6: Cross-sectional and unbalanced panel results for the EQ-5D descriptive system and VAS

Model 1: Cross-sectional Type of site# wave# contact

EQ-5D Utility	Coef.	Std. Err.	P value	95% CI Lower	Upper
Intervention	0.001	0.012	0.938	-0.022	0.024
Wave					
Wave 2	-0.002	0.012	0.871	-0.025	0.021
Wave 3	-0.005	0.012	0.655	-0.028	0.018
Type_site#Wave					
Intervention#wave 2	0.014	0.017	0.411	-0.019	0.047
Intervention#wave 3	-0.016	0.016	0.331	-0.048	0.016
Nature's visits					
ExpNat	0.005	0.011	0.646	-0.017	0.028
Type_site# Nature's visits					
Intervention#ExpNat	0.005	0.016	0.769	-0.026	0.036
Nature's visits #Wave					
ExpNat #wave 2	-0.005	0.016	0.751	-0.037	0.027
ExpNat #wave 3	0.012	0.017	0.458	-0.020	0.045
Type_site# Nature's visits #wave					
Intervention #wave 2# ExpNat	0.007	0.024	0.774	-0.040	0.053
Intervention #wave 3# ExpNat	0.014	0.024	0.551	-0.032	0.061
Age range (16-24)					
25-34	-0.019	0.011	0.083	-0.040	0.002
35-44	-0.038	0.011	0.001	-0.059	-0.016
45-54	-0.071	0.011	0.000	-0.092	-0.050
55-64	-0.099	0.011	0.000	-0.121	-0.077
65-74	-0.114	0.012	0.000	-0.137	-0.091
75+	-0.160	0.013	0.000	-0.185	-0.134
Gender (female)					
Male	0.008	0.005	0.111	-0.002	0.019
Social class (high managerial)					
Supervisory/clerical/junior managerial	-0.024	0.015	0.100	-0.053	0.005
Skilled manual worker	-0.022	0.015	0.146	-0.052	0.008
Semi-unskilled manual worker	-0.016	0.015	0.305	-0.046	0.014
Pensioner/casual/unemployed	-0.011	0.016	0.483	-0.042	0.020
Perceived income (Finding it difficult)					

Coping	0.029	0.007	0.000	0.014	0.043
Living comfortably	0.051	0.009	0.000	0.034	0.069
Distance band (150)					
300	-0.007	0.008	0.386	-0.023	0.009
500	0.010	0.008	0.248	-0.007	0.026
750	-0.019	0.009	0.027	-0.036	-0.002
1500	0.003	0.009	0.764	-0.015	0.021
Working status (No)					
Yes	0.043	0.008	0.000	0.028	0.058
Highest qualification (No qualification)					
1	0.036	0.007	0.000	0.023	0.050
2	0.019	0.009	0.037	0.001	0.036
3	0.017	0.011	0.107	-0.004	0.038
4	-0.004	0.012	0.741	-0.028	0.020
Car ownership (No)					
Yes	0.002	0.006	0.761	-0.010	0.014
Life events (Better)					
Much worse	-0.138	0.013	0.000	-0.163	-0.113
No different	0.012	0.011	0.269	-0.010	0.034
Nothing happened	0.030	0.011	0.004	0.009	0.051
Smoking (Currently smokes)					
Smoked in the past	0.002	0.007	0.772	-0.012	0.016
Never smoked	0.036	0.006	0.000	0.024	0.048
Disability (No)					
Yes	-0.315	0.008	0.000	-0.331	-0.299
Site pair (Mayfield-Glenrothes)					
Glasgow-Milliken park	0.004	0.007	0.557	-0.010	0.018
Linwood-Newarthill	0.001	0.007	0.852	-0.012	0.015
Constant	0.892	0.024	0.000	0.845	0.939

Random-effects Parameters	Estimate	Std. Err.	[95%CI]	
PersonID: Identity				
sd(_cons)	0.057	0.007	0.046	0.072
sd(Residual)	0.170	0.003	0.165	0.176

Model 2: Unbalanced panel Type site#wave#contact

EQ-5D Utility	Coef.	Std. Err.	P value	95% CI Lower	Upper
Intervention	-0.029	0.026	0.253	-0.080	0.021
Wave					
Wave 2	-0.029	0.025	0.240	-0.078	0.020
Wave 3	-0.027	0.023	0.242	-0.071	0.018

Type_site#Wave					
Intervention#wave 2	0.019	0.040	0.635	-0.060	0.098
Intervention#wave 3	0.000	0.036	0.990	-0.071	0.070
Nature's visits					
ExpNat	-0.009	0.023	0.691	-0.054	0.036
Type_site# Nature's visits					
Intervention#ExpNat	0.008	0.033	0.799	-0.057	0.074
Nature's visits #Wave					
ExpNat #wave 2	0.004	0.036	0.919	-0.066	0.074
ExpNat #wave 3	0.061	0.037	0.097	-0.011	0.132
Type_site# Nature's visits #wave					
Intervention #wave 2# ExpNat	0.047	0.057	0.403	-0.064	0.158
Intervention #wave 3# ExpNat	0.000	0.054	0.997	-0.105	0.106
Age range (16-24)					
25-34	-0.002	0.034	0.949	-0.069	0.064
35-44	-0.054	0.034	0.117	-0.121	0.013
45-54	-0.087	0.033	0.008	-0.151	-0.022
55-64	-0.115	0.033	0.000	-0.179	-0.051
65-74	-0.117	0.033	0.000	-0.182	-0.051
75+	-0.130	0.034	0.000	-0.196	-0.064
Gender (female)					
Male	-0.014	0.013	0.285	-0.038	0.011
Social class (high managerial)					
Supervisory/clerical/junior managerial	-0.010	0.034	0.775	-0.075	0.056
Skilled manual worker	-0.002	0.035	0.961	-0.071	0.067
Semi-unskilled manual worker	0.023	0.035	0.514	-0.045	0.091
Pensioner/casual/unemployed	0.009	0.036	0.807	-0.061	0.079
Perceived income (Finding it difficult)					
Coping	0.022	0.017	0.211	-0.012	0.056
Living comfortably	0.055	0.021	0.008	0.014	0.095
Distance band (150)					
300	-0.015	0.020	0.477	-0.055	0.026
500	0.005	0.020	0.807	-0.034	0.044
750	-0.010	0.021	0.615	-0.051	0.030
1500	0.010	0.024	0.669	-0.037	0.058
Working status (No)					
Yes	0.038	0.018	0.031	0.003	0.073
Highest qualification (No qualification)					
1	0.041	0.015	0.008	0.011	0.070
2	0.018	0.021	0.399	-0.023	0.059
3	0.013	0.026	0.623	-0.038	0.063
4	-0.015	0.027	0.575	-0.069	0.038
Car ownership (No)					
Yes	0.013	0.014	0.330	-0.013	0.040
Life events (Better)					
Much worse	-0.115	0.027	0.000	-0.168	-0.061
No different	0.038	0.026	0.139	-0.013	0.089

Nothing happened	0.045	0.024	0.062	-0.002	0.093
Smoking (Currently smokes)					
Smoked in the past	0.006	0.016	0.690	-0.025	0.038
Never smoked	0.043	0.015	0.004	0.014	0.072
Disability (No)					
Yes	-0.321	0.016	0.000	-0.353	-0.289
Site pair (Mayfield-Glenrothes)					
Glasgow-Milliken park	0.007	0.018	0.705	-0.029	0.043
Linwood-Newarthill	0.001	0.018	0.963	-0.035	0.037
Constant	0.866	0.060	0.000	0.748	0.983

Random-effects Parameters	Estimate	Std. Err.	[95%CI]	
PersonID: Identity				
sd(_cons)	0.058	0.013	0.037	0.089
sd(Residual)	0.195	0.005	0.187	0.205

EQ-5D VISUAL ANALOGUE SCALE

Model 1: Type site#wave#contact

EQ-5D VAS score	Coef.	Std. Err.	P value	[95% Conf. Interval]	
Intervention	-1.364	1.067	0.201	-3.456	0.727
Wave 2	-1.446	1.028	0.160	-3.462	0.570
Wave 3	-0.135	0.984	0.891	-2.062	1.793
Contact with woods	1.334	1.023	0.192	-0.671	3.338
Site#contact with woods					
Intervention#contact with woods	1.253	1.421	0.378	-1.533	4.039
Site#Wave					
Intervention#wave 2	3.892	1.423	0.006	1.102	6.682
Intervention#wave 3	-3.620	1.401	0.010	-6.366	-0.875
Contact with woods#Wave					
Contact with woods#wave 2	-2.020	1.419	0.155	-4.800	0.761
Contact with woods#wave 3	-0.799	1.423	0.574	-3.588	1.990
Site #wave#contact					
Intervention #wave 2#contact	-0.213	1.997	0.915	-4.126	3.701
Intervention #wave 3#contact	2.386	1.999	0.233	-1.532	6.305
Age range (16-24)					
25-34	-3.071	0.914	0.001	-4.862	-1.280
35-44	-6.095	0.944	0.000	-7.945	-4.245
45-54	-8.731	0.916	0.000	-10.526	-6.937
55-64	-12.147	0.994	0.000	-14.095	-10.200
65-74	-12.373	1.334	0.000	-14.988	-9.759
75+	-14.297	1.437	0.000	-17.113	-11.480
Male (female)	-0.361	0.459	0.432	-1.260	0.538
Social class (high managerial)					
Supervisory/clerical/junior managerial	-1.633	1.234	0.186	-4.052	0.786
Skilled manual worker	-1.189	1.279	0.352	-3.695	1.317
Semi-unskilled manual worker	-1.292	1.293	0.318	-3.827	1.243
Pensioner/casual/unemployed	-2.191	1.339	0.102	-4.815	0.433
Perceived income (Living comfortably)					
Coping	-3.631	0.513	0.000	-4.636	-2.625
Finding it difficult	-5.751	0.770	0.000	-7.260	-4.242
Finding it very difficult	-5.275	1.189	0.000	-7.605	-2.945
Distance band (150)					
300	0.406	0.732	0.579	-1.029	1.841
500	1.266	0.703	0.072	-0.112	2.644

750	-0.221	0.721	0.759	-1.635	1.193
1500	3.200	0.792	0.000	1.647	4.753
Working status (working)					
Other	-6.141	0.731	0.000	-7.574	-4.708
Retired	-6.480	1.033	0.000	-8.505	-4.455
Unemployed	-2.610	0.922	0.005	-4.416	-0.803
Highest qualification (No qualification)					
1	2.342	0.573	0.000	1.219	3.464
2	2.888	0.751	0.000	1.416	4.360
3	1.591	0.896	0.076	-0.165	3.347
4	0.756	1.016	0.457	-1.235	2.746
Car ownership (No)	0.587	0.508	0.248	-0.408	1.582
Life events (Better)					
Much worse	-10.362	1.042	0.000	-12.404	-8.320
No different	0.532	0.931	0.568	-1.294	2.357
Nothing happened	1.492	0.873	0.088	-0.220	3.203
Smoking (Never)					
Smoked in the past	-2.762	0.559	0.000	-3.857	-1.666
Currently smokes	-3.886	0.518	0.000	-4.901	-2.872
Disability (No) yes	-18.754	0.709	0.000	-20.144	-17.364
Site pair (Mayfield-Glenrothes)					
Glasgow-Milliken park	-1.980	0.573	0.001	-3.103	-0.857
Linwood-Newarthill	-1.458	0.566	0.010	-2.568	-0.348
Constant	95.086	2.010	0.000	91.142	99.030

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
PersonID: Identity				
sd(_cons)	7.387	0.393	6.656	8.199
sd(Residual)	14.585	0.219	14.162	15.020

Model 3: Type site#wave#contact unbalanced

EQ-5D VAS score	Coef.	Std. Err.	P value	[95% Conf. Interval]	
Intervention	-4.955	2.039	0.015	-8.952	-0.958
Wave 2	0.671	1.907	0.725	-3.066	4.408
Wave 3	2.913	1.783	0.102	-0.582	6.407
Contact with woods	2.568	1.839	0.163	-1.036	6.173
Site#contact with woods					
Intervention#contact with woods	0.335	2.593	0.897	-4.748	5.418
Site#Wave					
Intervention#wave 2	1.813	2.918	0.534	-3.906	7.532
Intervention#wave 3	-5.694	2.685	0.034	-10.956	-0.432
Contact with woods#Wave					
Contact with woods#wave 2	-4.658	2.767	0.092	-10.080	0.765
Contact with woods#wave 3	-5.575	2.898	0.054	-11.254	0.105
Site #wave#contact					
Intervention #wave 2#contact	5.531	4.028	0.170	-2.364	13.426
Intervention #wave 3#contact	7.674	3.992	0.055	-0.150	15.498
Age range (16-24)					
25-34	-4.183	2.597	0.107	-9.274	0.907
35-44	-6.577	2.653	0.013	-11.777	-1.376
45-54	-10.577	2.535	0.000	-15.545	-5.610
55-64	-15.700	2.621	0.000	-20.838	-10.563
65-74	-14.275	3.059	0.000	-20.271	-8.279
75+	-14.835	3.144	0.000	-20.998	-8.673
Male (female)	-1.043	1.039	0.315	-3.079	0.992
Social class (high managerial)					
Supervisory/clerical/junior managerial	-0.509	2.579	0.843	-5.563	4.545
Skilled manual worker	0.721	2.676	0.788	-4.525	5.967
Semi-unskilled manual worker	1.124	2.682	0.675	-4.133	6.380
Pensioner/casual/unemployed	-3.021	2.739	0.270	-8.389	2.346
Perceived income (Living comfortably)					
Coping	-4.169	1.029	0.000	-6.187	-2.152
Finding it difficult	-6.698	1.632	0.000	-9.897	-3.499
Finding it very difficult	-5.822	2.651	0.028	-11.020	-0.625
Distance band (150)					
300	-0.196	1.819	0.914	-3.761	3.369
500	-0.530	1.722	0.758	-3.905	2.845

750	-0.349	1.734	0.840	-3.748	3.050
1500	-1.325	2.064	0.521	-5.371	2.721
Working status (working)					
other	-6.520	1.546	0.000	-9.550	-3.490
Retired	-4.558	1.949	0.019	-8.378	-0.738
Unemployed	-2.532	2.041	0.215	-6.532	1.469
Highest qualification (No qualification)					
1	0.251	1.157	0.828	-2.016	2.519
2	1.515	1.604	0.345	-1.628	4.658
3	-0.291	1.890	0.878	-3.995	3.414
4	1.632	1.997	0.414	-2.282	5.546
Car ownership (No) Yes	1.291	1.037	0.213	-0.741	3.323
Life events (Better)					
Much worse	-6.451	2.007	0.001	-10.384	-2.518
No different	2.769	1.908	0.147	-0.970	6.508
Nothing happened	2.419	1.774	0.173	-1.058	5.895
Smoking (Never)					
Smoked in the past	-2.397	1.112	0.031	-4.576	-0.218
Currently smokes	-3.385	1.153	0.003	-5.645	-1.124
Disability (No) yes	-17.937	1.300	0.000	-20.485	-15.389
Site pair (Mayfield-Glenrothes)					
Glasgow-Milliken park	-1.715	1.313	0.191	-4.287	0.858
Linwood-Newarthill	-3.885	1.412	0.006	-6.653	-1.117
Constant	95.451	4.492	0.000	86.647	104.255

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
PersonID: Identity				
sd(_cons)	7.669	0.668	6.465	9.097
sd(Residual)	15.742	0.359	15.055	16.461

Appendix 7: Ethics approval letter

University of Glasgow | College of Medical, Veterinary & Life Sciences

4th November 2014

Dear Professor Briggs

MVLS College Ethics Committee

Project Title: An economic evaluation of woodlands' improvements as a public health intervention for mental well-being in Scotland
Project No: 200140011

The College Ethics Committee has reviewed your application and has agreed that there is no objection on ethical grounds to the proposed study. It is happy therefore to approve the project, subject to the following conditions:

- Project end date: **31 December 2014**
- The research should be carried out only on the sites, and/or with the groups defined in the application.
- Any proposed changes in the protocol should be submitted for reassessment, except when it is necessary to change the protocol to eliminate hazard to the subjects or where the change involves only the administrative aspects of the project. The Ethics Committee should be informed of any such changes.
- If the study does not start within three years of the date of this letter, the project should be resubmitted.
- You should submit a short end of study report to the Ethics Committee within 3 months of completion.

Yours sincerely

Dorothy McKeegan

Dr Dorothy McKeegan
College Ethics Officer

Dr Dorothy McKeegan

Senior Lecturer

R303 Level 3
Institute of Biodiversity Animal Health and Comparative Medicine
Jarrett Building
Glasgow G61 1QH Tel: 0141 330 5712
E-mail: Dorothy.McKeegan@glasgow.ac.uk

Appendix 8: SPDCE main study questionnaire

University of Glasgow

SURVEY TO ESTABLISH THE VALUE OF USING WOODLANDS FOR HEALTH IN SCOTLAND

Scientists are suggesting that **being in and around nature is good for you**. We find this interesting. The Health Economics and Public Health units of the University of Glasgow are carrying out research to establish the **value** of woodlands for the **well-being** of people. This is funded by the Forestry Commission Scotland.

You are kindly being asked to participate in this study. We would be interested in your **views** and your **preferences of woodland characteristics**. If you choose to take part, please complete this survey. The first section of the survey presents a series of questions about woodlands and asks your **preference** between **woodland A** and **woodland B** with different scenarios of **woodland characteristics** which make or would make you **visit** the woodlands and **do the things you want to do either alone or with others**. You may choose **neither** of these woodlands. The survey then asks some follow-up questions in the second section. The responses you provide in this survey will not affect you or your visits to woodlands in anyway. This will all take about **20 minutes** of your time.

Please note:
Although the **scenarios** of the woodland characteristics described in this survey are **hypothetical**, it is important that you to provide **answers** as you would in real life.

The survey is **anonymous** and all responses will be **strictly confidential** and only used for the purposes of this study. Remember that completing this survey is **voluntary** and you may choose to **stop** at any time in which case your responses **will not be used**.

The findings will be used for **academic purposes** and **will help in decision-making over woodlands in Scotland.**

If you agree to take part, please <u>"NEXT".</u>

<u>If you do not agree simply close your browser window</u>**. Your decision to complete** this survey will be **interpreted** as an indication of your **consent** to participate.

Thank you for your **time** and **cooperation**.

SECTION A: WOODLAND PREFERENCES

So **imagine** a situation where you have access to either **woodland (A)** or **woodland (B)** which are different from each other in the following characteristics:

1. The **woodland environmental support** which we define as **one which allows you to do the things you want to do, either on your own or with others** (such as exercise, relaxing, enjoying wildlife) and makes it **easy** and **enjoyable** to do them.
2. The **time** that it takes **to walk from home** to **the woodland**.
3. The **quality of the woodland environment** which include cleanliness; the condition of paths and entrances; the naturalness of its appearance; the views of plants and wildlife.

4. The **opportunities for social activities** that the woodland offers you such as meeting people, community events, guidance on how to use the woodland and about what is going on there.
5. The **cost for access** to the woodlands, if you **imagine** you lived in a country where you had to pay, in the form of an **annual subscription,** in order to go there.

Thinking about these characteristics of woodlands, please consider the following **choices** and say which **woodland: (A)** or **(B)** you would **prefer** by **clicking** in the appropriate button. You may also prefer to select **'Neither'** of them.

First, we provide you with an example of the **choice task**. **Please look at this example and think about how the two types of woodlands differ from each other in terms of their characteristics and choose the one you prefer**.

Example

Characteristics	**Woodland A**	**Woodland B**
1. The **supportive woodland environment** that allows you to do enjoyable activities easily	**Some** support	**A lot** of support
2. The **time** that it takes to walk from home to the woodland	**50**mins	**15**mins
3. The **quality of the woodland environment**	**Poor** quality	**Good** quality
4. The **opportunities for social activities**	**Some** opportunities	**Many** opportunities
5. The **cost for access** to the woodland	**£50**	**£50**

Which woodland would you prefer? Woodland A ☐ Woodland B ☐
Neither ☐

(Tick one box only)

Dominant alternative

Thank you for your preference in this example, now proceed to make more preferences below. Please answer every choice task.

Choice task 1

Characteristics	**Woodland A**	**Woodland B**
1. The **supportive woodland environment** that allows you to do enjoyable activities easily	**Some** support	**A lot** of support
2. The **time** that it takes to walk from home to the woodland	**50**mins	**5**mins
3. The **quality of the woodland environment**	**Poor** quality	**Average** quality
4. The **opportunities for social activities**	**No** opportunities	**Some** opportunities
5. The **cost for access** to the woodland	**£50**	**£0**

Which **woodland** would you **prefer**? **Woodland A** ☐ **Woodland B** ☐

Neither ☐

(Tick one box only)

No attribute-level overlap

Choice task 2

Characteristics	**Woodland A**	**Woodland B**
1. The **supportive woodland environment** that allows you to do enjoyable activities easily	**A lot** of support	**Some** support
2. The **time** that it takes to walk from home to the woodland	**5**mins	**50**mins
3. The **quality of the woodland environment**	**Poor** quality	**Average** quality
4. The **opportunities for social activities**	**Some** opportunities	**No** opportunities
5. The **cost for access** to the woodland	£**0**	£**50**

Which woodland would you prefer? Woodland A ☐ Woodland B ☐
Neither ☐

(Tick one box only)

Choice task 3

Characteristics	Woodland A	Woodland B
1. The **supportive woodland environment** that allows you to do enjoyable activities easily	**Some** support	**A lot** of support
2. The **time** that it takes to walk from home to the woodland	**50**mins	**15**mins
3. The **quality of the woodland environment**	**Poor** quality	**Good** quality
4. The **opportunities for social activities**	**Some** opportunities	**Many** opportunities
5. The **cost for access** to the woodland	**£50**	**£50**

Which woodland would you prefer? Woodland A ☐ Woodland B ☐
Neither ☐

(Tick one box only)

Attribute-level overlap

Then the questionnaire continues from Choice task 4 to choice task 20.

SECTION B: FOLLOW-UP QUESTIONS

1. Are there any other characteristics of woodlands that you consider important but were not included and you would want them included?

   ☐ Yes Please specify below:

   ....

   ☐ No

2. Please indicate how easy or difficult it was to make your choices in the first question of section A

   Please select one answer only

   ☐ Very easy

   ☐ Easy

   ☐ OK

   ☐ Difficult

   ☐ Very difficult

3. We are interested in knowing how many minutes you are prepared to walk from your home to your local woodland? ***Please enter your answer in the text box below***

   .... minutes.

4. Do you have any additional comments?

☐ Yes Please specify below:

……………………………………………………………………………………

☐ No

SECTION C: WILLINGNESS TO PAY

In this section, we are interested in knowing how much you value woodlands. One way of finding out this is to ask you how much you would be willing to pay for access to these woodlands per year in a form of an annual subscription. Remember you would not have to actually pay for access to woodlands in practice but imagine you lived in a country where you had to pay for it.

5. Please **tick** the amount you are sure you would be willing to pay per annum to access woodlands.

☐ £0	☐ £15	☐ £30	☐ £45
☐ £5	☐ £20	☐ £35	☑ £50
☐ £10	☐ £25	☐ £40	

6. If you would be willing to pay more than £50, please state the maximum amount you would be willing to pay ☐

7. If you are not prepared to pay any amount per annum to access woodlands, could you please state the reason?

………………………………………………………………………………………..

8. Please indicate how easy or difficult it was to make your choice in question 5 of this Section C

Please select one answer only

☐ Very easy

☐ Easy

☐ OK

☐ Difficult

☐ Very difficult

9. Do you have any additional comments about section C?

☐ Yes Please specify below:

....

☐ No

SECTION D: ABOUT YOU

We would like to understand your answers better and it is important that we ask a few questions about you. All this information will remain confidential and anonymous.

Please tick **(√).**

☐ Male ☐ Female

10. What is your age?

☐ 16-24 ☐ 65+

☐ 25-34

☐ 35-44

☐ 45-54

☐ 55-64

11. Are you working?

☐ Part-time ☐ Not working

☐ Full time ☐ Student

☐ Other

12. What is your highest level of education?
(Please tick one box only)

☐ Secondary school

☐ Vocational/trade/college

☐ Higher/A levels

☐ University

☐ Other

13. What is your ethnic background?

☐ White Scottish
☐ Any other European
☐ Any mixed or multiple ethnicity
☐ Asian, Asian Scottish or Asian British
☐ Pakistan, Pakistan Scottish or Pakistan British

- ☐ Indian, Indian Scottish or Indian British
- ☐ Chinese, Chinese Scottish or Chinese British
- ☐ Any other Asian
- ☐ African, African Scottish or African British
- ☐ Any other African
- ☐ Caribbean, Caribbean Scottish or Caribbean British
- ☐ Any other Caribbean
- ☐ Arab, Arab Scottish or Arab British
- ☐ Any other Arab
- ☐ I do not want to state my ethnicity

14. Please state which group represents your household total income in the last 12 months, before any deductions for tax?

☐ less than £3,999 ☐ £4,000-£5,999 ☐ £6,000-£7,999

☐ £8,000-£9,999

☐ £10,000-£11,999 ☐ £12,000-£13,999 ☐ £14,000-15,999

☐ £16,000-£17,999

☐ £18,000-£19,999 ☐ £20,000-£22,999 ☐ £23,000-£25,999

☐ £26,000-£28,999

☐ £29,000-£31,999 ☐ £32,000-£37,999 ☐ £38,000-£43,999

☐ £44,000-£49,999

☐ £50,000-£55,999 ☐ £56,000 or more

15. Are there any children under 16 years living in your household?
- ☐ Yes
- ☐ No

16. Do you own a dog?

- ☐ Yes
- ☐ No

17. Do you consider yourself disabled?
    - ☐ Yes
    - ☐ No

18. Are there any other comments that you would like to make about this questionnaire?
    - ☐ Yes   Please specify below:

    ....

    - ☐ No

Thank you very much for your time

References

Abadie, A. 2005. Semiparametric difference-in-differences estimators. *The Review of Economic Studies,* 72, 1-19.

Abraham, A., Sommerhalder, K. & Abel, T. 2010. Landscape and well-being: a scoping study on the health-promoting impact of outdoor environments. *International Journal of Public Health,* 55, 59-69.

Africa, J., Logan, A., Mitchell, R., Korpela, K., Allen, D., Tyrväinen, L., Nisbet, E., Li, Q., Tsunetsugu, Y. & Miyazaki, Y. 2014. The Natural Environments Initiative: Illustrative Review and Workshop Statement. *Center for Health and the Global Environment at the Harvard School of Public Health*.

Aizaki, H. & Aizaki, M. H. 2015. Package 'support. CEs'.

Albaladejo-Pina, I. P. & Díaz-Delfa, M. T. 2009. Tourist preferences for rural house stays: Evidence from discrete choice modelling in Spain. *Tourism Management,* 30, 805-811.

Albright, J. J. & Marinova, D. M. 2010. Estimating multilevel models using SPSS, Stata, SAS, and R. *Indiana University*.

Alcock, I., White, M. P., Wheeler, B. W., Fleming, L. E. & Depledge, M. H. 2014. Longitudinal effects on mental health of moving to greener and less green urban areas. *Environmental science & technology,* 48, 1247-1255.

Almanza, E., Jerrett, M., Dunton, G., Seto, E. & Pentz, M. A. 2012. A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. *Health & place,* 18, 46-54.

Alriksson, S. & Öberg, T. 2008. Conjoint analysis for environmental evaluation. *Environmental Science and Pollution Research,* 15, 244-257.

Altman, D. G. & Bland, J. M. 1995. Statistics notes: Absence of evidence is not evidence of absence. *Bmj,* 311, 485.

Alves, S., Aspinall, P. A., Thompson, C. W., Sugiyama, T., Brice, R. & Vickers, A. 2008. *Facilities,* 26, 433-453.

Angrist, J. D. & Pischke, J.-S. 2008. *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press.

Annerstedt, M., Norman, J., Boman, M., Mattsson, L., Grahn, P. & Währborg, P. 2010. Finding stress relief in a forest. *Ecol. Bull,* 53, 33-42.

Annerstedt, M., Ostergren, P. O., Bjork, J., Grahn, P., Skarback, E. & Wahrborg, P. 2012. Green qualities in the neighbourhood and mental health - results from a longitudinal cohort study in Southern Sweden. *BMC Public Health,* 12, 337.

Aravena, C., Martinsson, P. & Scarpa, R. 2014. Does money talk? — The effect of a monetary attribute on the marginal values in a choice experiment. *Energy Economics,* 44, 483-491.

Arnberger, A. & Eder, R. 2012. The influence of green space on community attachment of urban and suburban residents. *Urban Forestry & Urban Greening,* 11, 41-49.

Aspinall, P., Mavros, P., Coyne, R. & Roe, J. 2013. The urban brain: analysing outdoor physical activity with mobile EEG. *British Journal of Sports Medicine*.

Azur, M. J., Stuart, E. A., Frangakis, C. & Leaf, P. J. 2011. Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research,* 20, 40-49.

Baio, G. & Dawid, A. P. 2011. Probabilistic sensitivity analysis in health economics. *Statistical methods in medical research,* 0962280211419832.

Bala, M. V., Zarkin, G. A. & Mauskopf, J. A. 2002. Conditions for the Near Equivalence of Cost-Effectiveness and Cost-Benefit Analyses. *Value in Health,* 5, 338-346.

Balfour, R. & Allen, J. 2014. Local action on inequalities: Improving access to green spaces. *Health equity briefing 8.* UCL Institute of Health Equity (IHE) commissioned by Public Health England.

Barbosa, O., Tratalos, J. A., Armsworth, P. R., Davies, R. G., Fuller, R. A., Johnson, P. & Gaston, K. J. 2007. Who benefits from access to green space? A case study from Sheffield, UK. *Landscape and Urban Planning,* 83, 187-195.

Barrett, M. A., Miller, D. & Frumkin, H. 2014. Parks and Health: Aligning Incentives to Create Innovations in Chronic Disease Prevention. *Preventing chronic disease,* 11.

Bateman, I. J. 1996. Household willingness to pay and farmers' willingness to accept compensation for establishing a recreational woodland. *Journal of Environmental Planning and Management,* 39, 21-44.

Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S. & Özdemiroglu, E. 2002. Economic valuation with stated preference techniques: a manual. *Economic valuation with stated preference techniques: a manual*.

Bates, S. & Marquit, J. 2011. Space psychology: natural elements in habitation design. *Personal and Ubiquitous Computing,* 15, 519-523.

Bech, M. & Gyrd-Hansen, D. 2005. Effects coding in discrete choice experiments. *Health economics,* 14, 1079-1083.

Belfield, C. & Levin, H. M. 2010. Cost–Benefit Analysis and Cost–Effectiveness Analysis. *In:* Editors-in-Chief: Penelope, P., Eva, B., Barry McGawA2 - Editors-in-Chief: Penelope Peterson, E. B. & Barry, M. (eds.) *International Encyclopedia of Education (Third Edition)*. Oxford: Elsevier.

Bell, S. & Ward Thompson, C. 2008. Health, identity, and sense of place: The importance of local landscapes. *In:* Berlan-Darqué, M., Luginbühl, Y. & Terrasson, D. (eds.) *Landscape, from Knowledge to Action*. France: Editions Quae.

Bennett, J. 1999. *Some Fundamentals of Environmental Choice Modelling,* Australia, School of Economics and Management, University of New South Wales.

Bertram, C., Meyerhoff, J., Rehdanz, K. & Wüstemann, H. 2017. Differences in the recreational value of urban parks between weekdays and weekends: A discrete choice analysis. *Landscape and Urban Planning,* 159**,** 5-14.

Bertrand, M., Duflo, E. & Mullainathan, S. 2004. How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics,* 119**,** 249-275.

Birch, S. & Gafni, A. 1996. Cost-Effectiveness and Cost Utility Analyses: Methods for the Non-Economic Evaluation of Health Care Programmes and How We Can Do Better. *Managing Technology in Healthcare*. Springer.

Blamey, R. K., Bennett, J. W., Louviere, J. J., Morrison, M. & Rolfe, J. 2000. A test of policy labels in environmental choice modelling studies. *Ecological Economics,* 32**,** 269-286.

Bleichrodt, H. & Quiggin, J. 1999. Life-cycle preferences over consumption and health: when is cost-effectiveness analysis equivalent to cost–benefit analysis? *Journal of Health Economics,* 18**,** 681-708.

Bliemer, M. C. J., Rose, J. M. & Hess, S. 2008. Approximation of bayesian efficiency in experimental choice designs. *Journal of Choice Modelling,* 1**,** 98-126.

Bobinac, A., van Exel, N. J. A., Rutten, F. F. & Brouwer, W. B. 2012. Inquiry into the relationship between equity weights and the value of the QALY. *Value in Health,* 15**,** 1119-1126.

Botosaru, I. & Gutierrez, F. H. 2015. Difference-in-Differences When Treatment Status is Observed in Only One Period. *SSRN 2377483*.

Bowler, D., Buyung-Ali, L., Knight, T. & Pullin, A. 2010. A systematic review of evidence for the added benefits to health of exposure to natural environments. *BMC Public Health,* 10**,** 456.

Boxall, P., Adamowicz, W. L. & Moon, A. 2009. Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement*. *Australian Journal of Agricultural and Resource Economics,* 53, 503-519.

Bratman, G. N., Hamilton, J. P. & Daily, G. C. 2012. The impacts of nature experience on human cognitive function and mental health. *Annals of the New York Academy of Sciences,* 1249, 118-136.

Brazier, J. 2010. Is the EQ-5D fit for purpose in mental health? *The British Journal of Psychiatry,* 197, 348-349.

Brazier, J., Green, C., McCabe, C. & Stevens, K. 2003. Use of visual analog scales in economic evaluation. *Expert review of pharmacoeconomics & outcomes research,* 3, 293-302.

Brazier, J., Ratcliffe, J. & Tsuchiya, A. 2007. *Measuring and valuing health benefits for economic evaluation*, Oxford University Press.

Breffle, W. 2008. In pursuit of optimal design: a guide for choice experiment practitioners. *International Journal of Ecological Economics and Statistics,* 14.

Brent, R. J. 2003. *Cost- Benefit Analysis and Health Care Evaluations,* Cheltenham United Kingdom, Edward Elgar Publishing Limited.

Bridges, J. F. P., Hauber, A. B., Marshall, D., Lloyd, A., Prosser, L. A., Regier, D. A., Johnson, F. R. & Mauskopf, J. 2011. Conjoint Analysis Applications in Health—a Checklist: A Report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value in Health,* 14, 403-413.

Briggs, A. 2016. A View from the Bridge: Health Economic Evaluation — A Value-Based Framework? *Health Economics,* 25, 1499-1502.

Briggs, A. & Fenn, P. 1998. Confidence intervals or surfaces ? Uncertainty on the cost-effectiveness plane. *Health Economics,* 7.

Briggs, A., Sculpher, M. & Claxton, K. 2006. *Decision modelling for health economic evaluation*, Oxford university press.

Briggs, A. H. 2000. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics,* 17, 479-500.

Briggs, A. H. 2009. Cost-Benefit Analysis. *Encyclopedia of Medical Decision Making*. Thousand Oaks, CA: SAGE Publications, Inc.

Briggs, A. H. & O'Brien, B. J. 2001. The death of cost-minimization analysis? *Health Economics,* 10, 179-184.

Briggs, A. H., Parfrey, P. S., Khan, N., Tseng, S., Dehmel, B., Kubo, Y., Chertow, G. M. & Belozeroff, V. 2016. Analyzing Health-Related Quality of Life in the EVOLVE Trial The Joint Impact of Treatment and Clinical Events. *Medical Decision Making*, 0272989X16638312.

Briggs, A. H., Weinstein, M. C., Fenwick, E. A., Karnon, J., Sculpher, M. J. & Paltiel, A. D. 2012. Model parameter estimation and uncertainty analysis a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. *Medical Decision Making,* 32, 722-732.

Briggs, A. H., Wonderling, D. E. & Mooney, C. Z. 1997. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Economics,* 6.
Brousselle, A. & Lessard, C. 2011. Economic evaluation to inform health care decision-making: promise, pitfalls and a proposal for an alternative path. *Social science & medicine (1982),* 72, 832-839.
Brouwer, W. 2009. *Welfare, Welfarism, and Extrawelfarism,* Thousand Oaks, CA, SAGE Publications, Inc.
Brouwer, W. B., Culyer, A. J., van Exel, N. J. & Rutten, F. F. 2008. Welfarism vs. extra-welfarism. *J Health Econ,* 27, 325-38.
Brown, M. M. & Brown, G. C. 2005. How to interpret a healthcare economic analysis. *Current Opinion in Ophthalmology,* 16, 191-194.
Buchanan, J. & Wordsworth, S. 2015. Welfarism versus extra-welfarism: can the choice of economic evaluation approach impact on the adoption decisions recommended by economic evaluation studies? *PharmacoEconomics,* 33, 571-579.
Burgess, D., Finney, G., Matthews, D. & Patton, M. 2012. Landscape Valuation: Choice Experiments or Contingent Valuation? *Landscape,* 19, 20.
Burgess, L. & Street, D. J. 2005. Optimal designs for choice experiments with asymmetric attributes. *Journal of Statistical Planning and Inference,* 134, 288-301.
Burls, A. 2007. People and green spaces: promoting public health and mental well-being through ecotherapy. *Journal of Public Mental Health,* 6, 24-39.
Burton, C. R., Fargher, E., Plumpton, C., Roberts, G. W., Owen, H. & Roberts, E. 2014. Investigating preferences for support with life after stroke: a discrete choice experiment. *BMC health services research,* 14, 63.
Campbell, M., Fitzpatrick, R., Haines, A. & Kinmonth, A. L. 2000. Framework for design and evaluation of complex interventions to improve health. *British medical journal,* 321, 694.
Campbell, M. J., Machin, D. & Walters, S. J. 2007. *Medical statistics: a textbook for the health sciences*, John Wiley & Sons.
Can, Ö. & Alp, E. 2012. Valuation of environmental improvements in a specially protected marine area: A choice experiment approach in Göcek Bay, Turkey. *Science of The Total Environment,* 439, 291-298.
Canby Iv, J. B. 1995. Applying activity-based costing to healthcare settings. *hfm (Healthcare Financial Management),* 49, 50.
Canter, D. 1977. *The psychology of place,* London, Architectural Press.
Carey, K. & Stefos, T. 2011. Measuring the cost of hospital adverse patient safety events. *Health Economics,* 20, 1417-1430.

Carlsson, F. 2011. *Non-Market Valuation: Stated Preference Methods*, 'Oxford University Press'.
Carlsson, F., Mørkbak, M. R. & Olsen, S. B. 2012. The first time is the hardest: A test of ordering effects in choice experiments. *Journal of Choice Modelling,* 5**,** 19-37.
Carson, R. T. & Louviere, J. J. 2010. Experimental design and the estimation of willingness to pay in choice experiments for health policy evaluation. *Applied Methods of Cost-Benefit Analysis in Health Care,* 4**,** 185.
Champ, P. 2003. Collecting Survey Data for Nonmarket Valuation. *In:* Champ, P., Boyle, K. & Brown, T. (eds.) *A Primer on Nonmarket Valuation.* Springer Netherlands.
Chapko, M. K., Liu, C. F., Perkins, M., Li, Y. F., Fortney, J. C. & Maciejewski, M. L. 2009. Equivalence of two healthcare costing methods: bottom-up and top-down. *Health Econ,* 18**,** 1188-201.
Choicemetrics 2014. Ngene 1.1.2: User manual and reference guide. Choice Metrics Pty Ltd.
Clark, M., Determann, D., Petrou, S., Moro, D. & de Bekker-Grob, E. 2014. Discrete Choice Experiments in Health Economics: A Review of the Literature. *PharmacoEconomics***,** 1-20.
Claxton, K. & Cookson, R. 2012. *The Humble Economist: Tony Culyer on Health, Health Care and Social Decision Making*, University of York.
Claxton, K., Fenwick, E. & Sculpher, M. J. 2012. Decision-making with Uncertainty: The Value of Information. *In:* Jones, A. M. (ed.) *The Elgar Companion to Health Economics.* Second ed.: Edward Elgar Publishing.
Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P. C. & Sculpher, M. 2015. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health technology assessment (Winchester, England)***,** 1-542.
Claxton, K., Sculpher, M. & Culyer, A. 2007. Mark versus Luke? Appropriate methods for the evaluation of public health interventions. CHE Research Paper 2007, 31 ed. Alcuin College, University of York York, UK: Centre for Health Economics.
Claxton, K., Sculpher, M., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., Brazier, J. & O'Hagan, T. 2005. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health economics,* 14**,** 339-347.
Coast, J. 2004. Is economic evaluation in touch with society's health values? *BMJ,* 329.
Coast, J. 2009. Maximisation in extra-welfarism: A critique of the current position in health economics. *Soc Sci Med,* 69**,** 786-92.

Coast, J., Al-Janabi, H., Sutton, E. J., Horrocks, S. A., Vosper, A. J., Swancutt, D. R. & Flynn, T. N. 2012. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Economics,* 21, 730-741.

Coast, J. & Horrocks, S. 2007. Developing attributes and levels for discrete choice experiments using qualitative methods. *Journal of Health Services Research & Policy,* 12, 25-30.

Coast, J., Huynh, E., Kinghorn, P. & Flynn, T. 2015a. Complex Valuation: Applying Ideas from the Complex Intervention Framework to Valuation of a New Measure for End-of-Life Care. *PharmacoEconomics,* 1-10.

Coast, J., Kinghorn, P. & Mitchell, P. 2015b. The Development of Capability Measures in Health Economics: Opportunities, Challenges and Progress. *The Patient - Patient-Centered Outcomes Research,* 8, 119-126.

Coast, J., Smith, R. & Lorgelly, P. 2008a. Should the capability approach be applied in health economics? *Health Econ,* 17.

Coast, J., Smith, R. & Lorgelly, P. 2008b. Should the capability approach be applied in Health Economics? *Health Economics,* 17, 667-670.

Conedera, M., Del Biaggio, A., Seeland, K., Moretti, M. & Home, R. 2015. Residents' preferences and use of urban and peri-urban green spaces in a Swiss mountainous region of the Southern Alps. *Urban Forestry & Urban Greening,* 14, 139-147.

Cookson, R., Drummond, M. & Weatherly, H. 2009. Explicit incorporation of equity considerations into economic evaluation of public health interventions. *Health Econ Policy Law,* 4.

Coombes, E., Jones, A. P. & Hillsdon, M. 2010. The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Soc Sci Med,* 70, 816-22.

Craig, P., Cooper, C. & Gunnell, D. 2011. Using Natural Experiments to Evaluate Population Health Interventions. *Glasgow, UK: Medical Research Council.*

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. & Petticrew, M. 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ: British Medical Journal,* 337.

Culyer, A. J. 2015. Four Issues in Cost-Effectiveness Analysis and Health Technology Assessment: a View from the Touch-line. *In:* del Llano-Señarís, J. E. & Campillo-Artero, C. (eds.) *Health Technology Assessment and Health Policy Today: A Multifaceted View of their Unstable Crossroads.* Cham: Springer International Publishing.

Culyer, A. J. & Cookson, R. 2012. *The Humble Economist: Tony Culyer on Health, Health Care and Social Decision Making,* University of York.

Cunningham, S. J. 2001. An introduction to economic evaluation of health care. *Journal of orthodontics,* 28**,** 246-250.
Curtis, L. J. 2014. Economics, Health Economics, Evaluation, and Public Health. *Background Paper*. Canada: Department of Economics University of Waterloo.
Dadvand, P., de Nazelle, A., Figueras, F., Basagaña, X., Su, J., Amoly, E., Jerrett, M., Vrijheid, M., Sunyer, J. & Nieuwenhuijsen, M. J. 2012. Green space, health inequality and pregnancy. *Environment International,* 40**,** 110-115.
Dahlgren, G. & Whitehead, M. 1991. Policies and strategies to promote social equity in health. Background document to WHO-Strategy paper for Europe. Institute for Futures Studies.
Dakin, H. & Wordsworth, S. 2011. Cost-minimisation analysis versus cost-effectiveness analysis, revisited. *Health Economics***,** n/a-n/a.
Dallat, M. A. T., Soerjomataram, I., Hunter, R. F., Tully, M. A., Cairns, K. J. & Kee, F. 2013. Urban greenways have the potential to increase physical activity levels cost-effectively. *The European Journal of Public Health*.
Daly, A., Dekker, T. & Hess, S. 2016. Dummy coding vs effects coding for categorical variables: Clarifications and extensions. *Journal of Choice Modelling*.
Davalos, M. E., French, M. T., Burdick, A. E. & Simmons, S. C. 2009. Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association,* 15**,** 933-948.
David, E., Jake, M., Liz, O. B., Vadims, S. & Gregory, V. 2008. The Economic and social contribution of forestry for the people in Scotland. *Forestry Commission Scotland*.
De Ayala, A., Hoyos Ramos, D. & Mariel Chladkova, P. 2012. Landscape valuation through discrete choice experiments: Current practice and future research reflections. *Biltoki*.
de Bekker-Grob, E., Bliemer, M., Donkers, B., Essink-Bot, M., Korfage, I., Roobol, M., Bangma, C. & Steyerberg, E. 2013. Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. *British journal of cancer,* 109**,** 633-640.
de Bekker-Grob, E., Donkers, B., Jonker, M. & Stolk, E. 2015. Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. *The Patient - Patient-Centered Outcomes Research***,** 1-12.
de Bekker-Grob, E. W., Ryan, M. & Gerard, K. 2012. Discrete choice experiments in health economics: a review of the literature. *Health Economics,* 21**,** 145-172.

de Vries, S., van Dillen, S. M., Groenewegen, P. P. & Spreeuwenberg, P. 2013. Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. *Social Science & Medicine,* 94, 26-33.

Dean, J., van Dooren, K. & Weinstein, P. 2011. Does biodiversity improve mental health in urban settings? *Medical Hypotheses,* 76, 877-880.

Deber, R. 2009. Choice Theories. Encyclopedia of Medical Decision Making. SAGE Publications, Inc. Thousand Oaks, CA: SAGE Publications, Inc.

Devlin, N. 2016. 5 Things you should do with EQ-5D data. *Office of Health Economics News* [Online]. Available: https://www.ohe.org/news/5-things-you-should-do-eq-5d-data [Accessed 04/02/2016].

Devlin, N., Shah, K., Feng, Y., Mulhern, B. & van Hout, B. 2016. Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England. *Office of Health Economics (OHE) Research,* Research Paper 16/01.

Devlin, N. J. & Lorgelly, P. K. 2016. QALYs as a Measure of Value in Cancer. *Journal of Cancer Policy*.

Di Nardo, F., Saulle, R. & La Torre, G. 2012. Green areas and health outcomes: a systematic review of the scientific literature. *Italian Journal of Public Health*, 7.

Dimick, J. B. & Ryan, A. M. 2014. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA,* 312, 2401-2402.

Doherty, E., Campbell, D., Hynes, S. & van Rensburg, T. M. 2013. Examining labelling effects within discrete choice experiments: An application to recreational site choice. *Journal of Environmental Management*, 125, 94-104.

Dolan, P. 1997. Modeling valuations for EuroQol health states. *Medical care,* 35, 1095-1108.

Dolan, P. & Edlin, R. 2002. Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? *Journal of Health Economics,* 21, 827-43.

Domínguez-Torreiro, M. 2014. Alternative experimental design paradigms in choice experiments and their effects on consumer demand estimates for beef from endangered local cattle breeds: An empirical test. *Food Quality and Preference,* 35, 15-23.

Donaldson, C. 1998a. The (Near) Equivalence of Cost-Effectiveness and Cost-Benefit Analyses: Fact or Fallacy? *PharmacoEconomics,* 13, 389-396.

Donaldson, C. 1998b. The (Near) Equivalence of Cost—Effectiveness and Cost-Benefit Analyses. *Pharmacoeconomics,* 13, 389-396.

Donaldson, C., Baker, R., Mason, H., Jones-Lee, M., Lancsar, E., Wildman, J., Bateman, I., Loomes, G., Robinson, A. & Sugden, R. 2011. The social value of a QALY: raising the bar or barring the raise? *BMC health services research,* 11**,** 1.

Donovan, G. H., Michael, Y. L., Butry, D. T., Sullivan, A. D. & Chase, J. M. 2011. Urban trees and the risk of poor birth outcomes. *Health & Place,* 17**,** 390-393.

Dowless, R. M. 2007. Your guide to costing methods and terminology. *Nursing Management,* 38**,** 52-57.

Drummond, M., Brixner, D., Gold, M., Kind, P., McGuire, A. & Nord, E. 2009. Toward a consensus on the QALY. *Value in Health,* 12**,** S31-S35.

Drummond, M., Weatherly, H., Claxton, K., Cookson, R., Ferguson, B. & Godfrey, C. 2007. Assessing the challenges of applying standard methods of economic evaluation to public health programmes. *London: Public Health Research Consortium*.

Drummond, M. F., O'Brien, B. & Stoddart, G. L. 2005. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L. & Torrance, G. W. 2015. *Methods for the Economic Evaluation of Health Care Programmes*, Oxford University Press.

Edejer, T. T.-T. 2003. *Making choices in health: WHO guide to cost-effectiveness analysis*, World Health Organization.

Edlin, R., McCabe, C., Hulme, C., Hall, P. & Wright, J. 2015. *Cost Effectiveness Modelling for Health Technology Assessment: A Practical Course*, Springer.

Edwards, D., Elliott, A., Hislop, M., Martin, S., Morris, J., O'Brien, L., Peace, A., Sarajevs, V., Serrand, M. & Valatin, G. 2009. A valuation of the economic and social contribution of forestry for people in Scotland. Edinburgh: Forestry Commission.

Elizalde, A. 2016. NIHR WIAT project-Summary prepared for the project meeting. Glasgow: University of Glasgow.

Elliott, R. & Payne, K. 2005. *Essentials of economic evaluation in healthcare*, Pharmaceutical Press.

Evers, S., Aarts, M.-J. J. & Alayli-Goebbels, A. F. 2015. *Measurement challenges in the economic evaluation of public health interventions,* McGraw-Hill House United Kingdom, Open University Press - McGraw-Hil.

Fan, Y., Das, K. V. & Chen, Q. 2011. Neighborhood green, social support, physical activity, and stress: assessing the cumulative impact. *Health Place,* 17**,** 1202-11.

FCS 2005. WIAT-Woods In and Around Towns. *In:* Executive, S. (ed.) *Policy.* Edinburgh: Forestry Commission Scotland.

FCS 2008. WIAT Woods In and Around Towns: Phase II. *In:* Governement, T. S. (ed.) *Policy*. Edinburgh.
FCS 2011a. Woods In and Around Towns- The story so far. A report on achievements and developments to date. *In:* Scotland, F. C. (ed.) *WIAT*. Edinburgh.
FCS 2011b. Woods in and around towns Phase III next steps. *Policy*. Scotland: Forestry Commission Scotland.
FCS 2015. A Strategic Framework Woods In and Around Towns (WIAT) Programme 2015-2020: Phase 4 sustaining delivery. *In:* Government, T. S. (ed.). Forestry Commission Scotland.
Federowicz, M. H., Grossman, M. N., Hayes, B. J. & Riggs, J. 2010. A Tutorial on Activity-Based Costing of Electronic Health Records. *Quality Management in Healthcare,* 19, 86-89 10.1097/QMH.0b013e3181ccbd71.
Feng, Y., Parkin, D. & Devlin, N. 2014. Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Quality of Life Research,* 23, 977-989.
Fenwick, E. & Byford, S. 2005. A guide to cost-effectiveness acceptability curves. *The British Journal of Psychiatry,* 187, 106-108.
Fenwick, E., Marshall, D. A., Levy, A. R. & Nichol, G. 2006. Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. *BMC Health Services Research,* 6, 1-8.
Fenwick, E., O'Brien, B. & Briggs, A. 2004. Cost-effectiveness acceptability curves: facts, fallacies and frequently asked questions. *Health Economics,* 13.
Fifer, S., Rose, J. & Greaves, S. 2014. Hypothetical bias in Stated Choice Experiments: Is it a problem? And if so, how do we deal with it? *Transportation research part A: policy and practice,* 61, 164-177.
Flynn, T., Bilger, M., Malhotra, C. & Finkelstein, E. 2016. Are Efficient Designs Used in Discrete Choice Experiments Too Difficult for Some Respondents? A Case Study Eliciting Preferences for End-of-Life Care. *PharmacoEconomics,* 1-12.
Flynn, T. N. 2016. Where Next for Discrete Choice Health Valuation Exercises? *Social Science Research Network (SSRN 2810444)*.
Flynn, T. N., Louviere, J. J., Peters, T. J. & Coast, J. 2007. Best–worst scaling: what it can do for health care research and how to do it. *Journal of health economics,* 26, 171-189.
Fricke, H. 2015. Identification based on Difference-in-Differences Approaches with Multiple Treatments. University of St. Gallen, School of Economics and Political Science.

Frumkin, H. 2001. Beyond toxicity: human health and the natural environment. *American journal of preventive medicine,* 20, 234-240.

Fujiwara, D. & Campbell, R. 2011. *Valuation Techniques for Social Cost-Benefit Analysis: Stated preference, revealed preference and subjective well-being approaches: a discussion of the current issues*, HM Treasury.

Gafni, A. 2006. Economic Evaluation of Health-care Programmes: Is CEA Better than CBA? *Environmental and Resource Economics*, 407.

Gasparrini, A. & Lopez Bernal, J. 2015. Commentary: On the use of quasi-experimental designs in public health evaluation. *International Journal of Epidemiology*.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B. & Vermeersch, C. M. 2011. *Impact evaluation in practice*, World Bank Publications.

Gidlöf-Gunnarsson, A. & Öhrström, E. 2007. Noise and well-being in urban residential environments: The potential role of perceived availability to nearby green areas. *Landscape and Urban Planning,* 83, 115-126.

Gidlow, C. J., Randall, J., Gillman, J., Smith, G. R. & Jones, M. V. 2016. Natural environments and chronic stress measured by hair cortisol. *Landscape and Urban Planning,* 148, 61-67.

Giles-Corti, B., Broomhall, M. H., Knuiman, M., Collins, C., Douglas, K., Ng, K., Lange, A. & Donovan, R. J. 2005. Increasing walking: how important is distance to, attractiveness, and size of public open space? *American journal of preventive medicine,* 28, 169-176.

Giles, A. 2015. Cost–benefit analysis: a tool that is both useful and influential? The Tools of Policy Formulation Actors, Capacities, Venues and Effects. Cheltenham, UK: 'Edward Elgar Publishing, Inc.'.

Giuseppe, C., Sabine, P., Ylenia, P., Stefano, M., Massimiliano, S. & Gabriella, B. 2012. Contact with Nature and Children's Wellbeing in Educational Settings. *Journal of Social Sciences,* 8, 304.

Glick, H. A., Doshi, J. A., Sonnad, S. S. & Polsky, D. 2015. *Economic evaluation in clinical trials,* OUP Oxford.

Goetz, C., Coste, J., Lemetayer, F., Rat, A.-C., Montel, S., Recchia, S., Debouverie, M., Pouchot, J., Spitz, E. & Guillemin, F. 2013. Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. *Journal of Clinical Epidemiology,* 66, 710-718.

Gowdy, J. M. 2007. Toward an experimental foundation for benefit-cost analysis. *Ecological Economics,* 63, 649-655.

Grabich, S. C., Robinson, W. R., Engel, S. M., Konrad, C. E., Richardson, D. B. & Horney, J. A. 2015. County-level hurricane exposure and birth rates: application of difference-in-differences analysis for confounding control. *Emerging Themes in Epidemiology,* 12, 1-8.
Grahn, P. & Stigsdotter, U. K. 2010. The relation between perceived sensory dimensions of urban green space and stress restoration. *Landscape and Urban Planning,* 94, 264-275.
Gray, A. M., Clarke, P. M., Wolstenholme, J. L. & Wordsworth, S. 2010. *Applied methods of cost-effectiveness analysis in healthcare*, OUP Oxford.
Gray, A. M. & Wilkinson, T. 2016. Economic evaluation of healthcare interventions: old and new directions. *Oxford Review of Economic Policy,* 32, 102-121.
Greco, G., Lorgelly, P. & Yamabhai, I. 2016. Outcomes in Economic Evaluations of Public Health Interventions in Low- and Middle-Income Countries: Health, Capabilities and Subjective Wellbeing. *Health Economics,* 25, 83-94.
Green, C. & Gerard, K. 2009. Exploring the social value of health-care interventions: a stated preference discrete choice experiment. *Health Economics,* 18, 951-976.
Greenspace Scotland 2008. Greenspace and quality of life: a critical literature review. Scotland.
Greiner, R., Bliemer, M. & Ballweg, J. 2014. Design considerations of a choice experiment to estimate likely participation by north Australian pastoralists in contractual biodiversity conservation. *Journal of Choice Modelling,* 10, 34-45.
Griffin, S., Rice, N. & Sculpher, M. 2010. Economic evaluation of public health interventions. *Evidence-based Public Health: Effectiveness and Efficiency,* 111.
Groenewegen, P. P., van den Berg, A. E., de Vries, S. & Verheij, R. A. 2006. Vitamin G: effects of green space on health, well-being, and social safety. *BMC Public Health,* 6, 149.
Groenewegen, P. P., van den Berg, A. E., Maas, J., Verheij, R. A. & de Vries, S. 2012. Is a Green Residential Environment Better for Health? If So, Why? *Annals of the Association of American Geographers,* 102, 996-1003.
Guttmann, R., Castle, R. & Fiebig, D. G. 2009. Use of discrete choice experiments in health economics: An update of the literature. *CHERE WORKING PAPER 2009/2*
Gyrd-Hansen, D. 2005. Willingness to pay for a QALY. *Pharmacoeconomics,* 23, 423-432.
Gyrd-Hansen, D. 2013. Using the Stated Preference Technique for Eliciting Valuations: The Role of the Payment Vehicle. *PharmacoEconomics,* 31, 853-861.

Gyrd-Hansen, D. & Skjoldborg, U. S. 2008. The price proxy in discrete choice experiments: Issues of relevance for future research. *Using discrete choice experiments to value health and health care*. Springer.

Hailey, D., Roine, R. & Ohinmaa, A. 2002. Systematic review of evidence for the benefits of telemedicine. *Journal of telemedicine and telecare,* 8 Suppl 1, 1-30.

Hall, J., Viney, R., Haas, M. & Louviere, J. 2004. Using stated preference discrete choice modeling to evaluate health care programs. *Journal of Business Research,* 57, 1026-1032.

Hamilton, L. 2012. *Statistics with Stata: version 12*, Cengage Learning.

Hanley, N., Mourato, S. & Wright, R. E. 2001. Choice Modelling Approaches: A Superior Alternative for Environmental Valuatioin? *Journal of Economic Surveys,* 15, 435-462.

Hanley, N., Wright, R. & Adamowicz, V. 1998. Using Choice Experiments to Value the Environment. *Environmental and Resource Economics,* 11, 413-428.

Hanley, N., Wright, R. E. & Alvarez-Farizo, B. 2006. Estimating the economic value of improvements in river ecology using choice experiments: an application to the water framework directive. *Journal of Environmental Management,* 78, 183-193.

Hansen, B. O., Hougaard, J. L., Keiding, H. & Osterdal, L. P. 2004. On the possibility of a bridge between CBA and CEA: comments on a paper by Dolan and Edlin. *J Health Econ,* 23, 887-98.

Hanski, I., von Hertzen, L., Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., Karisola, P., Auvinen, P., Paulin, L., Mäkelä, M. J., Vartiainen, E., Kosunen, T. U., Alenius, H. & Haahtela, T. 2012. Environmental biodiversity, human microbiota, and allergy are interrelated. *Proceedings of the National Academy of Sciences,* 109, 8334-8339.

Hansmann, R., Hug, S.-M. & Seeland, K. 2007. Restoration and stress relief through physical activities in forests and parks. *Urban Forestry & Urban Greening,* 6, 213-225.

Härkänen, T., Maljanen, T., Lindfors, O., Virtala, E. & Knekt, P. 2013. Confounding and missing data in cost-effectiveness analysis: comparing different methods. *Health Economics Review,* 3, 8.

Hartig, T., Evans, G. W., Jamner, L. D., Davis, D. S. & Gärling, T. 2003. Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology,* 23, 109-123.

Hartig, T., Mitchell, R., de Vries, S. & Frumkin, H. 2014. Nature and Health. *Annual review of public health*.

HAS 2012. Choices in Methods for Economic Evaluation. *A methodological guide*. Haute Autorite de Sante.

Hasan-Basri, B. & Karim, M. Z. A. 2013. The Effects of Coding on the Analysis of Consumer Choices of Public Parks. *World Applied Sciences Journal,* 22, 500-505.

Hauber, A. B., González, J. M., Groothuis-Oudshoorn, C. G., Prior, T., Marshall, D. A., Cunningham, C., IJzerman, M. J. & Bridges, J. F. 2016. Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. *Value in Health*.

Healey, A. & Chisholm, D. 1999. Willingness to pay as a measure of the benefits of mental health care. *The journal of mental health policy and economics,* 2, 55-58.

Health Council of the Netherlands 2004. Nature and Health. The influence of nature on social, psychological and physical well-being. The Hague: Health Council of the Netherlands and RMNO.

Hensher, D. & Greene, W. 2003. The Mixed Logit model: The state of practice. *Transportation,* 30, 133-176.

Hensher, D., Rose, J. & Greene, W. 2015. *Applied choice analysis analysis,* Cambridge, Cambridge University Press.

Herzele, A. & Vries, S. 2011. Linking green space to health: a comparative study of two urban neighbourhoods in Ghent, Belgium. *Population and Environment,* 34, 171-193.

Herzog, T. R., Black, A. M., Fountaine, K. A. & Knotts, D. J. 1997. Reflection and attentional recovery as distinctive benefits of restorative environments. *Journal of Environmental Psychology,* 17, 165-170.

Hiligsmann, M., Dellaert, B. G., Dirksen, C. D., van der Weijden, T., Goemaere, S., Reginster, J. Y., Watson, V. & Boonen, A. 2014. Patients' preferences for osteoporosis drug treatment: a discrete-choice experiment. *Arthritis Res Ther,* 16, R36.

Hiligsmann, M., van Durme, C., Geusens, P., Dellaert, B. G., Dirksen, C. D., van der Weijden, T., Reginster, J.-Y. & Boonen, A. 2013. Nominal group technique to select attributes for discrete choice experiments: an example for drug treatment choice in osteoporosis. *Patient preference and adherence,* 7, 133.

Hillsdon, M., Panter, J., Foster, C. & Jones, A. 2006. The relationship between access and quality of urban green space with population physical activity. *Public Health,* 120, 1127-32.

Hoch, J. S. & Dewa, C. S. 2005. An introduction to economic evaluation: what's in a name? *The Canadian Journal of Psychiatry,* 50, 159-166.

Hoefman, R. J., van Exel, J., Rose, J. M., van de Wetering, E. J. & Brouwer, W. B. F. 2014. A Discrete Choice Experiment to Obtain a Tariff for Valuing Informal Care Situations Measured with the CarerQol Instrument. *Medical Decision Making,* 34, 84-96.

Hole, A. R. 2015. DCREATE: Stata module to create efficient designs for discrete choice experiments. *Statistical Software Components*.
Homik, J. E. & Suarez-Almazor, M. 2004. An economic approach to health care. *Best practice & research.Clinical rheumatology,* 18, 203-218.
Hordyk, S. R., Hanley, J. & Richard, É. 2015. " Nature is there; its free": Urban greenspace and the social determinants of health of immigrant families. *Health & place,* 34, 74-82.
Horsman, J., Furlong, W., Feeny, D. & Torrance, G. 2003. The Health Utilities Index (HUI®): concepts, measurement properties and applications. *Health and Quality of Life Outcomes,* 1, 54.
Howard, K. 2009. *Quality-Adjusted Life Years (QALYs). Encyclopedia of Medical Decision Making. SAGE Publications, Inc,* Thousand Oaks, CA, SAGE Publications, Inc.
Hoyos, D. 2010. The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics,* 69, 1595-1603.
Huber, J. & Zwerina, K. 1996. The importance of utility balance in efficient choice designs. *Journal of Marketing research*, 307-317.
Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., Leonard, B., Lorig, K., Loureiro, M. I. & van der Meer, J. W. 2011. How should we define health? *BMJ: British Medical Journal,* 343.
Hurley, J. 2014. Welfarism and Extra-Welfarism A2 - Culyer, Anthony J. *Encyclopedia of Health Economics.* San Diego: Elsevier.
Irvine, K. N., Warber, S. L., Devine-Wright, P. & Gaston, K. J. 2013. Understanding urban green space as a health resource: a qualitative comparison of visit motivation and derived effects among park users in Sheffield, UK. *Int J Environ Res Public Health,* 10, 417-42.
Jacobs, J. C. & Barnett, P. G. 2016. Emergent Challenges in Determining Costs for Economic Evaluations. *PharmacoEconomics,* 1-11.
Johannesson, M. 1995. The relationship between cost-effectiveness analysis and cost-benefit analysis. *Social Science & Medicine,* 41, 483-489.
John, H. 2011. Benefits of Green Space – Recent Research. Chantilly, USA: Environmental Health Research Foundation.
Johnson, F. R., Kanninen, B., Bingham, M. & Özdemir, S. 2007. Experimental design for stated-choice studies. *Valuing environmental amenities using stated choice studies.* Springer.
Johnson, R. F., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., Bresnahan, B. W., Kanninen, B. & Bridges, J. F. P. 2013. Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental

Design Good Research Practices Task Force. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research,* 16**,** 3-13.
Kaczynski, A. T. & Henderson, K. A. 2007. Environmental Correlates of Physical Activity: A Review of Evidence about Parks and Recreation. *Leisure Sciences,* 29**,** 315-354.
Kahneman, D. & Knetsch, J. L. 1992. Valuing public goods: the purchase of moral satisfaction. *Journal of environmental economics and management,* 22**,** 57-70.
Kanninen, B. J. 2002. Optimal Design for Multinomial Choice Experiments. *Journal of Marketing Research,* 39**,** 214-227.
Kaplan, R. & Kaplan, S. 1989. *The experience of nature: A psychological perspective*, CUP Archive.
Kehlbacher, A., Balcombe, K. & Bennett, R. 2013. Stated Attribute Non-attendance in Successive Choice Experiments. *Journal of Agricultural Economics,* 64**,** 693-706.
Kelly, M., Morgan, A., Ellis, S., Younger, T., Huntley, J. & Swann, C. 2010. Evidence based public health: A review of the experience of the National Institute of Health and Clinical Excellence (NICE) of developing public health guidance in England. *Social Science & Medicine,* 71**,** 1056-1062.
Kelly, M. P., McDaid, D., Ludbrook, A. & Powell, J. 2005. *Economic appraisal of public health interventions,* London, Health Development Agency
Kelly, M. P., Stewart, E., Morgan, A., Killoran, A., Fischer, A. & Threlfall, A. 2009. A conceptual framework for public health: NICE's emerging approach. *Public Health,* 123.
Kenkel, D. 2006. WTP- and QALY-Based Approaches to Valuing Health for Policy: Common Ground and Disputed Territory. *Environmental & Resource Economics,* 34**,** 419-437.
Kenkel, D. & Suhrcke, M. 2011. Economic evaluation of the social determinants of health, an overview of conceptual and practical issues. *Copenhagen, Denmark: World Health Organization Regional Office for Europe*.
Kerr, G. N. & Sharp, B. M. Efficiency benefits of choice model experimental design updating: a case study. 2009 Conference (53rd), February 11-13, 2009, Cairns, Australia, 2009. Australian Agricultural and Resource Economics Society.
Khandker, S. R., Koolwal, G. B. & Samad, H. A. 2010. *Handbook on impact evaluation: quantitative methods and practices*, World Bank Publications.
Kim, J. & Kaplan, R. 2004. Physical and psychological factors in sense of community new urbanist Kentlands and nearby Orchard Village. *Environment and behavior,* 36**,** 313-340.

Kjær, T. 2005. *A review of the discrete choice experiment-with emphasis on its application in health care*, Syddansk Universitet Denmark.
Kjær, T., Gyrd-Hansen, D. & Willaing, I. 2006. Investigating patients' preferences for cardiac rehabilitation in Denmark. *International Journal of Technology Assessment in Health Care,* 22, 211-218.
Kløjgaard, M. E., Bech, M. & Søgaard, R. 2012. Designing a Stated Choice Experiment: The Value of a Qualitative Process. *Journal of Choice Modelling,* 5, 1-18.
Knez, I. 2005. Attachment and identity as related to a place and its perceived climate. *Journal of Environmental Psychology,* 25, 207-218.
Kobelt, G. 2013. *Health economics: an introduction to economic evaluation*, Office of Health Economics.
Krabbe, P. & Weijnen, T. 2003. Guidelines for analysing and reporting EQ-5D outcomes. *The measurement and valuation of health status using EQ-5D: A European perspective*. Springer.
Kragt, M. & Bennett, J. 2012. Attribute Framing in Choice Experiments: How Do Attribute Level Descriptions Affect Value Estimates? *Environmental and Resource Economics,* 51, 43-59.
Kragt, M. E. 2013. Evidence-based Research in Environmental Choice Experiments. *Working Paper 1310.* Australia: School of Agricultural and Resource Economics, University of Western Australia.
Kreye, M. M., Adams, D. C. & Escobedo, F. J. 2014. The Value of Forest Conservation for Water Quality Protection. *Forests,* 5, 862-884.
Krupnick, A. & Adamowicz, W. 2007. Supporting Questions in Stated-Choice Studies. *Valuing environmental amenities using stated choice studies*. Springer.
Kuo, F. E. 2001. Coping with Poverty: Impacts of Environment and Attention in the Inner City. *Environment and Behavior,* 33, 5-34.
Kuo, M. 2015. How might contact with nature promote human health? Promising mechanisms and a possible central pathway. *Frontiers in psychology,* 6.
Lachowycz, K. & Jones, A. 2011. Greenspace and obesity: a systematic review of the evidence. *Obesity reviews,* 12, e183-e189.
Lachowycz, K. & Jones, A. P. 2012. Towards a better understanding of the relationship between greenspace and health: Development of a theoretical framework. *Landscape and Urban Planning*.
Lachowycz, K., Jones, A. P., Page, A. S., Wheeler, B. W. & Cooper, A. R. 2013. What can global positioning systems tell us about the contribution of different types of urban greenspace to children's physical activity? *Health & place,* 18, 586-594.

Lafortezza, R., Carrus, G., Sanesi, G. & Davies, C. 2009. Benefits and well-being perceived by people visiting green spaces in periods of heat stress. *Urban Forestry & Urban Greening,* 8, 97-108.
Lagarde, M. 2013. Investigating attribute non-attendance and its consequences in choice experiments with latent class models. *Health Economics,* 22, 554-567.
Lancaster, K. J. 1966. A new approach to consumer theory. *The journal of political economy,* 74, 132-157.
Lancsar, E., Fiebig, D. G. & Hole, A. R. 2017. Discrete Choice Experiments: A Guide to Model Specification, Estimation and Software. *PharmacoEconomics,* 1-20.
Lancsar, E. & Louviere, J. 2008. Conducting Discrete Choice Experiments to Inform Healthcare Decision Making. *PharmacoEconomics,* 26, 661-677.
Lancsar, E., Wildman, J., Donaldson, C., Ryan, M. & Baker, R. 2011. Deriving distributional weights for QALYs through discrete choice experiments. *Journal of Health Economics,* 30, 466-478.
Lawson, K., Mason, H., McIntosh, E. & Donaldson, C. 2014. Priority Setting in Public Health. *In:* Culyer, A. J. (ed.) *Encyclopedia of Health Economics.* San Diego: Elsevier.
Lee, A. C. K. & Maheswaran, R. 2011. The health benefits of urban green spaces: a review of the evidence. *Journal of Public Health,* 33, 212-222.
Lee, J., Park, B. J., Tsunetsugu, Y., Ohira, T., Kagawa, T. & Miyazaki, Y. 2011. Effect of forest bathing on physiological and psychological responses in young Japanese male subjects. *Public Health,* 125, 93-100.
Lee, J., Tsunetsugu, Y., Takayama, N., Park, B.-J., Li, Q., Song, C., Komatsu, M., Ikei, H., Tyrväinen, L. & Kagawa, T. 2014. Influence of forest therapy on cardiovascular relaxation in young adults. *Evidence-Based Complementary and Alternative Medicine,* 2014.
Lessard, C. 2007. Complexity and reflexivity: two important issues for economic evaluation in health care. *Social science & medicine (1982),* 64, 1754-1765.
Lewis, D. 2004. Economic Evaluation of Health Care Programs. *Australian Economic Review,* 37, 350-358.
Li, Q., Morimoto, K., Kobayashi, M., Inagaki, H., Katsumata, M., Hirata, Y., Hirata, K., Suzuki, H., Li, Y. & Wakayama, Y. 2008. Visiting a forest, but not a city, increases human natural killer activity and expression of anti-cancer proteins. *International journal of immunopathology and pharmacology,* 21, 117-127.
Lin, C. & Hsu, S. 2014. Constructing “Behavioral” Comparison Groups: A Difference-in-Difference Analysis of the Effect of Copayment

Based on the Patient's Price Elasticity. *Evaluation & the Health Professions,* 37, 434-456.
Liu, C. & Li, X. 2012. Carbon storage and sequestration by urban forests in Shenyang, China. *Urban Forestry & Urban Greening,* 11, 121-128.
Lloyd, A. J. 2003. Threats to the estimation of benefit: are preference elicitation methods accurate? *Health economics,* 12, 393-402.
Londoño, C. & Ando, A. W. 2013. Valuing preferences over stormwater management outcomes including improved hydrologic function. *Water Resources Research*.
Long, J. S. & Freese, J. 2014. *Regression models for categorical dependent variables using Stata*, Stata press.
Lorgelly, P. K., Lawson, K. D., Fenwick, E. A. & Briggs, A. H. 2010. Outcome measurement in economic evaluations of public health interventions: a role for the capability approach? *Int J Environ Res Public Health,* 7, 2274-89.
Louviere, J. J., Hensher, D. A. & Swait, J. D. 2000. *Stated choice methods: analysis and applications*, Cambridge University Press.
Louviere, J. J., Islam, T., Wasi, N., Street, D. & Burgess, L. 2008. Designing Discrete Choice Experiments: Do Optimal Designs Come at a Price? *Journal of Consumer Research,* 35, 360-375.
Luce, R. D. 1959. *Individual choice behavior: A theoretical analysis*, Wiley New York NY.
Luyten, J., Naci, H. & Knapp, M. 2016. Economic evaluation of mental health interventions: an introduction to cost-utility analysis. *Evidence Based Mental Health*.
Lynch, S. V., Wood, R. A., Boushey, H., Bacharier, L. B., Bloomberg, G. R., Kattan, M., O'Connor, G. T., Sandel, M. T., Calatroni, A. & Matsui, E. 2014. Effects of early-life exposure to allergens and bacteria on recurrent wheeze and atopy in urban children. *Journal of Allergy and Clinical Immunology,* 134, 593-601. e12.
Maas, J., Spreeuwenberg, P., Van Winsum-Westra, M., Verheij, R. A., de Vries, S. & Groenewegen, P. P. 2009a. Is green space in the living environment associated with people's feelings of social safety? *Environment and planning. A,* 41, 1763.
Maas, J., van Dillen, S. M., Verheij, R. A. & Groenewegen, P. P. 2009b. Social contacts as a possible mechanism behind the relation between green space and health. *Health Place,* 15, 586-95.
Maas, J., Verheij, R. A., Spreeuwenberg, P. & Groenewegen, P. P. 2008. Physical activity as a possible mechanism behind the relationship between green space and health: a multilevel analysis. *BMC Public Health,* 8, 206.
Madureira, H., Nunes, F., Oliveira, J. V., Cormier, L. & Madureira, T. 2015. Urban residents' beliefs concerning green space benefits in

four cities in France and Portugal. *Urban Forestry & Urban Greening,* 14, 56-64.
Maller, C., Townsend, M., Pryor, A., Brown, P. & St Leger, L. 2006. Healthy nature healthy people: 'contact with nature' as an upstream health promotion intervention for populations. *Health Promotion International,* 21, 45-54.
Manca, A., Hawkins, N. & Sculpher, M. J. 2005. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Economics,* 14, 487-496.
Marselle, M., Warber, S. & Irvine, K. 2012. P04.54. Natural health service: enhancing wellbeing with group walks in green spaces. *BMC Complementary and Alternative Medicine,* 12, P324.
Marsh, K., Ijzerman, M., Thokala, P., Baltussen, R., Boysen, M., Kaló, Z., Lönngren, T., Mussen, F., Peacock, S., Watkins, J. & Devlin, N. 2016. Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value in Health,* 19, 125-137.
Marsh, K., Lanitis, T., Neasham, D., Orfanos, P. & Caro, J. 2014. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. *Pharmacoeconomics,* 32, 345-365.
Marsh, K., Phillips, C. J., Fordham, R., Bertranou, E. & Hale, J. 2012. Estimating cost-effectiveness in public health: a summary of modelling and valuation methods. *Health economics review,* 2, 1-6.
Marshall, D., Bridges, J. P., Hauber, B., Cameron, R., Donnalley, L., Fyie, K. & Reed Johnson, F. 2010. Conjoint Analysis Applications in Health — How are Studies being Designed and Reported? *The Patient: Patient-Centered Outcomes Research,* 3, 249-256.
Mason, H., van Exel, J., Baker, R., Brouwer, W., Donaldson, C., Donaldson, C., Baker, R., Mason, H., Pennington, M., Bell, S., Jones-Lee, M., Wildman, J., Lancsar, E., Robinson, A., Bacon, P., Olsen, J. A., Gyrd-Hansen, D., Kjaer, T., Beck, M., Nielsen, J. S., Persson, U., Bergman, A., Protière, C., Moatti, J. P., Luchini, S., Pinto Prades, J. L., Mataria, A., Khatib, R., Jaralla, Y., Brouwer, W., van Exel, J., Topór-Madry, R., Kozierkiewicz, A., Poznanski, D., Kocot, E., Gulácsi, L., Péntek, M., Kharroubi, S., Manca, A. & Shackley, P. 2016. From representing views to representativeness of views: Illustrating a new (Q2S) approach in the context of health care priority setting in nine European countries. *Social Science & Medicine,* 166, 205-213.
Mason, M. 2009. Walk on the wild side. *Nursing standard (Royal College of Nursing (Great Britain): 1987),* 23, 20.

Matthews, J., Altman, D. G., Campbell, M. & Royston, P. 1990. Analysis of serial measurements in medical research. *Bmj,* 300, 230-235.
Mauskopf, J. A., Paul, J. E., Grant, D. M. & Stergachis, A. 1998. The Role of Cost-Consequence Analysis in Healthcare Decision-Making. *PharmacoEconomics,* 13, 277-288.
Mays, G. P. & Mamaril, C. 2015. Fundamentals of Economic Evaluation for Public Health. *Health Management and Policy Presentations,* Paper 109.
McCabe, C., Claxton, K. & Culyer, A. J. 2008. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics,* 26, 733-744.
McCrone, P., Patel, A., Knapp, M., Schene, A., Koeter, M., Ammadeo, F., Ruggeri, M., Giessler, A., Puschner, B. & Thornicroft, G. 2009. A comparison of SF-6D and EQ-5D utility scores in a study of patients with schizophrenia. *Journal of mental health policy and economics,* 12, 27-31.
McCrone, P. R., Dhanasiri, S., Patel, A., Knapp, M. & Lawton-Smith, S. 2008. *Paying the price: the cost of mental health care in England to 2026*, King's Fund.
McEachan, R., Prady, S., Smith, G., Fairley, L., Cabieses, B., Gidlow, C., Wright, J., Dadvand, P., van Gent, D. & Nieuwenhuijsen, M. 2015. The association between green space and depressive symptoms in pregnant women: moderating roles of socioeconomic status and physical activity. *Journal of epidemiology and community health,* jech-2015-205954.
McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers of econometrics*, 105-142.
McFadden, D. & Train, K. 2000. Mixed MNL models for discrete response. *Journal of applied Econometrics,* 15, 447-470.
McFarland, A. 2014. Economic evaluation of interventions in health care. *Nursing Standard,* 29, 49-58.
McIntosh, E. 2006. Using Discrete Choice Experiments within a Cost-Benefit Analysis Framework: Some Considerations. *PharmacoEconomics,* 24, 855.
McIntosh, E. 2010. A practical guide to reporting and presenting stated preference. *Applied Methods of Cost-Benefit Analysis in Health Care,* 4, 231.
McIntosh, E., Bond, L., Donaldson, C., Lawson, K. & Mason, H. 2014. Economics of population health. *In:* Guest, G. & Namey, E. (eds.) *Public Health Research Methods.* Thousand Oak California: SAGE Publications, Inc.
McIntosh, E., Clarke, P., Frew, E. & Louviere, J. 2010. *Applied Methods of Cost-Benefit Analysis in Health Care,* OUP Oxford.

McIntosh, E., Donaldson, C. & Mandy, R. 1999. Recent Advances in the Methods of Cost-Benefit Analysis in Healthcare: Matching the Art to the Science. *PharmacoEconomics,* 15, 357-367.
McIntosh, E. & Luengo-Fernandez, R. 2006a. Economic evaluation. Part 1: Introduction to the concepts of economic evaluation in health care. *The journal of family planning and reproductive health care / Faculty of Family Planning & Reproductive Health Care, Royal College of Obstetricians & Gynaecologists,* 32, 107-112.
McIntosh, E. & Luengo-Fernandez, R. 2006b. Economic evaluation. Part 2: frameworks for combining costs and benefits in health care. *The journal of family planning and reproductive health care / Faculty of Family Planning & Reproductive Health Care, Royal College of Obstetricians & Gynaecologists,* 32, 176-180.
Mengoni, A., Seghieri, C. & Nuti, S. 2013. The application of discrete choice experiments in health economics: a systematic review of the literature. *Working paper n. 01/2013* Istituto di Management-Scuola Superiore Sant'Anna di Pisa.
Mentzakis, E., Ryan, M. & McNamee, P. 2011. Using discrete choice experiments to value informal care tasks: exploring preference heterogeneity. *Health Economics,* 20, 930-944.
Mercer, E. & Snook, A. 2005. Analyzing Ex-Ante Agroforestry Adoption Decisions with Attribute-Based Choice Experiments. *Valuing Agroforestry Systems*. Springer Netherlands.
Miller, P. 2009. *Health economic evaluation,* East Midlands, The NIHR RDS for the East Midlands/Yorkshire and the Humber.
Mills, T. C. & Patterson, K. D. 2011. *Palgrave handbook of econometrics,* New York, Palgrave Macmillan.
Mitchell, R. 2012. Is physical activity in natural environments better for mental health than physical activity in other environments? *Social Science & Medicine*.
Mitchell, R. & Popham, F. 2007. Greenspace, urbanity and health: relationships in England. *Journal of Epidemiology and Community Health,* 61, 681-683.
Mitchell, R. & Popham, F. 2008. Effect of exposure to natural environment on health inequalities: an observational population study. *The Lancet,* 372, 1655-1660.
Mogyorosy, Z. & Smith, P. 2005. The main methodological issues in costing health care services: a literature review.
Mora Villarrubia, R. & Reggio, I. 2012. Treatment effect identification using alternative parallel assumptions. Universidad Carlos III, Departamento de Economía.
Morgan, G. A., Gliner, J. A. & Harmon, R. J. 2000. Quasi-Experimental Designs. *Journal of the American Academy of Child & Adolescent Psychiatry,* 39, 794-796.

Morris, S., Devlin, N. & Parkin, D. 2012. *Economic analysis in health care*, Wiley.
Morris, S., Patel, N. V., Dobson, J., Featherstone, R. L., Richards, T., Luengo-Fernandez, R., Rothwell, P. M., Brown, M. M. & investigators, I. C. S. S. 2016. Cost-utility analysis of stenting versus endarterectomy in the International Carotid Stenting Study. *International Journal of Stroke*, 1747493016632237.
Morrison, M. D., Blamey, R. K. & Bennett, J. W. 2000. Minimising Payment Vehicle Bias in Contingent Valuation Studies. *Environmental and Resource Economics*, 16, 407-422.
NICE 2008. Guide to the methods of technology appraisal. London: National Institute for Clinical Excellence.
NICE 2012. Methods for the development of NICE public health guidance 3rd ed. London.
NICE 2013. Guide to the methods of technology appraisal *Process and Methods Guide*. The National Institute for Health and Care Excellence.
Nielsen, T. S. & Hansen, K. B. 2007. Do green areas affect health? Results from a Danish survey on the use of green areas and health indicators. *Health Place*, 13, 839-50.
Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M. & Dadvand, P. 2016. Fifty Shades of Green: Pathway to Healthy Urban Living. *Epidemiology (Cambridge, Mass.)*.
NIHR 2012. How effective is the Forestry Commission Scotland's woodland improvement programme - 'Woods In and Around Towns' (WIAT) - at improving psychological wellbeing in deprived communities? *Project Ref: 10/3005/18*. Southampton, UK: National Institute for Health Research
Nixon, R. M. & Thompson, S. G. 2005. Methods for incorporating covariate adjustment, subgroup analysis and between-centre differences into cost-effectiveness evaluations. *Health Economics*, 14, 1217-1229.
Nowak, D. J., Crane, D. E. & Stevens, J. C. 2006. Air pollution removal by urban trees and shrubs in the United States. *Urban forestry & urban greening*, 4, 115-123.
O'Brien, B. & Viramontes, J. L. 1993. Willingness-to-pay: A Valid and Reliable Measure of Health State Preference? : Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, Canada.
O’Neill, S., Kreif, N., Grieve, R., Sutton, M. & Sekhon, J. S. 2016. Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services and Outcomes Research Methodology*, 16, 1-21.

Oemar, M. & Janssen, B. 2014. EQ-5D-5L User guide. Basic information on how to use the EQ-5D-5L instrument. EuroQol Group.
Oliver, D., Foot, C. & Humphries, R. 2014. Making our health and care systems fit for an ageing population. *London: The King's Fund*.
Olsson, T. M. 2011. Comparing top-down and bottom-up costing approaches for economic evaluation within social welfare. *Eur J Health Econ,* 12**,** 445-53.
Oostenbrink, J. B., Koopmanschap, M. A. & Rutten, F. F. H. 2002. Standardisation of Costs: The Dutch Manual for Costing in Economic Evaluations. *PharmacoEconomics,* 20**,** 443-454.
Ord, K., Mitchell, R. & Pearce, J. 2013. Is level of neighbourhood green space associated with physical activity in green space? *International Journal of Behavioral Nutrition and Physical Activity,* 10**,** 127.
Orme, B. 1998. Sample size issues for conjoint analysis studies. *Sawthooth Software Research paper Series. Squim, WA, USA: Sawthooth Software Inc*.
Orme, B. 2006. Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research Madison. *Wisconsin: Research Publishers*.
Ottosson, J. & Grahn, P. 2005. A comparison of leisure time spent in a garden with leisure time spent indoors: On measures of restoration in residents in geriatric care. *Landscape Research,* 30**,** 23-55.
Owen, L., Morgan, A., Fischer, A., Ellis, S., Hoy, A. & Kelly, M. P. 2011. The cost-effectiveness of public health interventions. *Journal of Public Health***,** fdr075.
Park, A.-L. 2014. The impacts of intergenerational programmes on the physical health of older adults. *Journal of Aging Science,* 2**,** 1-5.
Parkin, D. & Devlin, N. 2006. Is there a case for using visual analogue scale valuations in cost-utility analysis? *Health economics,* 15**,** 653-664.
Parkin, D., Morris, S. & Devlin, N. 2015. Economic appraisal in public healthcare: assessing efficiency and equity. *In:* Detels, R., Gulliford, M., Karim, Q. A. & Tan, C. C. (eds.) *Oxford Textbook of Global Public Health.* Oxford, UK: Oxford University Press.
Parkin, D., Rice, N. & Devlin, N. 2010. Statistical analysis of EQ-5D profiles: does the use of value sets bias inference? *Medical Decision Making,* 30**,** 556-565.
Paulden, M., Galvanni, V., Chakraborty, S., Kudinga, B. & McCabe, C. 2016. *Discounting and the Evaluation of Health Care Programs* [Online]. Canada: Canadian Agency for Drugs and Technologies in Health. Available: https://www.cadth.ca/sites/default/files/pdf/CP0008_Economic_

Evaluation_Guidelines_Discount_Rate_Report.pdf [Accessed 06/04 2016].

Payne, K., McAllister, M. & Davies, L. M. 2013. Valuing the economic benefits of complex interventions: When maximising health is not sufficient. *Health Economics,* 22**,** 258-271.

Payne, K. & Thompson, A. 2013. Economics of Pharmacogenomics: Rethinking Beyond QALYs? *Current Pharmacogenomics and Personalized Medicine,* 11**,** 187-195.

Payne, K. & Thompson, A. 2015. Economic evaluations of complex interventions. *In:* Richards, D. & Hallberg, I. (eds.) *Complex Interventions in Health: An overview of research methods.* 1st ed. London: Routledge.

Pearce, D., Özdemiroğlu, E. & Britain, G. 2002. *Economic valuation with stated preference techniques: summary guide*, Department for Transport, Local Government and the Regions.

Pearce, J. R., Richardson, E. A., Mitchell, R. J. & Shortt, N. K. 2010. Environmental justice and health: the implications of the socio-spatial distribution of multiple environmental deprivation for health inequalities in the United Kingdom. *Transactions of the Institute of British Geographers,* 35**,** 522-539.

Pearce, N. & Greenland, S. 2014. Confounding and Interaction. *In:* Ahrens, W. & Pigeot, I. (eds.) *Handbook of Epidemiology.* 2nd ed. New York: Springer.

Peng Yu, A. 2009. *Cost Measurement Methods. Encyclopedia of Medical Decision Making. SAGE Publications, Inc,* Thousand Oaks, CA, SAGE Publications, Inc.

Perkins, C., Steinbach, R., Tompson, L., Green, J., Johnson, S., Grundy, C., Wilkinson, P. & Edwards, P. 2015. Cost-benefit analysis: methodological challenges of evaluating large-scale public health interventions and a worked example of the costs and benefits of part-night lighting.

Peschardt, K. K. & Stigsdotter, U. K. 2013. Associations between park characteristics and perceived restorativeness of small public urban green spaces. *Landscape and Urban Planning,* 112**,** 26-39.

Pfarr, C., Schmid, A. & Schneider, U. 2014. Using Discrete Choice Experiments to Understand Preferences in Health Care. *Health Care Provision and Patient Mobility: Health Integration in the European Union***,** 27-48.

PHE 2014. Local action on health inequalities: improving access to green spaces. *In:* Equity, U. I. o. H. (ed.) *Health Equity Evidence Review.* London: Public Health England.

Phelps, C. E. & Mushlin, A. I. 1991. On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *International journal of technology assessment in health care,* 7**,** 12-21.

Phillips, C. 2009. *What is cost-effectiveness?*, Whatisseries UK.
Phillips, C. & Thompson, G. 1998. *What is a QALY?*, Hayward Medical Communications.
Phillips, C. J., Fordham, R., Marsh, K., Bertranou, E., Davies, S., Hale, J., Kingsley, M., Parke, S., Porteous, C. & Rance, J. 2011. Exploring the role of economics in prioritization in public health: what do stakeholders think? *The European Journal of Public Health,* 21, 578-584.
Pinder, R., Kessel, A., Green, J. & Grundy, C. 2009. Exploring perceptions of health and the environment: a qualitative study of Thames Chase Community Forest. *Health Place,* 15, 349-56.
Polinder, S., Toet, H., Panneman, M. & Beeck, E. 2011. Methodological approaches for cost-effectiveness and cost-utility analysis of injury prevention measures. *World Health Organization, Regional Office for Europe, Copenhagen*.
Potestio, M. L., Patel, A. B., Powell, C. D., McNeil, D. A., Jacobson, R. D. & McLaren, L. 2009. Is there an association between spatial access to parks/green space and childhood overweight/obesity in Calgary, Canada? *International Journal of Behavioral Nutrition and Physical Activity,* 6, 77.
Raanaas, R. K., Patil, G. G. & Hartig, T. 2012. Health benefits of a view of nature through the window: a quasi-experimental study of patients in a residential rehabilitation center. *Clinical Rehabilitation,* 26, 21-32.
Rabarison, K. M., Bish, C. L., Massoudi, M. S. & Giles, W. H. 2015. Economic evaluation enhances public health decision making. *Frontiers in public health,* 3.
Ratcliffe, J. & Longworth, L. 2002. Investigating the structural reliability of a discrete choice experiment within health technology assessment. *International Journal of Technology Assessment in Health Care,* 18, 139-144.
Rayner, G. & Lang, T. 2012. *Ecological Public Health: Reshaping the Conditions for Good Health*, Routledge.
Rayner, G. & Lang, T. 2015. What is the point of public health in the 21st century? *Public Health,* 129, 1309-1313.
Reed Johnson, F. 2012. Why Not Real Economics? *PharmacoEconomics,* 30, 127-131.
Regier, D. A., Watson, V., Burnett, H. & Ungar, W. J. 2014. Task complexity and response certainty in discrete choice experiments: An application to drug treatments for juvenile idiopathic arthritis. *Journal of Behavioral and Experimental Economics,* 50, 40-49.
Remler, D. K. & Van Ryzin, G. G. 2010. *Research methods in practice: Strategies for description and causation,* Sage Publications.

Richards, D. A. & Hallberg, I. R. 2015. *Complex Interventions in Health: An overview of research methods*, Routledge.

Richardson, E., Pearce, J., Mitchell, R., Day, P. & Kingham, S. 2010. The association between green space and cause-specific mortality in urban New Zealand: an ecological analysis of green space utility. *BMC Public Health,* 10, 240.

Richardson, E. A. & Mitchell, R. 2010. Gender differences in relationships between urban green space and health in the United Kingdom. *Soc Sci Med,* 71, 568-75.

Richardson, E. A., Mitchell, R., Hartig, T., de Vries, S., Astell-Burt, T. & Frumkin, H. 2012. Green cities and health: a question of scale? *J Epidemiol Community Health,* 66, 160-5.

Richardson, E. A., Pearce, J., Mitchell, R. & Kingham, S. 2013. Role of physical activity in the relationship between urban green space and health. *Public Health*.

Richardson, G. & Manca, A. 2004. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. *Health economics,* 13, 1203-1210.

Richardson, J. 2009. Is the incorporation of equity considerations into economic evaluation really so simple? A comment on Cookson, Drummond and Weatherly. *Health Econ Policy Law,* 4, 247-54; discussion 261-3.

Richardson, J., McKie, J. & Olsen, J. A. 2005. *Welfarism or non-welfarism? Public preferences for willingness to pay versus health maximisation*, Monash University Centre for Health Economics.

Ride, J., Rowe, H., Wynter, K., Fisher, J. & Lorgelly, P. 2014. Protocol for economic evaluation alongside a cluster-randomised controlled trial of a psychoeducational intervention for the primary prevention of postnatal mental health problems in first-time mothers. *BMJ open,* 4, e006226.

Roe, J. & Aspinall, P. 2011. The restorative benefits of walking in urban and rural settings in adults with good and poor mental health. *Health & Place,* 17, 103-113.

Rose, J. & Bliemer, M. J. 2013. Sample size requirements for stated choice experiments. *Transportation,* 40, 1021-1041.

Rose, J. M. & Bliemer, M. C. 2008. Stated preference experimental design strategies. *Handbook of transport modelling*, 151-180.

Rose, J. M. & Bliemer, M. C. 2009. Constructing efficient stated choice experimental designs. *Transport Reviews,* 29, 587-617.

Ruger, J. P. & Ng, N. Y. 2014. Ethics and social value judgments in public health. *Ng NY, and Ruger JP Ethics and Social Value Judgments in Public Health. In: Anthony J. Culyer (ed.), Encyclopedia of Health Economics,* 1, 287-291.

Ryan, A. M., Burgess, J. F. & Dimick, J. B. 2015. Why We Should Not Be Indifferent to Specification Choices for Difference-in-Differences. *Health services research,* 50, 1211-1235.
Ryan, M. & Gerard, K. 2003. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Applied health economics and health policy,* 2, 55-64.
Ryan, M., Gerard, K. & Amaya-Amaya, M. 2008a. Discrete Choice Experiments in a Nutshell. *In:* Ryan, M., Gerard, K. & Amaya-Amaya, M. (eds.) *Using Discrete Choice Experiments to Value Health and Health Care*. Springer Netherlands.
Ryan, M., Gerard, K. & Currie, G. 2012a. Using discrete choice experiments in health economics. *The Elgar companion to health economics*.
Ryan, M., Gerard, K., Watson, V., Street, D. & Burgess, L. 2008b. Practical Issues in Conducting a Discrete Choice Experiment. *In:* Ryan, M., Gerard, K. & Amaya-Amaya, M. (eds.) *Using Discrete Choice Experiments to Value Health and Health Care*. Springer Netherlands.
Ryan, M., Kolstad, J., Rockers, P. & Dolea, C. 2012b. How to conduct a Discrete Choice Experiment for health workforce recruitment and retention in remote and rural areas: a user guide with case studies. *CapacityPlus. World Bank and World Health Organization*
Rychetnik, L., Frommer, M., Hawe, P. & Shiell, A. 2002. Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health,* 56, 119-127.
Salazar, I., Jackson, S., Shiell, A. & Rice, M. 2007. *Guide to economic evaluation in health promotion,* Washington, Pan American Health Organisation.
Salingaros, N. A. 2015. Biophilia & Healing Environments

Healthy Principles For Designing the Built World. New York: Terrapin Bright Green, LLC.
Salmon, P. 2001. Effects of physical exercise on anxiety, depression, and sensitivity to stress: A unifying theory. *Clinical Psychology Review,* 21, 33-61.
Saulle, R. & La Torre, G. 2012. Good quality and available urban green spaces as good quality, health and wellness for human life. *J Public Health (Oxf),* 34, 161-2.
Scarpa, R. & Rose, J. M. 2008. Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why*. *Australian journal of agricultural and resource economics,* 52, 253-282.
Schäfer, T., Gericke, C. & Busse, R. 2014. Health Services Research. *In:* Ahrens, W. & Pigeot, I. (eds.) *Handbook of Epidemiology*. Springer New York.

Schipperijn, J., Bentsen, P., Troelsen, J., Toftager, M. & Stigsdotter, U. K. 2013. Associations between physical activity and characteristics of urban green space. *Urban Forestry & Urban Greening,* 12**,** 109-116.
Schipperijn, J., Ekholm, O., Stigsdotter, U. K., Toftager, M., Bentsen, P., Kamper-Jørgensen, F. & Randrup, T. B. 2010. Factors influencing the use of green space: Results from a Danish national representative survey. *Landscape and Urban Planning,* 95**,** 130-137.
Schlander, M. 2010. Measures of efficiency in healthcare: QALMs about QALYs? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen,* 104**,** 214-226.
Schoen, C., Osborn, R., How, S. K., Doty, M. M. & Peugh, J. 2009. In chronic condition: experiences of patients with complex health care needs, in eight countries, 2008. *Health affairs,* 28**,** w1-w16.
Schoenfeld, T. J., Rada, P., Pieruzzini, P. R., Hsueh, B. & Gould, E. 2013. Physical Exercise Prevents Stress-Induced Activation of Granule Neurons and Enhances Local Inhibitory Mechanisms in the Dentate Gyrus. *The Journal of Neuroscience,* 33**,** 7770-7777.
Schwarz, N. & Strack, F. 1999. Reports of subjective well-being: Judgmental processes and their methodological implications. *Well-being: The foundations of hedonic psychology,* 7**,** 61-84.
Scottish Government 2011. Mid Term Evaluation of Scotland Rural Development Programme. *In:* Government, S. (ed.). Scotland.
Scottish Natural Heritage 2008. Health Impact Assessment of greenspace A Guide.
Sculpher, M. J., Claxton, K., Drummond, M. & McCabe, C. 2006. Whither trial-based economic evaluation for health care decision making? *Health economics,* 15**,** 677-687.
Sculpher, M. J. & Price, M. 2003. Measuring costs and consequences in economic evaluation in asthma. *Respiratory Medicine,* 97**,** 508-520.
Shadish, W. R., Cook, T. D. & Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference,* Wadsworth Cengage learning.
Shah, K., Mulhern, B., Longworth, L. & Janssen, B. 2016. Important Aspects of Health Not Captured by EQ-5D: Views of the UK General Public. *Office of Health Economics (OHE) Research,* Research Paper 16/06.
Shah, K. K., Cookson, R., Culyer, A. J. & Littlejohns, P. 2013. NICE's social value judgements about equity in health and health care. *Health Economics, Policy and Law,* 8**,** 145-165.

Shah, K. K., Tsuchiya, A. & Wailoo, A. J. 2015. Valuing health at the end of life: A stated preference discrete choice experiment. *Social Science & Medicine,* 124**,** 48-56.
Shearer, J. & Byford, S. 2015. The basics of economic evaluation in mental healthcare. *BJPsych Advances,* 21**,** 345-353.
Sheill, A. 2009. Still waiting for the great leap forward. *Health Economics, Policy and Law,* 4**,** 255-260.
Shiell, A. 2007. In search of social value. *International Journal of Public Health,* 52**,** 333-334.
Shiell, A., Hawe, P. & Gold, L. 2008. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ,* 336.
Silveirinha de Oliveira, E., Aspinall, P., Briggs, A., Cummins, S., Leyland, A. H., Mitchell, R., Roe, J. & Ward Thompson, C. 2013. How effective is the Forestry Commission Scotland's woodland improvement programme—'Woods In and Around Towns' (WIAT)—at improving psychological well-being in deprived urban communities? A quasi-experimental study. *BMJ Open,* 3.
Simoens, S. 2009. Health economic assessment: a methodological primer. *International journal of environmental research and public health,* 6**,** 2950-2966.
SMC. 2015. *A Guide To Quality Adjusted Life Years* [Online]. Scotland: Scottish Medicines Consortium. Available: http://www.scottishmedicines.org.uk/About_SMC/Policy_statements/A_Guide_to_Quality_Adjusted_Life_Years [Accessed 20th August 2015 2015].
Smith, R., Lorgelly, P., Al-Janabi, H., Venkatapuram, S. & Coast, J. 2012. The capability approach: an alternative evaluation paradigm for health economics? *In:* Jones, A. M. (ed.) *The elgar companion to health economics.* Edward Elgar
Smith, R. D. & Petticrew, M. 2010. Public health evaluation in the twenty-first century: time to see the wood as well as the trees. *J Public Health (Oxf),* 32.
Squires, H., Chilcott, J., Akehurst, R., Burr, J. & Kelly, M. P. 2016. A Framework for Developing the Structure of Public Health Economic Models. *Value in Health,* 19**,** 588-601.
StataCorp, L. 2013. Stata Statistical Software. *Release 13.* College Station, TX StataCorp LP.
Stigsdotter, U. K., Ekholm, O., Schipperijn, J., Toftager, M., Kamper-Jorgensen, F. & Randrup, T. B. 2010. Health promoting outdoor environments--associations between green space, and health, health-related quality of life and stress based on a Danish national representative survey. *Scand J Public Health,* 38**,** 411-7.

Stigsdotter, U. K. & Grahn, P. 2011. Stressed individuals' preferences for activities and environmental characteristics in green spaces. *Urban Forestry & Urban Greening,* 10, 295-304.

Stinnett, A. A. & Mullahy, J. 1998. Net Health Benefits: A new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making,* 18.

Stoto, M. & Cosler, L. 2005. Evaluation of public health interventions. *In:* Novick, L. F. & Mays, G. P. (eds.) *Public health administration: principles for population-based management*

Jones & Bartlett Learning.

Street, D., Burgess, L., Viney, R. & Louviere, J. 2008. Designing Discrete Choice Experiments for Health Care. *In:* Ryan, M., Gerard, K. & Amaya-Amaya, M. (eds.) *Using Discrete Choice Experiments to Value Health and Health Care.* Springer Netherlands.

Street, D. J. & Burgess, L. 2007. *The construction of optimal stated choice experiments: theory and methods,* John Wiley & Sons.

Street, D. J., Burgess, L. & Louviere, J. J. 2005. Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing,* 22, 459-470.

Sugiyama, T., Gunn, L. D., Christian, H., Francis, J., Foster, S., Hooper, P., Owen, N. & Giles-Corti, B. 2015. Quality of Public Open Spaces and Recreational Walking. *American Journal of Public Health,* 105, 2490-2495.

Sugiyama, T. & Thompson, C. W. 2008. Associations between characteristics of neighbourhood open space and older people's walking. *Urban forestry & urban greening,* 7, 41-51.

Susan, D. C. & Henk, S. 2012. Restorative Environments. *Handbook of Environmental and Conservation Psychology.* Oxford London: Oxford University Press.

Swanwick, C., Dunnett, N. & Woolley, H. 2003. Nature, role and value of green space in towns and cities: An overview. *Built environment,* 29, 94-106.

Takano, T., Nakamura, K. & Watanabe, M. 2002. Urban residential environments and senior citizens' longevity in megacity areas: the importance of walkable green spaces. *Journal of Epidemiology and Community Health,* 56, 913-918.

Tan, M. C., Regier, D. A., Esdaile, J. M., Lynd, L. D., Anis, A. H. & Marra, C. A. 2006. Health economic evaluation: a primer for the practicing rheumatologist. *Arthritis and Rheumatism,* 55, 648-656.

Tchouaket, E. & Brousselle, A. 2013. Using the results of economic evaluations of public health interventions: Challenges and proposals. *The Canadian Journal of Program Evaluation,* 28, 43.

ten Brink, P., Mutafoglu, K., Schweitzer, J.-P., Kettunen, M., Twigger-Ross, C., Baker, J., Kuipers, Y., Emonts, M., Tyrväinen, L. & Hujala, T. 2016. The Health and Social Benefits of Nature and Biodiversity Protection. A report for the European Commission London/Brussels: Institute for European Environmental Policy.

Thokala, P., Devlin, N., Marsh, K., Baltussen, R., Boysen, M., Kalo, Z., Longrenn, T., Mussen, F., Peacock, S., Watkins, J. & Ijzerman, M. 2016. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value in Health,* 19, 1-13.

Thokala, P. & Duenas, A. 2012. Multiple Criteria Decision Analysis for Health Technology Assessment. *Value in Health,* 15, 1172-1181.

Thompson Coon, J., Boddy, K., Stein, K., Whear, R., Barton, J. & Depledge, M. H. 2011. Does Participating in Physical Activity in Outdoor Natural Environments Have a Greater Effect on Physical and Mental Wellbeing than Physical Activity Indoors? A Systematic Review. *Environmental Science & Technology,* 45, 1761-1772.

Tinelli, M., Ryan, M. & Bond, C. 2016. What, who and when? Incorporating a discrete choice experiment into an economic evaluation. *Health Economics Review,* 6, 1-9.

Torrance, G. W., Feeny, D. & Furlong, W. 2001. Visual Analog Scales Do They Have a Role in the Measurement of Preferences for Health States? *Medical Decision Making,* 21, 329-334.

Townsend, M. 2006. Feel blue? Touch green! Participation in forest/woodland management as a treatment for depression. *Urban Forestry & Urban Greening,* 5, 111-120.

Train, K. E. 2009. *Discrete choice methods with simulation,* Cambridge university press.

Trueman, P. & Anokye, N. K. 2013. Applying economic evaluation to public health interventions: the case of interventions to promote physical activity. *Journal of Public Health,* 35, 32-39.

Tsuchiya, A. & Williams, A. 2001. Welfare economics and economic evaluation. *Economic evaluation in health care: merging theory with practice,* 27-28.

Tsuchiya, A., Williams, A., Drummond, M. & McGuire, A. 2001. Welfare economics and economic evaluation. *Economic evaluation in health care: merging theory with practice,* 27-28.

Tyrväinen, L., Ojala, A., Korpela, K., Lanki, T., Tsunetsugu, Y. & Kagawa, T. 2014. The influence of urban green environments on stress relief measures: A field experiment. *Journal of Environmental Psychology,* 38, 1-9.

Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J. & James, P. 2007. Promoting ecosystem and human

health in urban areas using Green Infrastructure: A literature review. *Landscape and Urban Planning,* 81, 167-178.
Ulrich, R. 1983. Aesthetic and affective response to natural environment. *Human Behavior & Environment: Advances in Theory and Research,* 6, 85-125.
Ulrich, R. 1984. View through a window may influence recovery. *Science*, 224-5.
van den Berg, A. E. & van den Berg, M. M. Health benefits of plants and green space: Establishing the evidence base. XI International People Plant Symposium on Diversity: Towards a New Vision of Nature 1093, 2012. 19-30.
van der Pol, M. & McKenzie, L. 2010. Costs and benefits of tele-endoscopy clinics in a remote location. *Journal of Telemedicine and Telecare,* 16, 89-94.
van Hout, B., Janssen, M. F., Feng, Y.-S., Kohlmann, T., Busschbach, J., Golicki, D., Lloyd, A., Scalone, L., Kind, P. & Pickard, A. S. 2012. Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets. *Value in Health,* 15, 708-715.
van Mastrigt, G. A., Paulus, A. T., Aarts, M.-J., Evers, S. M. & Alayli-Goebbels, A. F. 2015. A qualitative study on the views of experts regarding the incorporation of non-health outcomes into the economic evaluations of public health interventions. *BMC Public Health,* 15, 1-17.
Vecchiato, D. & Tempesta, T. 2013. Valuing the benefits of an afforestation project in a peri-urban area with choice experiments. *Forest Policy and Economics,* 26, 111-120.
Villa, J. M. 2012. Simplifying the estimation of difference in differences treatment effects with Stata. University Library of Munich, Germany.
Viney, R., Lancsar, E. & Louviere, J. 2002. Discrete choice experiments to measure consumer preferences for health and healthcare. *Expert Review of Pharmacoeconomics & Outcomes Research,* 2, 319-326.
Viney, R., Savage, E. & Louviere, J. 2005. Empirical investigation of experimental design properties of discrete choice experiments in health care. *Health Economics,* 14, 349-362.
Vojáček, O. & Pecáková, I. 2010. Comparison of Discrete Choice Models for Economic Environmental Research. *Prague Economic Papers,* 19, 35-53.
Walter, E. & Zehetmayr, S. 2006. Guidelines on Health Economic Evaluation-Consensus paper Austria: Institute for Pharmaeconomic Research

Walters, S. & Brazier, J. 2005. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Quality of Life Research,* 14**,** 1523-1532.

Wanless, D. 2004. Securing good health for the whole population. *Securing good health for the whole population*.

Ward Thompson, C., Aspinall, P., Bell, S. & Findlay, C. 2005. "It Gets You Away From Everyday Life": Local Woodlands and Community Use—What Makes a Difference? *Landscape Research,* 30**,** 109-146.

Ward Thompson, C., Aspinall, P., Bell, S., Findlay, C., Wherrett, J. & Travlou, P. 2004. Open space and social inclusion local woodland use in central scotland Edinburgh: Openspace: the Research Centre for Inclusive Access to Outdoor Environments in association with Landscape Architecture, Edinburgh College of Art and the School of the Built Environment, Heriot-Watt University.

Ward Thompson, C., Roe, J. & Alves, S. 2007. Woods in and around towns (WIAT). Evaluation: Baseline survey. *Unpublished report for Forestry Commission Scotland, Edinburgh, UK: OPENspace research centre*.

Ward Thompson, C., Roe, J. & Aspinall, P. 2013. Woodland improvements in deprived urban communities: What impact do they have on people's activities and quality of life? *Landscape and Urban Planning,* 118**,** 79-89.

Ward Thompson, C., Roe, J., Aspinall, P., Mitchell, R., Clow, A. & Miller, D. 2012. More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns. *Landscape and Urban Planning,* 105**,** 221-229.

Weatherly, H., Drummond, M., Claxton, K., Cookson, R., Ferguson, B., Godfrey, C., Rice, N., Sculpher, M. & Snowden, A. 2009. Methods for assessing the cost-effectiveness of public health interventions: key challenges and recommendations. *Health Policy,* 93.

Weatherly, H. L. A., Cookson, R. A. & Drummond, M. F. 2014. Economic Evaluation of Public Health Interventions: Methodological Challenges A2 - Culyer, Anthony J. *Encyclopedia of Health Economics*. San Diego: Elsevier.

Weinstein, M. C. 1988. A QALY is a QALY is a QALY–or is it. *J Health Econ,* 7**,** 289-90.

Weinstein, M. C., Torrance, G. & McGuire, A. 2009. QALYs: The Basics. *Value in Health,* 12**,** S5-S9.

Wells, N. M. & Evans, G. W. 2003. Nearby Nature: A Buffer of Life Stress among Rural Children. *Environment and Behavior,* 35**,** 311-330.

Westra, T. A., Parouty, M., Brouwer, W. B., Beutels, P. H., Rogoza, R. M., Rozenbaum, M. H., Daemen, T., Wilschut, J. C., Boersma, C. & Postma, M. J. 2012. On discounting of health gains from human papillomavirus vaccination: effects of different approaches. *Value*

in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research, 15, 562-567.
Wheeler, B. W., Lovell, R., Higgins, S. L., White, M. P., Alcock, I., Osborne, N. J., Husk, K., Sabel, C. E. & Depledge, M. H. 2015. Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality. *International Journal of Health Geographics,* 14, 17.
White, H. & Sabarwal, S. 2014. Quasi-experimental design and methods. *Methodological Briefs: Impact Evaluation,* 8.
White, M. P., Alcock, I., Wheeler, B. W. & Depledge, M. H. 2013. Would You Be Happier Living in a Greener Urban Area? A Fixed-Effects Analysis of Panel Data. *Psychological Science*.
Whitehead, S. J. & Ali, S. 2010. Health Outcomes in Economic Evaluation: the QALY and Utilities. *British Medical Bulletin*.
WHO 2010. Mental health: strengthening our response fact sheet no. 220, September 2010.
WHO 2014. The case for investing in public health, the strengthening public health services and capacity: A key pillar of the Europe regional health policy framework, Health 2020. *In:* Nurse, J., Dorey, S., Yao, L., Sigfrid, L. & Yfantopolous, P. (eds.) *A public health summary report for EPHO 8.* World Health Organisation.
WHO 2016. Urban green spaces and health - a review of evidence. Copenhagen WHO Regional Office for Europe.
Wildman, J., McMeekin, P., Grieve, E. & Briggs, A. 2016. Economic evaluation of integrated new technologies for health and social care: Suggestions for policy makers, users and evaluators. *Social Science & Medicine,* 169, 141-148.
Wilson, E. O. 1984. Sociobiology (1980) and Biophilia: The Human Bond to Other Species. Harvard University Press, Cambridge, Mass.
Wilson, E. O. & Kellert, S. R. 1993. The biophilia hypothesis. *Washington DC: Island*.
Wolf, K. L. F., K.; 2010. Mental Health and Function- A Literature Review. . *Green Cities: Good Health*.
Wolowacz, S. E., Briggs, A., Belozeroff, V., Clarke, P., Doward, L., Goeree, R., Lloyd, A. & Norman, R. 2016. Estimating health-state utility for economic models in clinical studies: an ISPOR Good Research Practices Task Force Report. *Value in Health,* 19, 704-719.
Wonderling, D., Black, N. & Raine, R. 2005. *Introduction to health economics*, McGraw-Hill Education (UK).
Wooldridge, J. 2012. *Introductory econometrics: A modern approach*, Cengage Learning.
Xiaolu, Z. & Md Masud, P. R. 2012. Social benefits of urban green space: A conceptual framework of valuation and accessibility

measurements. *Management of Environmental Quality: An International Journal,* 23, 173-189.

Zethraeus, N., Johannesson, M., Jönsson, B., Löthgren, M. & Tambour, M. 2003. Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. *Pharmacoeconomics,* 21, 39-48.