Operationalising the Capability Approach as an Outcome Measure in Public Health: the development of the OCAP-18

Abstract

There is growing interest in operationalising the capability approach to measure quality of life. This paper reports the results of a research project undertaken in 2007 that sought to reduce and refine a longer survey in order to provide a summary measure of wellbeing and capability in the realm of public health. The reduction and refinement of the questionnaire took place across a number of stages, using both qualitative (five focus group discussions and 17 in-depth interviews) and quantitative (secondary data analysis, N=1,048 and primary data collection using postal surveys and interviews, N=45) approaches. The questionnaire was reduced from its original 60+ questions to 24 questions (including demographic questions). Each of Nussbaum's ten Central Human Capabilities are measured using one (or more) of the 18 specific capability items which are included in the questionnaire (referred to as the OCAP-18). Analysis of the questionnaire responses (N=198) found that respondents differed with respect to the levels of capabilities they reported, and that these capabilities appear to be sensitive to one's gender, age, income and deprivation decile. An index of capability, estimated by assuming equal weight for each capability question, found that the average level of capability amongst respondents was 12.44 (range 3-17.75). This index was found to be highly correlated with a measure of health (EQ-5D) and wellbeing (global QoL), although some differences were apparent. This project operationalised the capability approach to produce an instrument to measure the effectiveness (and cost effectiveness) of public health interventions; the resulting OCAP-18 appears to be responsive and measure something supplementary to health and wellbeing, thus offers a promising addition to the current suite of outcome measures that are available.

Keywords: United Kingdom; capability approach; public health; outcome measurement; questionnaire development; survey instrument

Word Count: 9884 (abstract, text, references)

1. Introduction

Public health interventions are intended to promote health or prevent ill health in communities or populations, and can be distinguished from clinical or medical interventions which intend to prevent or treat ill health in individuals (Rychetnik et al. 2002). Public health focuses in part on behavioural risk factors like obesity and smoking (Petersen, Petersen & Lupton 1996), and population-level problems of inequality and poverty (Marmot 2005; Wilkinson & Pickett 2006). This has resulted in interventions which promote public health or seek to improve population health becoming more complex; this complexity can be evident in the intervention, the outcomes or the evaluation itself (Shiell, Hawe & Gold 2008).

This paper specifically focuses on the issue of public health outcomes. While public health interventions are predominantly interested in improving physical (and more recently mental) health, this is not necessarily their only outcome; the assessment of a broad range of outcomes is not uncommon, especially when there are a number of stakeholders involved (Sridharan, Campbell & Zinzow 2006). For example, the GoWell urban regeneration programme in Glasgow is interested in the potential benefits to both individuals and communities across various domains including housing, health, employment, the environment and crime (Egan et al. 2010). Similarly a community maternal health programme in a Malawi is not just interested in improving the health of mother and child, but also empowering women, building capacity and imparting knowledge (Lewycka et al. 2010). Such multiple and complex outcomes can pose a problem for evaluators, particularly economists tasked with estimating the cost effectiveness of such interventions.

The aim of economic evaluation is to identify whether a proposed change in service provision is a good use of scarce resources (Drummond 2005). This requires a comparison of the

additional costs associated with the change and the additional outcomes achieved by the change. The definition, assessment and measurement of the outcomes are key issues for economic evaluation. In healthcare, outcomes are commonly assessed using Quality Adjusted Life Years (QALYs) and results are presented through an incremental cost effectiveness ratio (ICER) indicating the additional costs per additional QALY gained from the intervention. Public health interventions with diverse outcomes (such as the urban regeneration programme or community maternal health programme described above) are unlikely to be wholly captured within the QALY framework. One reason for this is that most, if not all, of the multi-attribute utility instruments (MAUIs) that are used to estimate utilities or values for QALY estimation focus on *health* related quality of life (HRQoL). Therefore, QALYs and their associated HRQoL measures like the EQ-5D or SF-6D are likely to underestimate the relative benefits of public health interventions. This creates a dilemma for economic evaluation. Do we persevere with the cost per QALY approach even though we are aware it may not capture all the important outcomes? Do we present the cost and the diverse outcomes (consequences) separately (in the form of a cost consequence analysis (Coast 2004)) even though this does not provide a single answer to the question regarding the use of resources? Or do we attempt to find a new measure which can incorporate all of the outcomes of interest/importance in order to address the question of resource allocation (Lorgelly et al. 2010)?

Sen's Capability Approach (Sen 1985; 1993) would appear to provide a possible solution to the limitations of QALYs, as it expands the evaluative space (so it can include non-health outcomes like empowerment, participation, housing, and crime) to consider whether a programme/policy/intervention enhances an individual's capability (Lorgelly et al. 2010). Previous evaluative approaches focus on subjective-wellbeing (utilitarism) or the availability

of means for a good life (resourcism). The alternative paradigm of the capability approach instead suggests that the focus of wellbeing should be a set of valuable 'beings and doings' (for example being in good health or having loving relationships), which can be measured by opportunities (capabilities) or outcomes (functionings) (Sen 1992). Sen desires that policies ought to promote the "capabilities of persons to lead the kind of lives they value – and have reason to value' (Sen 1999, p.18). Of interest in its application to public health is the evaluative space; it diverges from narrow utility space, which is concerned with the pleasure obtained from the consumption of goods and services, and instead encapsulates an informational space, where evaluative judgements are about an individual's freedom. Therefore, Sen's approach is based on value judgments, which ultimately relate to an individual's capability set. Because it moves away from mental states, utility or welfare, in this sense it can be described as 'non-welfarist' (Coast, Smith & Lorgelly 2008a; 2008b).

One of the limitations of the approach is that "Sen has not specified how the various value judgments that inhere in his approach and are required in order for its practical use (whether at the micro or macro level) are to be made" (Alkire 2002, p.3). He believes that various selection and deliberation are an intrinsic part of the approach. Nussbaum, however, has identified what she regards as central human capabilities, and provides a list of ten capabilities: life; bodily health; bodily integrity; senses, imagination and thought; emotions; practical reason; affiliation; other species; play; and control over one's environment (Nussbaum 2000). Other prescriptive capability lists also exist, which have varying degrees of abstraction and generalisation (e.g. Robeyns 2003). The existence of such lists is crucial in the evaluation of capability sets (that is, the identification of freedoms) and the subsequent operationalisation of the approach (that is, evaluating whether such freedoms are achievable).

This paper describes an attempt to operationalise the capability approach by refining and reducing a previously developed questionnaire (Anand & van Hees 2006; Anand et al. 2009). Anand et al.'s (2009) operationalisation of the approach, by assessing capabilities using pre-established questions, is a useful platform from which to develop a measure of outcome for use in evaluations of public health interventions. This is partly because its survey design is practical for use in large research projects which involve self-completing questionnaires or interviews. It is also a generic approach, much like the SF-36 is a generic measure of health (Ware & Sherbourne 1992), and so offers the potential to provide a summary measure of wellbeing and capability. This negates the need to develop specific instruments for every evaluation of complex public health interventions. The drawback of their approach, however, in terms of outcome measurement, is that there are over 60 indicators of capability, making its usability (particularly any wide scale adoption) limited.

This paper reports on a project which sought to further develop and refine the survey instrument as proposed by Anand et al. (2009); validate the instrument for use in public health evaluations; and propose how future evaluations might employ the capability approach. The paper proceeds by first detailing the methods used in the refinement and reduction of the questionnaire, both qualitative and quantitative approaches were employed. The methods for validating the final version of the questionnaire, using mainly quantitative analyses, are then described. Due to space constraints the results section focuses on the validation of the instrument. A discussion section concludes the paper.

2. Methods

Anand et al.'s (2009) questionnaire was reduced and refined in two stages. It is important to reiterate that the questionnaire had already been used to generate information on capability

(Anand et al. 2009) and as such some initial pretesting was employed in that project; this included discussions with colleagues and with the professional social research company who administered the questionnaire (YouGov). In this current study, after each stage of data collection the questionnaire was revised given the findings of both qualitative and quantitative analyses that were conducted in each stage. Across the course of the project this produced three versions of the questionnaire, version 1 (that is Anand's OCAP) was employed in the first stage, version 2 was used in the second stage, while the last stage of reduction and refinement culminated in the third and final version of the questionnaire (OCAP-18), which was ultimately employed to measure capability in a population of the UK public.

The qualitative and quantitative methods are described separately, but were employed concurrently; the approach to reduction and refinement was to remove, replace or combine questions but to do so while considering the context of the questions and their interpretation. The methodology draws heavily on research into questionnaire design (DeMaio et al. 2006; Presser et al. 2004; Sheatsley et al. 1983), as well as other work in the area of measuring capabilities (Comim 2008; Alkire, Kakwani & Silber 2008).

The project received ethical approval from the University of Glasgow's Medical Faculty Ethics Committee (FM00606).

2.1. Qualitative methods – focus groups (Stage One)

Qualitative data from focus groups have been identified as being particularly useful for informing the actual content of scale construction (Barbour et al. 1999). Focus groups can be used to refine information previously known about a topic, and they can also stimulate new

ideas or concepts and offer the opportunity to collect data from group interactions, exploring issues that individuals in a one-to-one interview may not raise. For this reason focus groups were employed in the first stage.

There was an attempt to target recruitment of the focus groups in order to include a plurality of voices (Silverman 2009); young, middle-aged and older individuals as well as individuals from affluent and deprived areas were purposively sampled from various community groups in Glasgow, United Kingdom – including a book group, a carers group, a youth group, a mental health service user group and a group recruited from the University. Five focus groups were organised, with approximately eight individuals in each group; participants were offered nominal monetary compensation for their time and effort.

Focus group participants were told they would be participating in a study that sought to develop a tool to evaluate public health interventions, and that discussions would centre upon notions of health and wellbeing. After consent was gained the groups commenced with participants being asked to complete two sections of the questionnaire and discuss their views towards the meaning of the questions and their general understanding of the questions. Note that in a pilot focus group, participants were asked to complete the whole questionnaire, and then focus on a specific section, but this was found to be too time consuming. As there were five focus groups, the questionnaire was split into five logical sections, and each group discussed two sections, such that each section was discussed by two groups, thus providing maximum crossover for minimal effort.

Focus group participants were encouraged to interact with each other rather than respond individually with the moderator. Focus group participants were asked to identify any specific

questions they found problematic, confusing or objectionable. Those questions identified during this process, were discussed in detail amongst the group, in an attempt to determine if there was a consensus within the group or if the issue was just held by one individual. The general layout of the questionnaire and other aesthetic issues were also discussed. Discussions of the questionnaire constituted the first part of the focus group discussion.

The second part of the focus group discussion involved participants reading two vignettes before being asked to make normative statements about the vignettes, including the individual involved and their set of social circumstances. Vignettes allow for beliefs, attitudes, values and norms to be revealed in a context-specific way (Finch 1987). It is a method which acknowledges that meanings are social and it provides a way to express meanings which do not restrict the participant to choices which may be contrary to their belief (as can happen in survey methods) (Finch 1987). Lay perceptions of health and wellbeing may not be easily conveyed, as it is a broad subject area that individuals may find difficult to articulate if they have not previously considered it in any meaningful and systematic way. Additionally, some people may not wish to divulge personal information about themselves. As such, vignettes provide a means to overcome these issues by encouraging responses from participants in the way that they are prompted to consider a topic and to do so with their 'personal world' protected through distance.

The vignettes that were discussed by each focus group are presented in Appendix 1. The purpose of the vignettes was to explore participants' lay understandings of both this explicit capability and capabilities more generally. Additionally, the value participants placed on this capability and their beliefs around how they could practically achieve a capability, such as this, within their own life context was examined. The moderator probed for contextual

influences impacting on their understanding and views towards capabilities. Common themes to emerge were fed back to participants at the end of the discussion to verify the views and beliefs that participants provided. All focus group discussions were audio recorded (with verbal consent from participants) and transcribed verbatim; these were supplemented with field notes taken by an observer noting the non-verbal interaction in the group. The vignettes focused on Nussbaum's capability of 'life', that is "being able to live to the end of a human life of normal length; not dying prematurely, or before one's life is so reduced as to be not worth living" (Nussbaum 2000, p.78). The vignettes differ in terms of the situation that the individual is exposed to, both externally (the Calton is a very deprived area of Glasgow, while Bearsden is very affluent) and internally (drinking, smoking, diet and exercise); and as such were to probe for contextual influences impacting on their understanding and views towards capabilities. During discussions around the vignettes, the moderator probed for participants' views on life expectancy so to assess the value participants' placed upon this capability. The moderator noted common themes to emerge and fed this back to participants at the end of the focus group to verify the views and beliefs that participants provided. All focus group discussions were audio recorded (with verbal consent from participants) and transcribed verbatim, and these were supplemented with field notes taken by an observer noting the non-verbal interaction in the group.

The analysis of the focus group data was dictated by the fact that the focus groups had a dual purpose: to gather information on the usability and user comprehension of the questionnaire, and to gather information on participants' views towards their health and wellbeing. In the first instance, the transcripts were reviewed to extract information on comprehension problems for each questionnaire item. As the same questions (sections of the questionnaire) were discussed by two groups, the transcripts from both groups were analysed together. The

groups differed in composition and these confounders (age, gender or social status) were considered during the analysis. The data from the vignette discussion were analysed thematically both within and across groups (Aronson 1994). Emerging themes identified by the primary coder were discussed with other members of the research team in order to verify the analysis. During this stage, competing and alternate explanations were considered in order to ensure the analysis is verifiable and therefore ensuring the trustworthiness of the data (Tashakkori & Teddlie 1998).

2.2. Qualitative methods – cognitive interviews (Stage Two)

Cognitive interviews were employed in the second stage of the study, after the first revision of the questionnaire had been completed. The interviews were used as a further means of pre-testing and to check for face validity. Semi-structured interviews have been used extensively to capture data that can assist with survey development (Prieto, Thorsen & Juul 2005; Storck et al. 2006; Wamcata et al. 2005). Used in this way, the semi-structured technique can identify salient issues and explore meanings attached to particular items. Conducted alongside a questionnaire they can reveal the process of replying to the survey questions from the respondent's perspective. Cognitive interviews can unpack the four stages respondents work through in order to reply to a survey question: comprehension (understanding the question); recalling information; judgment (deciding upon the question relevance); and response (formulating an answer in the format provided by the interviewer) (Willis 2005).

Interview participants were identified using a postcode address file (PAF) to identify postal addresses in the Greater Glasgow area. A random sampling algorithm (based on postcode sectors), stratified to over-sample in deprived areas to compensate for the expected low

response rate in such areas, selected 400 addresses to which invitations for interviews were sent. It was envisaged that around 30 semi-structured interviews would be conducted; however, in practice this could be less if saturation was reached before all 30 interviews are conducted. In total 37 individuals indicated an interest in being interviewed and saturation was reached after 17 interviews.

The interviews began with participants completing the questionnaire and then responding to questions, which sought to understand participants' comprehension of and difficulties with the questionnaire. The interviewer utilised the 'verbal probe' technique, to explore the basis for the response; this is an increasingly common technique, used as an alternative to 'think aloud'. General probes ('How did you arrive at that answer?') were used along with specific probes to explore comprehension and recall. Particular attention was paid to questions requiring revision, or new questions introduced during the previous stage. Interviews were kept to no more than one hour to avoid respondent fatigue. All interviews were recorded on digital recorders (with the respondent's consent). Individual transcripts were read repeatedly by the qualitative researcher and coded according to identified emerging themes; subsequent recurring themes were then identified across the transcripts. Another member of the research team also read a sample of transcripts and the thematic analysis was jointly discussed until a consensus was reached on the main themes to emerge.

2.3. Quantitative methods – factor analysis (Stage One and Two)

The data previously generated through the YouGov web survey (Anand et al., 2009) were also analysed in stage one of the project. The original survey, while internet based, was essentially identical to the first version of the paper based questionnaire employed here. Anand et al.'s survey elicited a large amount of data (N=1,048), such that considerable

quantitative analysis could be undertaken. In the first instance, the responses to each question were tabulated to provide some sense of how often the range of answer options was utilised (most questions offered one of seven answer options, e.g. agree strongly, agree moderately, agree a little, neither agree nor disagree, disagree a little, disagree moderately, disagree strongly). If the range of answers was not widely used then this implied that the questionnaire could be refined in terms of simply reducing the number of answer options available. This is in keeping with Comim's (2008) suggestion regarding the choice of categories to appropriately represent the scale.

Subsequent analysis employed factor analysis, a statistical technique which aims to simplify complex sets of data, by attempting to describe correlations between variables (Klein 1994; Lelli et al. 2008). It does so by identifying a set of factors with factor loadings, that is the correlation of a variable with a factor. In this sense it can be used for item reduction, identifying questions that may have similar loadings, suggesting one of them, although not necessarily identifying which one, is redundant. For example in the OCAP questionnaire there were some 15 questions encompassed within one item (Nussbaum's 'affiliation'), so one might expect that a number of these questions are redundant. In the first instance 'factor analysis of the whole' was undertaken, whereby all questions were considered together and the analysis sought to identify whether the pattern of factor loadings was as expected. If each question (or group of questions) is independently and accurately measuring one of the ten explicit capabilities as put forward by Nussbaum, then ten factors should be evident from the factor analysis. However, given that there are multiple questions for some capabilities (and not for others – which could introduce problems of dominance into the analysis), 'factor analysis of the parts' was also undertaken; that is for specific capability domains (within) factor analysis was carried out to see if some questions were more dominant that others.

This, together with simple correlation plots, provides further insight regarding potentially redundant questions.

Factor analysis was also employed in the second stage. In addition to the cognitive interviews using version 2 of the questionnaire, this version of the questionnaire was also sent out to 200 randomly selected households in the Greater Glasgow area (using the same sampling approach described above). The questionnaires completed during the interviews were combined with the postal questionnaires, and the complete sample (N=45) was again subjected to response category analysis (that is tabulation of frequencies to compare the distribution of responses) and factor analysis. Additional comparative analysis compared two versions of the postal questionnaire, half of the postal sample received a questionnaire where the answers read positively from left to right ('difficult' to 'easy', 'unsafe' to 'safe') and the other half received a questionnaire where the answers read negatively from left to right ('easy' to 'difficult', 'safe' to 'unsafe'). These were used to test for response set bias, that is the tendency for respondents to answer a series of questions in a certain direction regardless of their content (Fox & Tracy 1986).

Note that all but one question (in all versions of the questionnaire) had categorical response options. The life expectancy question, which corresponds to Nussbaum 'life' capability "Being about to live to the end of a human life of normal length …" asked respondents to provide an estimate of their life expectancy given their family history, dietary habits, lifestyle and health status. In order to put this question into the context of a capability, the difference between one's actual life expectancy (given each respondent's age and gender, as estimated from life tables for Glasgow City) and predicted (or expected) life expectancy (as reported by respondents) was calculated. This deviation in life expectancy is used in all analyses.

2.4. Process of reduction and refinement

The first reduction and revision of the questionnaire was informed by the quantitative analysis of the YouGov responses and also by the emerging themes from the qualitative analysis of the focus group data. An advisory panel, consisting of the five members of the project team and two experts in the field, convened and each of the 64 questions were discussed in turn in relation to the qualitative and quantitative findings. Essentially the qualitative analysis was informative for question refinement, while the quantitative analysis was more informative for question reduction, although there were instances where the findings overlapped, for example where reduction/removal was supported by both the quantitative and qualitative analysis. Initially it was proposed that collective agreement from all members of the panel was required to remove or revise a question, but in practice, due to variations in opinions, the decision often came down to a democratic vote with majority rule.

The questionnaire was revised (version 2) and then subjected to a second stage of reduction and revision, as described above. Again the advisory panel met, results were presented and issues raised were discussed and a majority consensus was required to remove or revise a question. This resulted in the final version of the questionnaire, version 3 (subsequently referred to as the OCAP-18).

2.5. Validation of the reduced/refined questionnaire

In order to validate the condensed questionnaire the final stage of the project tested this version (version 3). Further semi-structured interviews were employed, loosely following the previous 'verbal probe' technique, and it was also sent out as a postal survey. These interviews additionally allowed us to explore respondents' values and preferences regarding

functioning and capability (see section 2.6). The questionnaire included the reduced and refined set of capability questions (the OCAP-18 instrument), questions relating to respondents socio-demographic status (gender, age, race, education level, employment status, marital status, income, etc), and personality questions. It also included two commonly used quality of life (QoL) instruments, the health-focused EQ-5D-3L and the wellbeing-focused global QoL scale.

The EQ-5D-3L is a commonly used measure of health status in health economics. Five questions/domains each with three levels are used to elicit information on an individual's health profile. Each profile corresponds to a tariff (a utility, value or preference) which was estimated from interviews with the general public (EuroQol 1991). A value of 1 represents perfect health and 0 represents dead, although there are some states considered to be worse than dead. The global QoL scale is argued to provide a *global* – that is overall – estimate of QoL; as it is a single question it is distinguished from other *total* measures of QoL which aggregate across items (Hyland & Sodergren 1996). Its creators argue that it provides a nomative overall judgment made by the respondent, of all the different aspects of what the respondent means by QoL and is therefore devoid of any researcher imposed value. It is a categorical rating scale with labelled end points (100 is perfect quality of life and 0 is no quality of life) and eight additional quantifiers placed at defined points (as determined by research subjects) along the scale.

Personality profiles offer further understanding of the traits of respondents, and may help explain variations in capabilities. Personality has been found to be highly correlated with life satisfaction (Schimmack et al. 2002; 2004), and others have used it as a measure of psychological capital when analysing the capability approach from a stocks and flows

approach (Muffels & Headey 2009). To measure personality a brief inventory was included, ten questions which assessed extraversion, agreeableness, conscientiousness, emotional stability and openness to experiences (Gosling, Rentfrow & Swann 2003).

2.5.1. Validation dataset

The PAF was again employed as the sampling frame, and 1000 private residential addresses in Glasgow City were randomly sampled (with some stratification for deprivation) to receive a postal questionnaire. While an additional 400 households (again stratified but with convenience sampling) were sent invitations for interview.

The data collected from the interviews and returned postal questionnaires were analysed together. Each of the individual capability questions was considered in terms of the mean response, but also in terms of variation to understand the use of the categorical response options. Note, as described above, reported life expectancy was compared to actual life expectancy (given life tables) and the analyses considered deviations in life expectancy. Correlations across individual capability questions were also explored, this particularly focused on questions contained within the same Nussbaum capability domain to understand if there was scope for further refinement. Factor analysis was again used to determine if further refinement or reduction was possible.

2.5.2. Assessment of subgroup differences

Inequalities in reported capabilities were explored; four groupings or types of inequalities were of interest: deprivation (as measured by Carstairs deprivation deciles, taken from respondents' postcodes (Morris & Carstairs 1991)), income, gender and age. In order to undertake meaningful comparisons it was necessary to combine some of the categories for

both deprivation and income. Three deprivation groups were created, those in postcode sectors with a deprivation decile of 1 to 6 were grouped together, as were those in postcode sectors with a deprivation decile of 7 to 9. The final group included respondents who resided in deprivation decile 10 (considered to be the most deprived postcode sectors in Scotland in 2001). Likewise, household income has been grouped into 4 groups of: less than £10,000 per year; between £10,000 and £20,000 per year; between £20,000 and £40,000; and household income greater than £40,000 per year. Age was categorised as less than 40 years old, between 40 and 60 years old and greater than 60 years old. Significant differences were examined using a chi-squared test (except the comparison of mean deviations in life expectancy which was undertaken using an F test).

2.5.3. Estimation of an index of capability

The instrument (the set of capability questions) would be of most use if the questions could be collapsed into an index, such that a single number could be generated and compared. This would mean that every individual would have an index of capability, and for evaluation purposes the index could be estimated at multiple time points, and improvements (or reductions) in capability could be easily measured. There are, however, two criteria that must be satisfied in order to estimate an index of capability. First it is necessary to consider whether the instrument itself is actually measuring capability, and whether a different composite instrument (with different questions and/or domains) might exist. Secondly, it is necessary to consider the weights (or tradeoffs) of the different components of the instrument (that is the specific capabilities) and how they might relate to each other.

When combining questions, the simplest approach to take is to assume equal weight for each capability. For instance, not having the capability to live a normal length of life (as proposed

by Nussbaum) is regarded as important as not being capable of having self respect, which is considered as equally important as whether one is capable of having adequate shelter or forming concept of good and engaging in critical reflection. Therefore, to estimate this index each question is given the same weight, and an index is generated by aggregating each question. Deviations in life expectancy, a continuous variable, was dichotomised such that those with deviations above average are coded as a 1 (that is their expectations are higher than average), and those below average are coded as 0; using quintiles was found to make little difference to the results.

The analysis compared the capability index across four inequality domains (deprivation, income, age and gender), considering the difference in the mean index value. The index was also employed in a multivariate regression to understand the independent effect of these groupings, and particularly whether any bivariate relationship identified in the analysis of mean differences held in the presence of other confounders.

A final analysis assessed the correlation between the OCAP-18 capability index and the EQ-5D-3L and the global QoL scale. This provides insight as to how similar or different a measure of capability is to accepted measures of health and wellbeing.

2.6. Capability vs functioning

Given the importance placed upon participatory processes in operationalising the capability approach, it is important that people are able to engage with and understand these concepts. Functioning, capability, opportunity and freedoms are easily understood (and distinguished) by an academic, but is this also the case for the respondents, those whose capabilities we are trying to measure? The focus group discussions and interviews in Stage One and Two of the

project suggested that there were difficulties among respondents in understanding the language of capability. The interview based data collection using version 3 of questionnaire provided an opportunity to explore this further.

At the completion of the interview-based questionnaire, each respondent was asked to provide some insight regarding what he or she valued more: the capability (being able to do something) or the functioning (actually doing something). Specifically they were asked to provide a preference for the capability domains of 'bodily health', 'senses, imagination and thought', 'emotions', 'practical reason' and 'control over one's environment'. Respondents were asked "What, in your view, do you value more":

- Being *able* to be adequately nourished or *actually* being adequately nourished?
- Being *able* to express your views, including political views or *actually* expressing them?
- Being *able* to plan how you would like your life to be or *actually* doing so?
- Being *able* to enjoy the love, care and support of your family and friends or *actually* enjoying it?
- Being *able* to influence decisions affecting my local area or *actually* doing so?

Their responses were analysed (including any qualitative comments) and compared across the various capability domains, in order to offer insight on the general public's understanding of capability vis-à-vis functioning.

3. Results

3.1. Item Reduction and Questionnaire Refinement

Due to space constraints this paper will focus on the analysis of the final version (version 3) of the questionnaire, including the tests of the instrument's validity. However, for

completeness, a short description of the process of item reduction and questionnaire refinement that was undertaken in each stage is detailed below.

A graphical representation of the reduction/refinement process is presented in Appendix 2. In Appendix 2 the first column presents Nussbaum's list of central human capabilities (life; bodily health; bodily integrity; senses imagination and thought; emotions; practical reason; affiliation; other species; play; and control over one's environment), while the second column (version 1) presents the questions from Anand et al.'s (2009) original questionnaire (the OCAP), classified into each of Nussbaum's ten capabilities. The questions used in the first revision of the questionnaire are presented in the third column (version 2), while the last column presents the final version of the questionnaire, version 3 (the OCAP-18). Reading from left to right shows the process of item reduction and question refinement.

During the first stage of item reduction, which employed factor analysis on Anand et al.'s (2009) YouGov data, questions were removed if: factor loadings suggested correlation with other questions; pairwise correlations were significant; and there were multiple questions measuring a specific capability; or questions measured functioning rather than capability. The remaining questions were refined given the analysis of the focus group discussion data in Stage One. Issues that were addressed included: ordering; merging; consistency in question wording and answer options (including reduction in answer options); understanding and interpretation of terminology. Key changes to version 1 of the questionnaire included:

- Changed from seven option answers to five options, also four to five, so that there is commonality across the questions. The question wording was changed to reflect this.
- Limited the different types of response options. The revised questionnaire only used six different categorical scales: always to never, strongly agree to strongly disagree, very easy

to very difficult, very suitable to very unsuitable, very safe to very unsafe, very likely to very unlikely;

- Conflated questions which ask to agree/disagree with statements, and conflated the discrimination questions (layout issue);
- Removed multiple questions, e.g. safety before and after dark became one question;
- Used more established questions, in particular the adequately nourished question was changed;
- Refined the wording, such that 'recently' was replaced with 'in the past 4 weeks', so to ensure consistency;
- Changed the ordering of the question, so that questions sit together in a more logical order and certain questions become less obtrusive (for example the initial questionnaire opened with "what age do you expect to live to", rather confronting as the first question).

In Stage Two, the quantitative analysis of the postal (N=28, response rate of 14%) and interview (N=17) data were combined to inform the item reduction. Questions were removed if: strong correlations were found; they appeared not be a measure of capability, rather a measure of functioning (given qualitative analysis), this was complemented by the quantitative analysis (in terms of correlations and factor loadings); or they were considered to be a capability in a developing country context, rather than specific to public health interventions (given the capability approach was developed with respect to poverty and human development, some of the concepts and questions are not relevant to the domain of public health, i.e. choices in matters of reproduction). Thematic analysis of the cognitive interview data informed questionnaire refinement, questions were refined according to: ordering; understanding; and interpretation of terminology. There was a particular focus on the wording, such that the statements or questions explicitly focused on capability, for

example “I respect, value and appreciate people around me” became “I *am able to* respect, value and appreciate people around me”.

Anand et al.’s original 64 capabilities – nested within Nussbaum’s list of ten – were, after the analysis at Stage One, reduced to 43 capabilities. The Stage Two analyses produced a final set of 18 specific capabilities (see Appendix 2, version 3). The final version of the questionnaire: the OCAP-18 questions and the socio-demographic questions, is available in the Supplementary Material [INSERT LINK TO ONLINE FILE].

3.2. Measurement of Capabilities

In October 2007, version 3 of the questionnaire was sent out to 1000 addresses within Glasgow City. 32 were returned with incomplete or as ineligible addresses, 180 questionnaires were returned completed. This resulted in a response rate of 18.6%. In addition, during October and November 2007, 18 respondents completed the questionnaire in an interview setting. Due to the small proportion of interview questionnaires relative to postal questionnaires, it was not feasible to undertake any comparative analysis by elicitation method (that is to compare postal with interview responses); therefore all questionnaires were analysed together giving a total sample size of 198.

Table 1 presents a detailed description of the demographics of the sample. In summary the majority of respondents were: white (97%), female (62%), employed full-time (50%), had some form of higher education (45%) or no qualifications (24%), either married (30%) or never married (34%), with no dependent children (69%), had no religion (35%), were Presbyterian (26%) or Catholic (28%), with a household income of under £30,000 per year (61%). The average age of respondents was 46 years old (range 19 to 91 years). Recall that

557 the original sampling algorithm (based on postcode sectors) was stratified to over-sample in
558 deprived areas to compensate for the expected low response rate from such areas; Table 1
559 shows that as a consequence of this strategy the proportion of respondents living in each
560 deprivation decile are relatively representative of the Glasgow population. Just over half of
561 the survey respondents live in a decile 10 postcode sector, compared with 54% of the
562 population of Glasgow.
563
564 *3.3. Sensitivity to inequalities*
565 An analysis of the questionnaire responses found that respondents had a range of capabilities
566 (see Figures 1 and 2 for two examples), and that these capabilities appear to be sensitive to
567 one's gender, age, income and deprivation decile (see Table 2).
568
569 An analysis of inequalities within individual capabilities and questions about capabilities
570 found that males were seemingly more accurate at predicting their life expectancy ('life'
571 capability), whilst males also believed that they are more likely to be victims of assault
572 ('bodily integrity' capability).
573
574 The elderly (those older than 60 years of age) were more likely to report that their health
575 limited their activities of daily life relative to younger respondents ('bodily health'
576 capability), while a higher proportion of younger respondents (those aged under 60 years old)
577 felt they were likely to experience discrimination outside of their place of employment
578 compared to older respondents ('affiliation' capability). This is likely to be a consequence of
579 those over 60 having limited employment opportunities.
580
581 Those living in more deprived areas were found to report greater limitations in their daily

activities due to their health status ('bodily health' capability), as well as feel less safe walking in their neighbourhood ('bodily integrity' capability), reported having fewer opportunities to socialise ('emotions' capability) and were less able to afford to own property than respondents in the more affluent areas of Glasgow City ('control over one's life' capability).

Those in low income groups were found to have worse health in terms of limiting daily activities ('bodily health' capability), and to predict life expectancies well below that expected given their age and gender ('life' capability), compared to those in higher income groups. Respondents with low household incomes also reported limitations in terms of socialising with friends and family ('emotions' capability) and owning property ('control over one's life' capability). They were also less likely to feel they could influence local decision making ('control over one's life' capability), more likely to report losing sleep over worry ('emotions' capability) and were rarely able to enjoy recreational activity ('play' capability) relative to respondents with high household incomes.

3.4. Index of Capability

To estimate an index of capability each capability question is given the same weight, and then these were aggregated to generate an index. Taking this approach and applying it only to the sample of respondents who answered every one of the 18 questions (N=166), it is estimated that the mean index of capability for the sample is 12.44 (range: 3–17.75). Figure 3 presents a histogram of the index.

Given a number of significant differences were found when considering the specific capabilities across the predefined groups of interest – gender, age, deprivation and income –

it is of interest to analyse whether such differences also exist in the index of capability. Table 3 presents a descriptive analysis of the mean index for each group and in the final column provides evidence of the level of significance of any difference. Notably, there are no evident gender or age differences; however, both those in deprived areas and those of low income are found to have lower capability as measured by the index.

In order to determine whether these significant differences in mean capability scores are independent of the effects of other variables, a multivariate regression was undertaken. Capability was estimated as a function of gender, age, income and deprivation. The regressions results are presented in Table 4. Table 4 shows that respondents aged over 60 years have marginally higher capability that those aged under 40 years old (p value < 0.10), while those with a household income between £10,000 and £19,000 also have marginally higher capability than those respondents in the lower income group. Respondents earning more than £20,000 were found to have significantly higher capability than those in the reference category (earning less than £10,000 per year). Notably the significant (pairwise) relationship between deprivation and capability (as presented in Table 3), is not found to hold in this multivariate regression, suggesting that income is a more dominant explanatory variable.

3.5. Test of validity

Given that the EQ-5D-3L is an accepted measure of outcome in economic evaluations (at least for health care interventions), it is interesting to see how similar or different the index of capability is to EQ-5D-3L as a measure of health. Figure 4 provides a graphical representation of this, and statistical analysis finds that they are highly correlated (pairwise correlation: 0.576; p-value: <0.001). Notably there are some deviations from the mean,

632 which suggests that they are capturing or measuring some concepts differently. Figure 5,
633 shows a similar relationship between global QoL and the capability index.
634
635 *3.6. Functionings versus Capabilities*
636 Interview respondents preferences for capability and functioning across the domains of
637 'bodily health', 'senses, imagination and thought', 'emotions', 'practical reason' and 'control
638 over one's environment' are presented in Table 5. There appears to be a significant support
639 for having the capability to express one's views rather than the actual expression of them, and
640 also to have the capability to influence decisions, rather than actually influencing them. The
641 other capability domains have a more mixed response.
642
643 **4. Discussion**
644 Public health interventions are becoming more complex, their numerous and broad
645 consequences require a new approach in order to evaluate the success of such interventions
646 (Smith & Petticrew 2010). Sen's capability approach has been argued to provide many
647 benefits particularly when seeking to evaluate the cost effectiveness of such interventions and
648 programmes Lorgelly et al. 2010). It offers a much richer set of dimensions for evaluation,
649 thereby potentially capturing all relevant outcomes, rather than focusing solely on health
650 status (as is the current approach in health economics) (Coast, Smith & Lorgelly 2008a). The
651 equitable underpinnings of the approach are also appropriate for use with public health
652 interventions that often seek to reduce/remove inequalities across groups (namely improving
653 deprivation) as an overriding aim. In terms of the practical issues of operationalising the
654 approach and measuring capabilities, it would appear that the questionnaire reduced and
655 refined here provides one means of doing this. There is evidence that it is responsive to
656 different groups of individuals, and it appears to measure something in addition to health and

wellbeing, although is still highly correlated with these measures.

The questionnaire was reduced and refined in a high income setting with a focus of future evaluations of public health interventions. The approach that was employed was highly participatory, respondent lead, and it could easily be replicated in another setting with another interest in mind. It is noted that some questions were removed as they were not considered relevant to the domain of public health in the setting of interest (e.g. matters of reproduction), but such questions could be re-introduced if other researchers regarded them as important. A recent example of adaptation is the refinement of OCAP-18 for use with patients with mental illness (Simon et al. 2013). The researchers identified four questions that needed modification given the patient group: discrimination at work, meeting socially with work colleagues, life expectancy, love and support – because they weren't relevant, caused distress or were not easily interpretable. They also identified a further dimension (access to activities/employment) that was deemed important for people with mental health problems. The adapted OCAP-18 has been renamed the OxCAP-MH, and the instrument has been found to be both feasible and valid for measuring capabilities for the mentally ill (Vergunst et al. 2014). This recent adaptation and the reduction/refinement presented here raise an outstanding issue: how many questions are required to capture capability? From Appendix 2 it is evident that some capabilities have more questions than others, compare 'Affiliation' and 'Practical Reason'. The answer is undoubtedly context specific, and Nussbaum (2011) suggests that there could be more capabilities as a result of changing context. Our approach sought to refine an already long questionnaire, but the focus groups could have been directed to discuss broader capabilities and dimensions if the context warranted it.

While the benefits of using the capability approach to evaluate public health interventions are

numerous (e.g. a richer evaluative space with a focus of equity) implementing the approach does involve a number of challenges. Specifically, in order to operationalise the approach for use in economic evaluations, it will be most useful if an index is generated, whereby an individual's capability is described by a composite single number, which reflects the relative importance of the different dimensions/domains. The current approach is rather simplistic, assuming that all capabilities have equal weight, that is all are equally important. Arguably this is no different to the fact that fundamental human rights cannot be traded (Devidal 2004; Cornwall & Nyamu-Musembi 2004), and similar to the UNDP's approach when constructing the Human Development Index (Anand & Sen 2000). Ideally, this should be tested.

There are a number of approaches which health economists currently employ to understand the 'value' that individuals place on a health state: a standard gamble, a time trade-off, a rating scale, and more recently a discrete choice experiment (Brazier & Ratcliffe 2007). However, it has been argued that capability states (or capability sets) cannot be valued (Cookson 2005); while Sen rejected the use of choice or desires, and instead notes a preference for value judgment (Sen 1985; 1992). Such value judgments would avoid issues of adaptation as well; adaptation is where individuals may not recognise their own lack of capability because they have adapted to their situation (Menzel et al. 2002; Burchardt 2009). Sen advocates for objectivity (Sen, 2010), but notes that external and internal views of one's health (perception versus observation) can diverge (Sen, 2002). The application of this impartiality in the context of valuation as health economists require, would involve public health professionals or policy makers providing values for capabilities that individuals have identified as important (Lorgelly et al. 2010). Notably, this conflicts with the movement in health towards patient and public involvement in decision making (Coast, Smith & Lorgelly 2008b).

Despite this, capability indices do exist, there is currently a suite of instrument that resulted from the Investigating Choice Experiments for the Preferences of Older People (ICEPOP) project. The ICECAP-O is for older people (Coast et al. 2008), the ICECAP-A is for adults more generally (Al-Janabi, Flynn & Coast 2012) and the ICECAP-SCM (Sutton & Coast 2013) will be utilised in end of life care. These indices were developed by health economists for use in economic evaluations, and their valuation task used a best-worst scaling approach (Flynn et al. 2007) (where respondents are asked to only specify the attribute levels which they think are the best and worst), which they argue elicits 'values' (as Cookson (2005) suggests) rather than 'choices', because the elicitation exercise does not ask individuals to risk or sacrifice, as would be the case in a standard gamble or time trade-off exercise, respectively. The use of these instruments in economic evaluations is, however, in its infancy, and the jury is still out as to what role they will have in decision making; that is whether they can or will replace a QALY or alternatively supplement the standard instruments Coast, Smith & Lorgelly 2008a; 2008b). Despite this the benefits cannot be overlooked, and there is considerable scope to operationalize the capability approach to measure the effectiveness (and thus cost effectiveness) of public health interventions and programmes across all development settings.

Appendix 1

Vignette 1

Robert is 43 years old and lives, as he has always done, in the Calton, Glasgow. Robert's grandparents and parents are no longer alive. One of Robert's older brothers died last year of a heart attack. Robert smokes 20-30 cigarettes each day and suffers from bronchial (chest) problems, for which he receives medication. Robert also enjoys a few pints each day with his friends in the local pub. Robert does not take regular exercise as he becomes breathless with any form of exercise.

Robert says he expects to live for a 'few more years'.

Vignette 2

Robert is 27 years old and lives with his parents in Bearsden, Glasgow. All 4 of his grandparents are still alive and aged 80 years +. Robert eats fresh fruit and vegetables each day and takes moderate exercise at least three times each week. He particularly enjoys playing football with his mates. Robert currently has no ill health and would describe himself as being 'healthy'. Robert is slim and does not smoke. He occasionally drinks small amounts of alcohol.

Robert says he expects to live to around 95 years.

Appendix 2

Capabilities
"What you can do, not what you actually do"

Capabilities	*Version 1 - OCAP*	*Version 2*	*Version 3 – OCAP-18*
Life Being able to live to the end of a human life of normal length . . . ; not dying prematurely . . .	Given your family history, dietary habits, lifestyle and health status until what age do you expect to live?	Until what age do you expect to live, given your family history, dietary habits, lifestyle and health status?	Until what age do you expect to live, given your family history, dietary habits, lifestyle and health status?
Bodily Health Being able to have good health, including reproductive health; being adequately nourished . . . ; being able to have adequate shelter . . .	Does your health in any way limit your daily activities, compared to most people of your age? Do you eat fresh meat, chicken or fish at least twice a week? If not, why not? Are you able to have children? If not, why not? Is you current accommodation adequate or inadequate for your current needs? Are you prevented from moving home?	Does your health in any way limit your daily activities, compared to most people of your age? How often do you eat fresh fruit and vegetables? Why do you not eat 5 portions each day? Are you currently physically able to have children? If not why not? How suitable or unsuitable is your accommodation for your current needs?	Does your health in any way limit your daily activities, compared to most people of your age? How suitable or unsuitable is your accommodation for your current needs?
Bodily Integrity Being able to move freely from place to place; being able to be secure against violent assault, including sexual assault . . . ; having opportunities for sexual satisfaction and for choice in matters of reproduction	Are you prohibited from using any of the following: contraception, abortion, fertility treatment? Do you have sufficient opportunities to satisfy your sexual needs/desires? Please indicate how safe you feel walking alone in the area near your home (daylight and after dark): Have you ever been a victim of sexual/domestic/violent assault? How vulnerable do you feel to sexual /domestic/ violent assault in the future.	Are you prohibited from using any of the following: contraception, abortion, fertility treatment? Do you have sufficient opportunities to satisfy your sexual needs/desires? How safe do you feel walking alone in the area near your home? How likely to you believe it to be that you will be assaulted in the future (including sexual and domestic assault)?	How safe do you feel walking alone in the area near your home? How likely do you believe it to be that you will be assaulted in the future (including sexual and domestic assault)?
Senses, Imagination and Thought Being able to use the senses; being able to imagine, to think, and to reason--and to do these things in . . . a way informed and cultivated by an adequate education . . . ; being able to use imagination and thought in connection with experiencing, and producing expressive works and events of one's own choice . . . ; being able to use one's mind in ways protected by guarantees of freedom of expression with respect to both political and artistic speech and freedom of religious exercise; being able to have pleasurable experiences and to avoid nonbeneficial pain	I am free to express my political views I am free to practice my religion How often do you use your imagination/reasoning? Have you been able to enjoy your normal day to day activities? What is the highest educational or work related qualification you have?	I am free to express my views, including political and religious views In the past 4 weeks, how often have you been able to enjoy your normal day to day activities? What is the highest educational qualification you have?	I am able to express my views, including political and religious views. I am free to use my imagination and to express myself creatively (e.g. through art, literature, music etc).
Emotions Being able to have attachments to things and persons outside ourselves; being able to love those who love and care for us; being able to grieve at their absence, to experience longing, gratitude, and justified anger; not having one's emotional development blighted by fear or anxiety. . . .	How easy/difficult do you find it to enjoy the love, care and support of you immediate family? Do you find it easy/difficult to express feelings of love, grief, long, gratitude and anger? How difficult do you find it to make friends? Have you recently lost much sleep over worry? Have you recently felt under constant stain?	How easy/difficult do you find it to enjoy the love, care and support of your family and friends? How easy/difficult do you find it to express feelings of love, grief, long, gratitude and anger? How easy/difficult do you find it to make lasting friendships? In the past 4 weeks, how often have you lost sleep over worry? In the past 4 weeks, how often have you felt under constant strain?	At present how easy or difficult do you find it to enjoy the love, care and support of your family and friends? In the past 4 weeks, how often have you lost sleep over worry?
Practical Reason Being able to form a conception of the good and to engage in critical reflection about the planning of one's own life. (This entails protection for liberty of conscience.)	My idea of a good life is based on my own judgment. I have a clear plan of how I would like my life to be. How often do you evaluate how you lead your life and where you are going in life? Outside of work, have you recently felt that you were playing a useful part in things?	My idea of a good life is based on my own judgment. I have a clear plan of how I would like my life to be. In the past 4 weeks, how often have you felt that you were playing a useful part in things?	I am free to decide for myself how to live my life.
Affiliation Being able to live for and in relation to others, to recognize and show concern for other human beings, to engage in various forms of social interaction; being able to imagine the situation of another and to have compassion for that situation; having the capability for both justice and friendship. . . . Being able to be treated as a dignified being whose worth is equal to that of others.	I respect, value and appreciate other people. Do you tend to find it easy or difficult to imagine the situation of other people? Have you recently been thinking of yourself as a worthless person? Do you normally have at least one week's holiday away from home? If not, why not? Do you normally meet up with friends/family for a drink or a meal at least once a month? If not, why not? Outside of work, have you ever experienced discrimination because of your:	I respect, value and appreciate people around me. Do you normally meet up with friends/family for a drink or a meal at least once a month? If not, why not? In the past 4 weeks, how often have you been thinking of yourself as a worthless person? Outside of any employment, in your everyday life, how likely do you think it is that you will experience discrimination because of your: Race; Sexual orientation; Gender; Religion; Age; Health/disability?	I am able to respect, value and appreciate people around me. Are you able to meet socially with friends, relatives or work colleagues? Outside of any employment, in your everyday life, how likely do you think it is that you will experience discrimination?
Species Being able to live with concern for and in relation to animals, plants, and the world of nature.	I appreciate and value plants, animals and the world of nature.	I appreciate and value plants, animals and the world of nature.	I am able to appreciate and value plants, animals and the world of nature
Play Being able to laugh, to play, to enjoy recreational activities.	Have you recently been enjoying your recreational activities?	In the past 4 weeks, how often have you been able to enjoy your recreational activities?	In the past 4 weeks, how often have you been able to enjoy your recreational activities?
Control over one's life (A) *Political*: being able to participate effectively in political choices that govern one's life; having the rights of political participation, free speech and freedom of association . . . (B) *Material*: being able to hold property (both land and movable goods); having the right to seek employment on an equal basis with others . . .	I am able to participate in the political activities that affect my life if I want to. At work, have you recently felt that you were playing a useful part in things? Which of these applies to your home? Why have you not bought your home? How likely do you think it is that you will be stopped and searched by the police? When seeking work in the past, have you ever experienced discrimination because of your: Race; Sexual orientation; Gender; Religion; Age When seeking work in the future, how likely do you think it is that you will experience discrimination because of your: Race; Sexual orientation; Gender; Religion; Age To what extent does your work make use of your skills and talents? Do you tend to find it easy or diffcult to relate to your colleagues at work? At work, are you treated with respect?	I am able to participate in the political activities that affect my life if I want to. Which of these applies to your home? For which of the following reasons, if any, have you NOT bought your home? How likely do you think it is that within the next 12 months you will be 'stopped and searched' by the police when it is not warranted? In your current or future employment, how likely do you think it is that you will experience discrimination because of your: Race; Sexual orientation; Gender; Religion; Age; Health/disability To what extent are your skills and talents made use of either in or outside of any employment? How easy or difficult do you find it to relate to people?	I am able to influence decisions affecting my local area. Which of these applies to your home? For which of the following reasons, if any, have you NOT bought your home? In your current or future employment, how likely do you think it is that you will experience discrimination?

References

Al-Janabi, H., T. N. Flynn and J. Coast (2012). "Development of a self-report measure of capability wellbeing for adults: the ICECAP-A." *Quality of Life Research* **21**: 167-176.

Alkire, S. (2002). *Valuing Freedoms: Sen's Capability Approach and Poverty Reduction*. Oxford, Oxford University Press

Alkire, S., N. Kakwani and J. Silber (2008). The Capability Approach: Mapping Measurement Issues and Choosing Dimensions. *Many Dimensions of Poverty*. London, Palgrave-Macmillan.

Anand, P., G. Hunter, I. Carter, K. Dowding, F. Guala and M. van Hees (2009). "The development of capability indicators." *Journal of Human Development and Capabilities* **10**(1): 125-152.

Anand, P. and M. van Hees (2006). "Capabilities and achievements: An empirical study." *The Journal of Socio-Economics* **35** 268-284.

Anand, S. and A. Sen (2000). "Human development and economic sustainability." *World development* **28**(12): 2029-2049.

Aronson, J. (1994). "A Pragmatic View of Thematic Analysis." *The Qualitative Report* **2** (1): Spring.

Barbour, R. S., J. Kitzinger, R. Barbour and J. Kitzinger (1999). *Developing Focus Group Research: Politics, Theory and Practice*. London Sage.

Brazier, J. and J. Ratcliffe (2007). *Measuring and valuing health benefits for economic evaluation*, Oxford University Press, USA.

Burchardt, T. (2009). "Agency goals, adaptation and capability sets." *Journal of Human Development and Capabilities* **10**(1): 3-19.

Coast, J. (2004). "Is economic evaluation in touch with society's health values?" *British Medical Journal* **329**(7476): 1233-1236.

Coast, J., T. N. Flynn, L. Natarajan, K. Sproston, J. Lewis, J. J. Louviere and T. J. Peters (2008). "Valuing the ICECAP capability index for older people." *Social science & medicine* **67**(5): 874-882.

Coast, J., R. D. Smith and P. Lorgelly (2008a). "Should the capability approach be applied in health economics?" *Health Economics* **17**(6): 667-670.

Coast, J., R. D. Smith and P. Lorgelly (2008b). "Welfarism, extra-welfarism and capability: The spread of ideas in health economics." *Social Science & Medicine* **67**(7): 1190-1198.

Comim, F. (2008). Measuring capabilities. *The Capability Approach: Concepts, Measures and Applications*. F. Comim, M. Qizilbash and S. Alkire. Cambridge, Cambridge University Press**:** 157-200.

Cookson, R. (2005). "QALYs and the capability approach." *Health Economics* **14** (8): 817-829.

Cornwall, A. and C. Nyamu-Musembi (2004). "Putting the 'rights-based approach'to development into perspective." *Third World Quarterly* **25**(8): 1415-1437.

DeMaio, T. J., N. Bates, D. Willimack and J. Ingold (2006). Questionnaire Testing and Evaluation Methods for Censuses and Surveys, US Census Bureau.

Devidal, P. (2004). "Trading away human rights? The GATS and the right to education: a legal perspective." *The Journal for Critical Education Policy Studies* **2**(2): 22.

Drummond, M. F. (2005). *Methods for the economic evaluation of health care programmes*. Oxford, Oxford University Press.

Egan, M., A. Kearns, P. Mason, C. Tannahill, L. Bond, J. Coyle, S. Beck, F. Crawford, P. Hanlon and L. Lawson (2010). "Protocol for a mixed methods study investigating the impact of investment in housing, regeneration and neighbourhood renewal on the health and wellbeing of residents: the GoWell programme." *BMC Medical Research Methodology* **10**(1): 41.

EuroQol, G. (1991). "EuroQol: a new facility for the measurement of health related quality of life." *Health Policy* **16** 199-208.

Finch, J. (1987). "The Vignette Technique in Survey Research." *Sociology* **21**(1): 105-114.

Flynn, T. N., J. J. Louviere, T. J. Peters and J. Coast (2007). "Best-worst scaling: What it can do for health care research and how to do it." *Journal of Health Economics* **26**(1): 171-189.

Fox, J. and P. Tracy (1986). *Response: A Method for Sensitive Surveys*. Beverly Hills, CA, Sage.

Gosling, S. D., P. J. Rentfrow and W. B. Swann (2003). "A very brief measure of the Big-Five personality domains." *Journal of Research in Personality* **37**(6): 504-528.

Hyland, M. and S. Sodergren (1996). "Development of a new type of global quality of life scale, and comparison of performance and preference for 12 global scales." *Quality of Life Research* **5**(5): 469-480.

Klein, P. (1994). *An easy guide to factor analysis*, London: Routledge.

Lelli, S., F. Comim, M. Qizilbash and S. Alkire (2008). Operationalising Sen's capability approach: the influence of the selected technique. *The Capability Approach: Concepts, Measures and Applications*. Cambridge, Cambridge University Press.

Lewycka, S., C. Mwansambo, P. Kazembe, T. Phiri, A. Mganga, M. Rosato, H. Chapota, F. Malamba, S. Vergnano and M. Newell (2010). "A cluster randomised controlled trial of the community effectiveness of two interventions in rural Malawi to improve health care and to reduce maternal, newborn and infant mortality." *Trials* **11**(1): 88.

Lorgelly, P. K., K. Lawson, E. A. L. Fenwick and A. H. Briggs (2010). "Outcome measurement in economic evaluations of public health interventions: a role for the capability approach." *International Journal of Environmental Research and Public Health* **7**.

Marmot, M. (2005). "Social determinants of health inequalities." *The Lancet* **365**(9464): 1099-1104.

Menzel, P., P. Dolan, J. Richardson and J. A. Olsen (2002). "The role of adaptation to disability and disease in health state valuation: a preliminary normative analysis." *Social Science & Medicine* **55**(12): 2149-2158.

Morris, R. and V. Carstairs (1991). "Which deprivation? A comparison of selected deprivation indexes." *Journal of Public Health* **13**(4): 318-326.

Muffels, R. and B. Headey (2009). Testing Sen's capability approach to explain objective and subjective well-being using German and Australian Panel Data? *HILDA Survey Research Conference 2009*. Melbourne University.

Nussbaum, M. C. (2000). *Women and Human Development: The Capabilities Approach*. New York, Cambridge University Press.

Nussbaum, M. C. (2011). *Creating Capabilities: The Human Development Approach*. Harvard University Press.

Petersen, A. R., A. Petersen and D. Lupton (1996). *The new public health: health and self in the age of risk*, Sage Publications Ltd.

Presser, S., J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin and E. Singer (2004). *Methods for Testing and Evaluating Survey Questionnaires*. New York Wiley.

Prieto, L., H. Thorsen and K. Juul (2005). "Development and validation of a quality of life questionniare for patients with colostomy of ileostomy." *Health and Quality of Life Outcomes* **3** 62.

Robeyns, I. (2003). "Sen's capability approach and gender inequality: selecting relevant capabilities." *Feminist Economics* **9**(2-3): 61-92.

Rychetnik, L., M. Frommer, P. Hawe and A. Shiell (2002). "Criteria for evaluating evidence on public health interventions." *Journal of Epidemiology and Community Health* **56**(2): 119-127.

Schimmack, U., P. Radhakrishnan, S. Oishi, V. Dzokoto and S. Ahadi (2002). "Culture, personality, and subjective well-being: Integrating process models of life satisfaction." *Journal of Personality and Social Psychology* **82**(4): 582.

Schimmack, U., S. Oishi, R. M. Furr and D. C. Funder (2004). "Personality and life satisfaction: A facet-level analysis." *Personality and Social Psychology Bulletin* **30**(8): 1062-1075.

Sen, A. (1985). *Commodities and Capabilities*. New York, Elsevier.

Sen, A. (1993). Capability and well-being. *The Quality of Life*. M. C. Nussbaum and A. Sen. Oxford, Claredon Press**:** 30-54.

Sen, A. (1999). *Development as Freedom*. Oxford, Oxford University Press.

Sen, A. (2002). "Health: perception versus observation." *British Medical Journal* **324**(7342): 860-861.

Sen, A. K. (1992). *Inequality Re-examined*. Oxford, Clarendon Press.

Sen, A. K. (2010). The place of capability in a theory of justice. *Measuring Justice: Primary Goods and Capabilities*. H. Brighouse, I. Robeyns. Cambridge, Cambridge University Press: 239-253.

Sheatsley, P., P. Rossi, J. Wright and A. Anderson (1983). Questionnaire construction and item writing. *Handbook of Survey Research*. New York, Academic Press.

Shiell, A., P. Hawe and L. Gold (2008). "Complex interventions or complex systems? Implications for health economic evaluation." *Bmj* **336**(7656): 1281-1283.

Silverman, D. (2009). *Doing qualitative research*, Sage Publications Ltd.

Simon, J., P. Anand, A. Gray, J. Rugkåsa, K. Yeeles and T. Burns (2013). "Operationalising the capability approach for outcome measurement in mental health research." *Social Science & Medicine* **98**: 187-196.

Smith, R. D. and M. Petticrew (2010). "Public health evaluation in the twenty-first century: time to see the wood as well as the trees." *Journal of Public Health* **32**(1): 2-7.

Sridharan, S., B. Campbell and H. Zinzow (2006). "Developing a stakeholder-driven anticipated timeline of impact for evaluation of social programs." *American Journal of Evaluation* **27**(2): 148-162.

Storck, A. J., K. B. Laupland, R. R. Read, M. W. Mah, J. M. Gill, D. Nevett and T. J. Louie (2006). "Development of a Health-Related Quality of Life Questionnaire (HRQL) for patients with Extremity Soft Tissue Infections (ESTI)." *BMC Infectious Diseases* **6** 148.

Sutton, E. J. and J. Coast (2013). "Development of a supportive care measure for economic evaluation of end-of-life care using qualitative methods." *Palliative medicine*: 0269216313489368.

Tashakkori, A. and C. Teddlie (1998). *Mixed methodology: Combining qualitative and quantitative approaches*, Sage Publications, Inc.

Vergunst, F., C. Jenkinson, T. Burns and J. Simon (2014). "Application of Sen's Capability Approach to Outcome Measurement in Mental Health Research: Psychometric Validation of a Novel Multi-dimensional Instrument (OxCAP-MH)." *Human Welfare* **3**: 1-4.

Wamcata, J., M. Krautgartner, J. Berner, R. Alexandrowicz, A. Unger, G. Kaiser, B. Marquart and M. Weiss (2005). "The Carers' Needs Assessment for Dementia (CNA-D): Development, validity and reliability." *International Psychogeriatrics* **17** (3): 393-406.

Ware, J. E., Jr. and C. D. Sherbourne (1992). "The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection." *Medical Care* **30** (6): 473-483.

Wilkinson, R. G. and K. E. Pickett (2006). "Income inequality and population health: a review and explanation of the evidence." *Social Science & Medicine* **62**(7): 1768-1784.

Willis, G. B. (2005). *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. London, Sage.

Table 1: Respondents Characteristics (Descriptive Statistics and Frequencies; N=198)

	Count* or Mean	Percentage or Std Dev
Age	45.84	16.13
Gender		
Male	72	37.3
Female	121	62.7
Marital Status		
Married	58	30.1
Living as married	18	9.3
Separated (after being married)	14	7.3
Divorced	22	11.4
Widowed	16	8.3
Never married	65	33.7
Number of children	0.46	0.81
Employment status		
Working full time (30 or more hours per week)	98	50.8
Working part time (8 to 29 hours per week)	13	6.7
Full time student	9	4.7
Retired	35	18.1
Unemployed	15	7.8
Not working for other reason	23	11.9
Annual Household Income		
£0 (nothing)	2	1.0
£1 to £9,999 per year	48	24.5
£10,000 to £19,999 per year	44	22.4
£20,000 to £29,999 per year	26	13.3
£30,000 to £39,999 per year	22	11.2
£40,000 to £59,999 per year	21	10.7
£60,000 or more per year	18	9.2
Prefer not to answer	8	4.1
Don't know	7	3.6
Highest Educational Attainment		
Postgraduate degree	23	11.7
First degree	41	20.8
Higher education below degree	25	12.7
Highers/A Levels or equivalent	19	9.6
Standard Grades 1-3/GCSEs or equivalent	18	9.1
Standard grades 4-7/CSE or equivalent	7	3.6
Foreign or other qualification	6	3.0
No qualification	47	23.9
Don't know	11	5.6
Ethnicity		
White	188	97.4
Mixed ethnic group	1	0.5
Asian or Asian British	2	1.0
Black or Black British	1	0.5
Other ethnic group	1	0.5

Table 1: Respondents Characteristics continued…

	Count* or Mean	Percentage or Std Dev
Religious Denomination		
Church of England	1	0.5
Church of Scotland	53	27.0
Muslim	2	1.0
Other Christian	8	4.1
Roman Catholic	54	27.6
Another religion	4	2.0
None	69	35.2
Prefer not to answer	5	2.6
Deprivation decile		
1	1	0.5
2	13	6.7
3	0	0.0
4	6	3.1
5	6	3.1
6	7	3.6
7	16	8.2
8	9	4.6
9	35	17.9
10	102	52.3
Health/QoL		
EQ-5D score (0: dead, 1: full health)	0.76	0.28
Global QoL (0: no QoL, 100: perfect QoL)	69.55	19.86

* Note the counts may not sum to 198 as not every question was completed by every respondent, however the percentage takes these missing responses into account

Table 2: Summary test statistics (chi-sq and F tests) for differences in individual capability questions by gender, age, deprivation and income groupings

	Gender	Age	Deprivation	Income
Life				
Life expectancy (deviations)[a]	5.514**	0.137	0.352	6.655**
Bodily Health				
Daily activities[b]	0.850	11.655**	8.374*	39.831**
Suitable Accommodation	2.895	3.906	4.409	16.120
Bodily integrity				
Neighbourhood safety	8.559	5.355	29.991**	12.314
Potential for assault	10.755*	9.548	8.202	13.601
Senses imagination and thought				
Freedom of expression	4.535	6.006	4.523	14.734
Imagination and creativity	6.717	14.895	6.817	14.304
Emotions				
Love and support	4.347	13.616	14.859	20.056
Losing sleep	3.244	5.223	10.080	21.750*
Practical Reason				
Planning one's life	5.947	6.989	14.423	12.382
Affiliation				
Respect and appreciation	7.121	5.807	1.527	14.450
Social networks[b]	0.037	2.418	8.025*	13.458**
Discrimination	2.586	18.569*	5.514	16.180
Species				
Appreciate nature	0.764	2.017	10.363	12.133
Play				
Enjoy recreation	0.209	2.584	11.447	25.648*
Control over one's environment				
Influence local decisions	2.452	12.778	14.869	31.934**
Property ownership[b]	1.912	2.057	14.602**	55.575**
Employment discrimination	2.218	3.302	5.501	10.039

[a] as this is a continuous variable, the test statistic employed is an F-test, all other variables are categorical and as such a chi-squared test is used.

[b] note these questions have binary answers, as such they have fewer degrees of freedom

** significant at 1% level; * significant at 5% level

Table 3: Descriptive statistics for the capability index by interest group

	Mean	Std Dev	Minimum	Maximum	p-value
Gender					
Male	12.53	2.41	5.50	17.75	
Female	12.40	2.62	3.00	17.25	0.761
Age					
Under 40	12.50	2.50	3.00	17.75	
40 to 60	12.30	2.65	4.50	17.25	
Over 60	12.70	2.42	6.50	16.00	0.772
Deprivation					
deciles 1 to 6	13.45	1.79	8.50	16.50	
deciles 7 to 9	12.88	2.43	4.50	17.75	
decile 10	11.92	2.66	3.00	17.25	0.006*
Income					
less than £10k	10.73	2.70	4.50	14.75	
£10k to £19k	11.85	2.66	3.00	17.25	
£20k to £40k	13.25	1.95	7.50	16.50	
more than £40k	13.94	1.54	10.50	17.75	<0.001*

Note: Significant differences are indicated by an asterisk.

Table 4: Multiple regression results for the capability index

	Coefficient	Std error	p-value
Gender			
Female	0.066	0.395	0.868
Age			
40 to 60	0.304	0.416	0.467
Over 60	1.134	0.626	0.072
Deprivation			
deciles 7 to 9	-0.100	0.578	0.863
decile 10	-0.549	0.542	0.313
Income			
£10k to £19k	1.080	0.527	0.042*
£20k to £40k	2.599	0.531	0.000*
more than £40k	3.239	0.574	0.000*

Note: males, under 40 years old, residing in decile 1 to 6 areas, and earning less than £10,000 per year is the reference category. The dependent variable is the capability index. Significant differences are indicated by an asterisk.
N= 155, R^2=0.258

Table 5: Which do interviewees value more: capability [C] or functioning [F]?

Interviewee	Adequately nourished	Expressing views	Love, care & support	Planning of one's own life	Influencing decisions affecting local area
1	C	C	C	C	F
2	C	C	F	F	C
3	C	C	C	F	C
4	C	C	F	C	C
5	F	C	F	F	C
6	C	C	F	C	F
7	F	C	C	F	C
8	C	C	C	C	C
9	C	C	C	C	C
10	C	C	C	C	C
11	F	F	F	F	C
12	F	C	C	C	C
13	C	C	C	C	C
14	F	C	C	C	C
15	C	C	F	C	C
16	C	C	NOT SURE	C	C
17	C	C	F	BOTH	C
18	C	F	C	F	F

Figure 1: Life Capability: Until what age do you expect to live, given your family history, dietary habits, lifestyle and health status? (difference between actual life expectancy, given each respondent's age and gender – using life tables for Glasgow City, and elicited life expectancy)

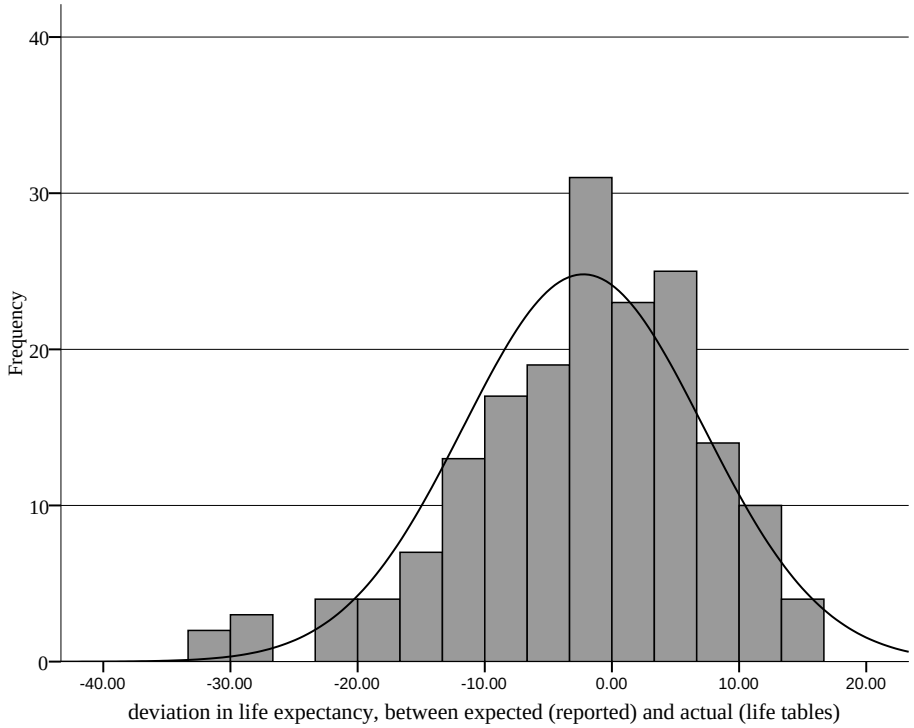

Figure 2: Affiliation Capability: Outside any employment, in your everyday life, how likely do you think it is that you will experience discrimination?

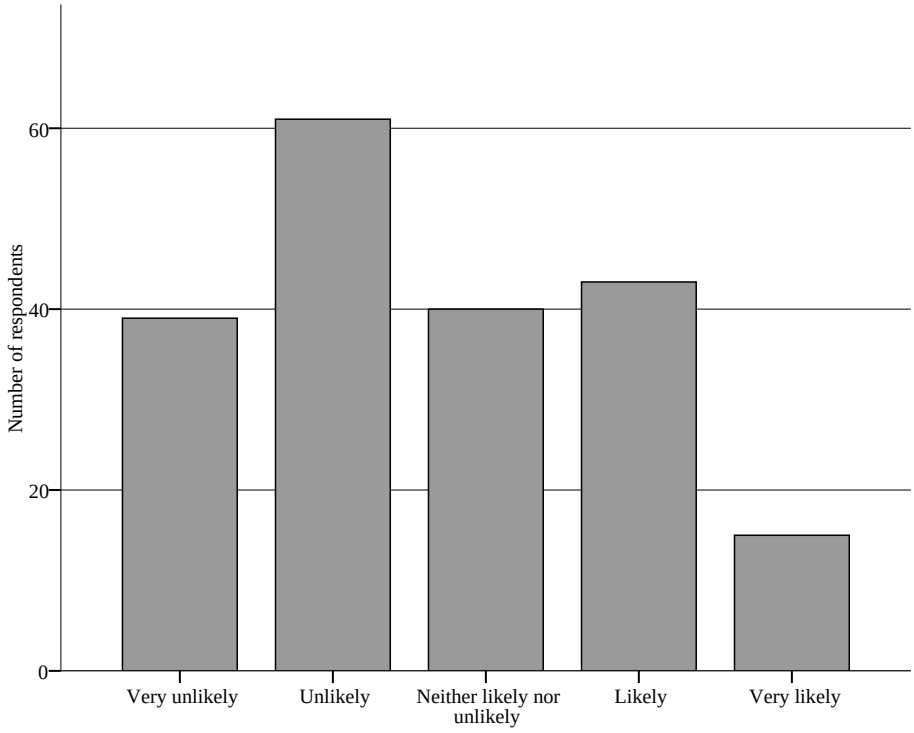

Figure 3: Histogram of the capability index

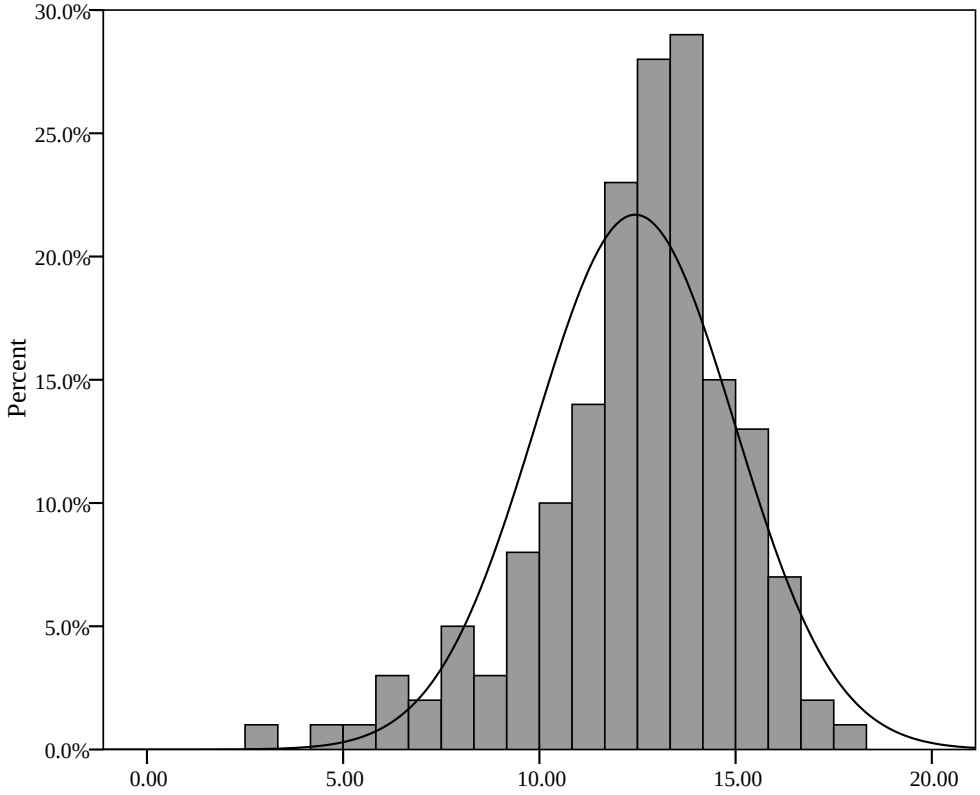

Figure 4: Scatterplot of relationship between EQ5D and capability index

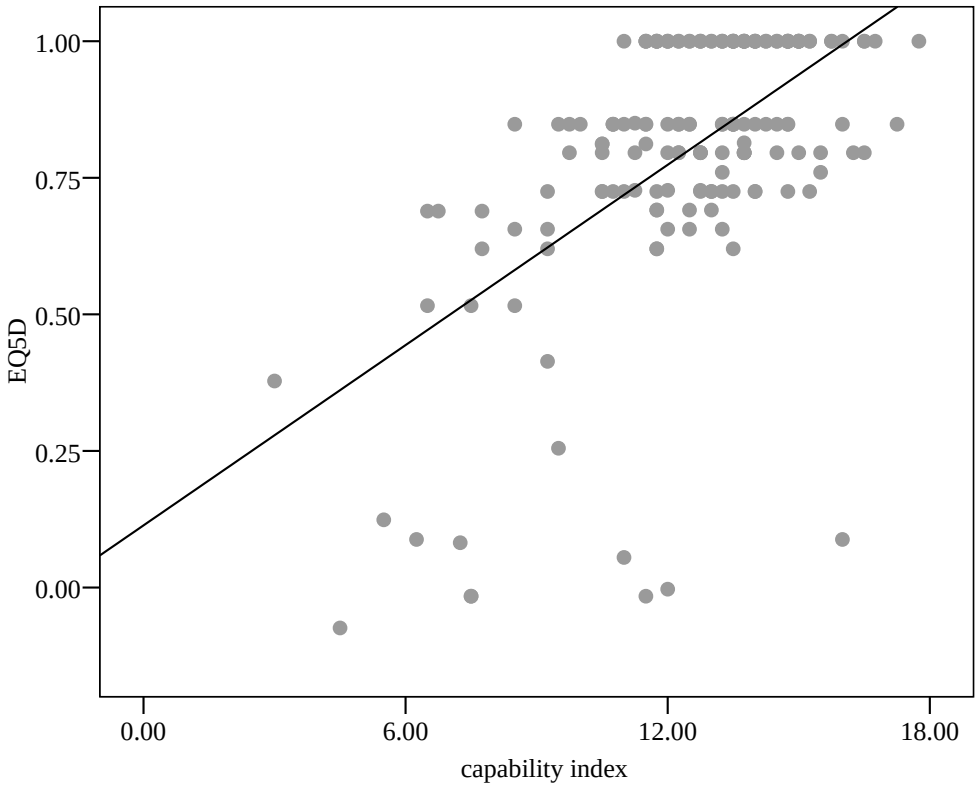

Figure 5: Scatterplot of relationship between QoL and capability index