



Development
Progress

Project Note

04

Measuring what matters

The role of well-being methods in development policy and practice

Allister McGregor, Sarah Coulthard and Laura Camfield



1. Introduction

Since the publication of the Final Report of the ‘Commission on Measuring Economic Performance and Social Progress’ in 2009 (hereafter The Stiglitz Commission: Stiglitz, Sen and Fitoussie, 2009) there has been a tremendous upsurge of interest in and initiatives to develop measures of human well-being as a yardstick of societal progress. The basic message of the report was that the measurement of development in terms of income is inadequate and has misguided public policy and practice in both developed and developing countries. The focus on income measures encourages an unbalanced focus on economic growth and policies that, at their worst, have proven to be unsustainable, unequal and not poverty-reducing. In order to assess whether societal development equates with progress¹ we need measures that tell us whether development is improving

human well-being or not. The main recommendation and challenge to policy-makers, academics and civil society actors set out by the Stiglitz Report was to ‘shift emphasis from measuring economic production to measuring people’s well-being’ (2009: 12).

The initiatives to develop a policy relevant concept of well-being that have followed this, and the number that were already in progress before the publication of that report (Gough and McGregor; 2007; Anand et al., 2009), are many and varied. Indeed there is something of a cacophony of competing conceptualisations, methodologies and ideologies that underpin the various initiatives to measure progress, well-being or happiness. Despite the potential for confusion, however, there is a belief that there is a growing consensus around what

This and other Development Progress materials are available at developmentprogress.org

Development Progress is an ODI project that aims to measure, understand and communicate where and how progress has been made in development.

ODI is the UK’s leading independent think tank on international development and humanitarian issues. Further ODI materials are available at odi.org.uk

Indonesian children near the town of Glebruk. Photo: © US Navy.

¹ Or ‘good development’, as Robert Chambers (2005) puts it.

needs to be measured and how it might be measured (Durand, 2013; OECD, 2013a).

In this paper we explore how the measurement of human well-being might contribute to making development policy and practice more effective. While most of the current initiatives are to develop measures of well-being at the level of the nation state (for the purposes of national policy deliberation or for between state comparisons), this paper focuses on a less prominent discussion about how the concept of well-being might be made relevant for the development frontline: for development policy and practice at the level of the project or programme (Rojas, 2008; Camfield and McGregor, 2009). In particular it considers how a comprehensive assessment of human well-being, and particularly one that integrates consideration of subjective well-being,² might be adopted into the routines of development.

This paper is particularly concerned with poverty policy and it emphasises that the interests of poor people themselves must be recognised as fundamental for the success or failure of these policies. But this concern is not confined just to poverty policies; it is about all policies that affect poor and marginalised people (in other words, most public policies). More broadly, understanding the interests and motivations of poor people, who are usually excluded or disenfranchised in elite-dominated policy processes, is vital to producing development and progress that is more oriented to socially just outcomes (Devereux and McGregor 2014). Knowing what people's interests are and what goals drive them is a challenge from the very outset of the policy process. Interests play a role in determining how the problem that policy will address is defined and the influence of different and differentially influential interests continues through the successive stages of the policy process.

While participatory development approaches have made some headway in articulating the types of thing that matter for people in their own development, producing lists that are effectively criteria for well-being, this is only one part of what might be required in an assessment of what really matters for people. There are three different things that must be done in order to make the exploration of 'what matters for people' relevant for policy and practice.

- The *first* is to identify systematically what is important to people for them to live their lives well, and to do so in a way that is universally comprehensible but is nevertheless sensitive to particular social, economic and cultural contexts (Gough and McGregor, 2004; McGregor 2004).
- The *second* is to find ways of assessing how well people are doing in their achievements in respect of the things that they regard as important for them to live well.

- The *third* is to establish ways of understanding how the different things that are important for well-being relate to each other. This may involve understanding how they are prioritised and what trade-offs may exist between them. From the policy perspective this relates to the challenge of establishing weightings in respect to the different things that matter.

This paper is organised around a discussion of approaches to each of these three tasks, but before we move on to discuss each in turn it is first necessary to review the current state of the art in the measurement of well-being field.

2. The different approaches and conceptual foundations for the measurement of human well-being

In a recent paper, Melamed (2011) points out that there are often disconnects between the views and priorities of poor people and those involved in development decision-making. This suggests that in the interests of both greater aid effectiveness and also improved transparency and accountability in development practice, there is a need for methods and practices that would better enable mutual comprehension. As she and her co-authors argue in a subsequent paper developing these ideas, 'If the drive for results and value for money is to deliver improved benefits for poor people, and better evidence for politicians trying to justify aid spending to taxpayers, we need an improved metric that allows us to measure outcomes in a standard way, and identify which – of the possible outcomes that aid spending could achieve – poor people would value the most. To be effective, this metric needs to be communicated in a way that is easy to use for decision makers in the development business (Melamed et al., 2012).'³

The struggle to have the idea accepted that people might participate in their own development has been a long one and has gained some success (Chambers, 1997). It has been taken on board at a level of high rhetoric and then in some particular set-piece studies (e.g. 'Voices of the Poor')⁴ and development routines such as the staged elements that have produced Poverty Reduction Strategy Papers (PRSPs – see Booth 2005 for a critical review). But knowledge generated by participatory means is not consistently incorporated in everyday policy and practice routines. Participatory methods may or may not be used at different stages in the policy or the project cycle and the data that they generate tends to have less traction than other forms of data in policy or project decision-making processes (Eyben, 2013). In short, participatory methods have not made sufficient headway in the transition to becoming an accepted, legitimate and routine element of development policy and

2 Subjective well-being refers to a person's own life experiences and evaluations about their quality of life; it includes emotional reactions and (dis)satisfactions with life and is subjective because it concerns the perspective of the individual in question (the subject) (Diener, Oishi and Lucas, 2009).

3 See project website <http://www.odi.org.uk/programmes/development-progress/valuing-progress>

4 'Voices of the poor' was a World Bank funded participatory poverty assessment that involved over 40,000 people across 50 countries (see Narayan et al 2000).

practice, although they have had more traction within specific NGOs in relation to needs assessment, planning and evaluation (Catley et al., 2008).

However, the basic message that it is necessary to take better account of what people themselves say is important to them has received renewed impetus from the Post-Stiglitz Commission measuring well-being movement. In a major international conference in 2012 on the measurement of progress and well-being, the Chief Statistician of the OECD felt compelled to state that,

'We are witnessing a convergence in our understanding of well-being with a common core set of well-being dimensions, and national priorities reflected in more specific domains and measures' (Durand, M. Plenary Address, 4th OECD Global Forum, New Delhi 2012).

Perhaps the first and main point of convergence is the acceptance of the view that human well-being must be understood as 'multi-dimensional' and cannot be captured adequately by considering a single measure such as 'happiness'. This runs counter to the work of the happiness economists such as Layard, Oswald and Frey and Stutzer whose arguments imply that a single happiness score contains enough information for it to be relied on in policy decision-making. This school of thought suggests that it is not necessary to know the detail of what lies below a person's feeling of well-being or happiness and that a single happiness or utility score can adequately capture the underlying complexity, contradiction and fluidity of human judgements about the quality of their lives across its whole range. This is not a widely accepted view when considering the possibility of the detailed use of well-being measures in particular policy spheres. Although a single overarching concept and measure of subjective well-being may have some uses as a broad brush comparator or as a variable to slot into regression equations, it cannot provide the richness of information required for most policy purposes (see also Alexandrova, 2005).

A second fundamental point in the growing consensus is that it is necessary to take account of both objective and subjective aspects of well-being in some kind of integrated framework (UK ONS, 2011; OECD, 2011a; UN, 2012). It is important to know not just what people have or have achieved in an objectives sense, but also to understand how they evaluate their achievements. These evaluations are likely to be as, if not more important for understanding why people then act in the ways that they do.⁵

Having agreed that there is some consensus that well-being is multi-dimensional and that both objective and subjective information are important we do however then move into a much more indeterminate discussion about which dimensions matter and how we are to define these.

This discussion hinges around which (or whose) framework is to be adopted. Currently there are a large and growing number of different multi-dimensional frameworks and each offer slightly different lists of what are proposed to be the set (or list) of universal dimensions.⁶

There are two main ways in which these lists have been established: one is in a 'top-down' manner in which the particular ingredients of what is required for well-being are identified from a particular philosophical position, conceptual framework, or ideology; the other is in a 'bottom-up' way, in which the 'list' of what is required for well-being is built up from observation or engagement with the people whose well-being you are concerned to understand. The desire for universality has tended to favour the 'lists' produced by top-down approaches and most of the multi-dimensional frameworks on offer arise from particular theoretical or philosophical positions.

Acknowledging the tension and the politics that lie between these two approaches, Amartya Sen had declined to propose such a list, arguing that any such list should arise from a process of public reasoning and deliberation within a particular polity. However many of his fellow travellers have been less hesitant. Martha Nussbaum, for example, has argued that there are 10 universal capabilities (what she calls 'spheres of existence').⁷ The Human Development Index has provided another multi-dimensional list based on Sen's capabilities but like the later and much lauded Multi-dimensional Poverty Index (MPI) developed by Alkire and Foster (2011), this is not a multi-dimensional *well-being* measure. It is a measure of poverty that deals with only a limited range of things that might be considered to represent poverty or refer to the human conditions which reproduce poverty. In the MPI, as with the Human Development Index, there are only three 'dimensions' (health, education and standard of living), and the rationale for these three is drawn from the human development and capabilities framework. In the MPI these three dimensions are then assessed using 11 objective indicators.

The 'Voices of the Poor' study conducted by the World Bank provides an interesting example that falls between 'top-down' and 'bottom-up' approaches (Narayan et al., 2000). In a study that took evidence from over 60,000 poor people in many countries of the world they confirmed that people viewed and experienced poverty as multi-dimensional. From this relatively 'bottom-up' perspective they assert that there are eight dimensions of experienced poverty: 1) material well-being (i.e. having enough food, assets, work); 2) bodily well-being (health, appearances and physical environment); 3) social well-being (being able to care for, bring up, marry and settle children); 4) self-respect and dignity; 5) peace, harmony, good relations in the family/community; 6) security (civil peace, a physically

5 The issue of 'adaptation' or 'adaptive preferences' (where people adjust their subjective assessments in relation to their objective circumstances) is discussed later in the paper.

6 For a review of the different domains of various prominent frameworks see Samman (2007).

7 Life, bodily health, bodily integrity, senses, imagination and thought, emotions, practical reason, affiliation, other species, play, control over one's environment.

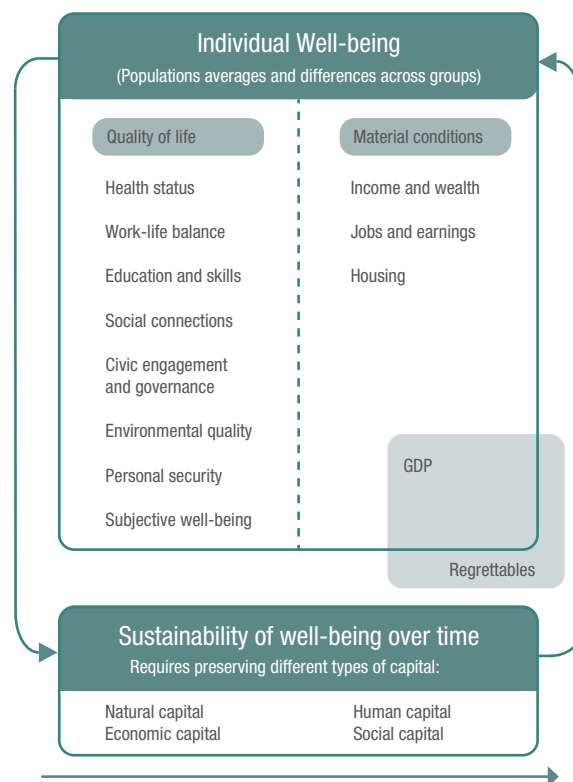
safe and secure environment; personal physical security; security in old age; confidence in the future); 7) freedom of choice and action; and 8) psychological well-being (peace of mind, happiness, harmony including a spiritual life and religious observance).

These domains/dimensions are broadly similar to those that are identified in the many well-being frameworks arising out of social psychology. For example Cummins (1998) has spearheaded work in this area and he and his colleagues in many countries have carried out extensive testing of the relevance and meaning of the various possible domains. The Personal Well-being Index that is corroborated by an International Working Group and is one of the most globally prominent subjective well-being frameworks proposes seven domains that contribute to Quality of Life universally (material well-being, health, productivity, intimacy, safety, community, emotional well-being – spiritual well-being was added later). In this approach these are all assessed subjectively. Other prominent frameworks arising out of social psychology or health such as the WHOQOL (World Health Organisation Quality of Life) or the SF-36 have between four and eight domains.

Coming from a tradition in social policy that focuses on human needs, Doyal and Gough (1991) suggest that well-being is underpinned by only two fundamental and universal human needs and these are for health and autonomy. However, in a sophisticated schema, they propose that these two basic needs are met in different social contexts through achievements in respect of 11 universal intermediate needs: adequate nutritional food and water, adequate protective housing, non-hazardous work and physical environments, appropriate health care, security in childhood, significant primary relationships, physical and economic security, safe birth control and childbearing, and appropriate basic and cross-cultural education (Doyal and Gough, 1991).

The Stiglitz Commission itself also proposed a ‘multidimensional’ framework to shape the measurement of well-being and this argued that there were eight ‘dimensions’ that needed to be taken into account (material living standards, health, education, personal activities, political voice, personal relationship, environment and physical and economic security). The OECD’s ‘How’s Life’ Framework (2011a),⁸ which is an element of their Better Life Initiative, is a direct descendent of the Stiglitz Commission thinking but expands the number of ‘dimensions’ to eleven (11). As is shown in Figure 1, these 11 dimensions are divided between two ‘pillars’ that are labelled ‘Quality of Life’ and ‘Material Conditions’. The Quality of Life pillar contains mainly human development dimensions but draws on other frameworks and (somewhat confusingly) includes a dimension labelled ‘Subjective Well-being’.⁹ A third and important element of the How’s Life

Figure 1: The OECD Well-being Framework: Better Life Initiative



Source: OECD (2013c).

Framework is labelled ‘sustainability’ and this elaborates the context for the reproduction of material conditions and quality of life. It identifies four types of ‘capitals’ that are seen as providing the enabling relationships through which present well-being can be translated (or not) into future well-being outcomes. This third element provides the dynamic that is necessarily part of any understanding of how human well-being is produced and how it is distributed across populations in a society.

3. Clarifying frameworks: dimensions, domains and indicators

It is a source of confusion for the further development of well-being measures that this range of frameworks are broadly similar but nevertheless make different claims about ‘dimensions’ and ‘domains’. At first glance the interchangeable use of ‘dimensions’ and ‘domains’ may look like a trivial semantic problem but we contend here that it is a more important source of confusion for the development of well-being measurement than is recognised. The interchangeable use is driven by a lack of conceptual clarity that will ultimately inhibit the successful, multi-disciplinary development of these measures.

⁸ <http://www.oecd.org/statistics/howslife.htm>

⁹ In the OECD Guidelines on Subjective Well-being (2013b) they emphasise that their interest in this domain is concerned with subjective well-being assessments that are ‘global’ in character (life satisfaction as a whole, happiness, eudaimonic well-being) and distinguish these from the types of approaches used in broader Quality of Life assessments, which are often concerned with domain satisfactions.

The OECD framework provides a good illustration of how there can be confusion between what are ‘dimensions’ and what are ‘domains’. As we have noted some frameworks refer to multiple ‘dimensions’, while other approaches refer to these as ‘domains’ (see Schmidt and Bullinger, 2005). In the Quality of Life Literature, arising out of the fields of health, social psychology and sociology, the term ‘dimension’ is used to distinguish between objective and subjective aspects of life: thus they refer to an ‘objective dimension of quality of life’ and a ‘subjective dimension of quality of life’ (see Axford et al., 2014). This is not the same sense in which the term dimension is used in the ‘multi-dimensional poverty’ and international development literatures. In those literature the term is used to encompass dimensions of poverty beyond income or consumption and is therefore closer to the terminology of ‘domains’ or ‘domains of life’ (Rojas, 2004) used in quality of life approaches. The confusion arises around what kind of data is required for assessment: some of these other dimensions can be regarded as ‘objective’, while others can be considered more ‘subjective’, while others still are social in character. An important difficulty with the use of the term ‘dimension’ in ‘multi-dimensional’ approaches to well-being, poverty or development is that these mix up different types of things in ways that confound empirical study (sometimes we are seeking to measure an object, sometimes a relationship, and sometimes a feeling – often it is not clear which it is).

The reason for the different uses of the terms and the numbers of dimensions or domains arises because each of these frameworks has been developed from a different intellectual tradition and because they are being developed with different purposes and audiences in mind. If, for example, we take the frameworks that arise out of social psychology and those that are applied in health (such as the WHOQOL) it is not surprising that they are relatively poorly developed in terms of their interest in and appreciation of social and economic factors that affect a person’s overall well-being. Those who have developed these frameworks are primarily interested in what is happening inside people’s heads either to assess their mental health or decide on an appropriate health treatment. In this respect they are able to limit their frameworks to identifying ‘domains’ within the dimension of subjective well-being that are most germane to health.

Those developing multi-dimensional poverty frameworks are primarily interested in providing a view of poverty that is expanded beyond a narrow income or consumption measure and their intended audience is a set of poverty policy-makers who have been long wedded to a limited range of objective indicators.¹⁰ It is not a surprise therefore that their framework would be limited both in terms of what they want to include in their framework and what measures they want to use to provide an analysis of the situation. There has been considerable neglect until recently in economics and other disciplines of psychological

or cognitive issues and physical health has been a matter of more concern than mental well-being. At the conceptual level these frameworks have been constrained both by the hegemony of the idea of ‘parsimony’ (unwillingness to accept notions of complexity) and also by the limits of their mathematical manipulation (too many variables produce intractable mathematical problems).

The purpose of frameworks such as that developed by the OECD is much grander, more ambitious and more complex than that of the frameworks that arise out of health and social psychology or those concerned with multi-dimensional poverty. Since the purpose of such a framework is to measure development or societal progress in terms of whether it is improving human well-being, it cannot be confined just to what is happening to the poor (albeit that providing an insightful analysis of poverty dynamics and distributions may be one important by-product of such an approach). Nor can it be interested only in measuring well-being in terms of what is happening inside peoples’ heads; it is interested in what is happening to their bodies also and to what is happening to their social relationships. It is this combination of considerations of the material aspects of well-being, the assessments of perceived quality of life and the relationships that give these their dynamic, that suggest a distinctive and three dimensional framework.

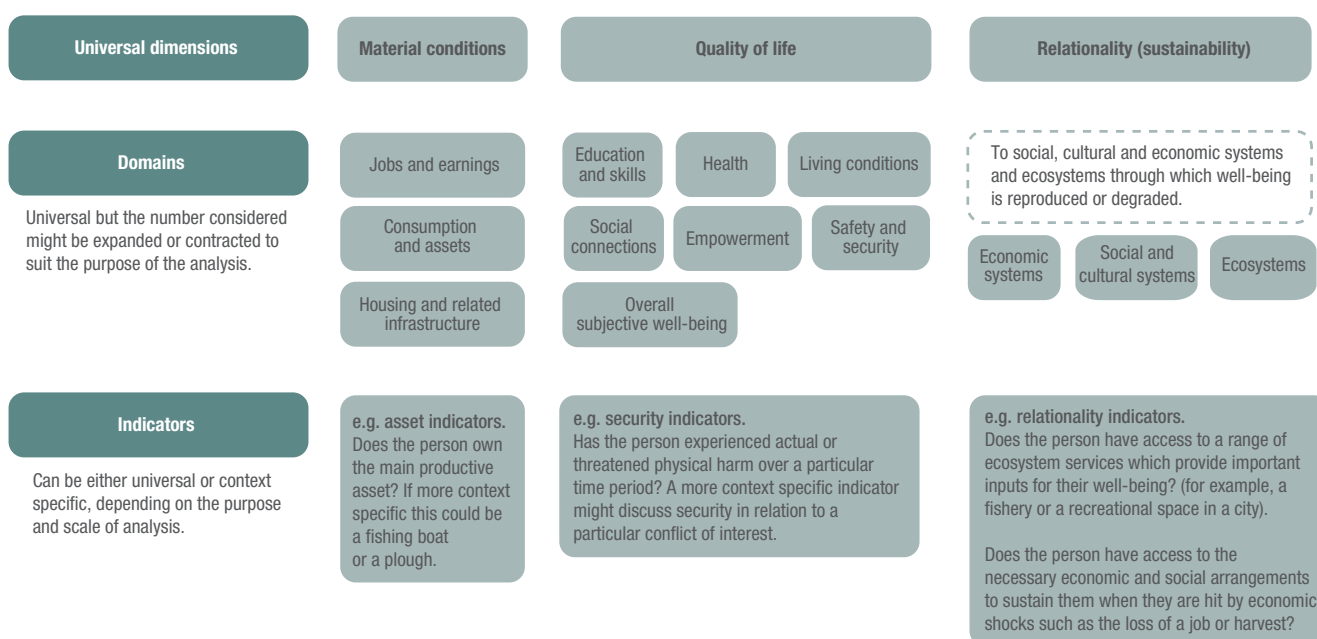
4. Using a framework to structure discussion of the measurement of well-being in development policy and practice

When we compare the many multi-dimensional frameworks currently that are on offer as a means to assess well-being or quality of life we find that most operate with some variant of three dimensions: a dimension that refers primarily to the material conditions of life; a dimension that refers to a broader notion of a person’s quality of life, which includes human development components such as education and health, but also extends to ‘conditions of being’ like security and subjective well-being; and a third dimension that can be understood as a relational dimension, which refers to the relationships that a person must be able to enter into in order to continue meeting the needs (or achieving the ‘functionings’) that are important for their well-being (for example, the relationships in society that allow them to act with autonomy or the relationships amongst friends and family that provide care and love).

Of the frameworks that are on offer and that we have discussed thus far the OECD framework comes closest to capturing this rich essence and also most closely relates to the purpose set out in this paper: to develop ways of assessing and measuring whether development interventions are having a positive impact on what matters to people for their well-being. The audience for which it is intended is also similar: policy-makers and

¹⁰ These are usually objectively verifiable indicators of deprivation(s).

Figure 2: Dimensions, domains and indicators (illustrated by recent work based on the OECD How's Life Framework)



Source: McGregor, et al. (2014) and Gupte, et al. (2015).

practitioners in international development or social policy. We can distil some order into the subsequent discussion by reminding ourselves of the main purpose of human well-being measurement – to understand what makes for a good society, in which human beings can survive and flourish. Good society is produced by a combination of efforts involving governments, citizens and all types of organisations in between (businesses, civil society, formal and informal). Considering these different levels and the ranges of activity that a well-being framework must speak to suggests that it is useful to distinguish between ‘dimensions’ and ‘domains’. The proposition here is that there is a standard and universal set of ‘dimensions’ (material conditions, quality of life, and relationality – the three basic dimensions of the OECD framework) and that each of these dimensions then contains a set of domains.

In order to reduce confusion in the following discussion of the three component steps to building well-being measures for development policy and practice, this paper will adopt this convention of the language of dimensions, domains and indicators. To recap: in the logic of this schema ‘dimensions’ are small in number and are fully universal; domains are universal but their number can be expanded or collapsed depending on context and the purpose for which the framework is to be used; and indicators can be universal but may also need to be context specific depending on the purpose for which the data is being generated.

5. Identifying what matters for well-being

The remainder of this paper uses this type of framework to structure a discussion of the range of methods that can be used to investigate a multidimensional assessment of human well-being. Using the three headline and universal dimensions of ‘Material Conditions’, ‘Quality of Life’ and ‘Relationality’ and drawing on the experiences of a number of ongoing measurement initiatives, it allows us to consider how to identify which domains are appropriate under each of these dimensions. Given that we have reviewed the provenance and content of a variety of the ‘top-down’ frameworks we now turn to consider what participatory methods have been used in other studies either to provide *sui generis* insights into what people need for their well-being in particular contexts or to provide some kind of ‘bottom-up’ justification for the proposed ‘top-down’ lists of domains and indicators.

Participatory approaches to the identification of well-being items and domains

There are an increasing number of studies worldwide that have used group-based methods to elicit relevant well-being domains (for example Narayan’s ‘Voices of the Poor’ and ‘Moving out of Poverty’ studies - see Nambiar (2012) for an overview). This can be done using individual interviews or by various forms of collective deliberations, such as community meetings or structured focus groups.

Focus group discussions are a relatively quick means of assessment, and can, if well conducted, produce an insightful debate as to what matters within a particular community, and why. If these are sufficiently open in their

approach they can also be a useful means of evidencing the need to broaden out ways of measuring well-being from methods that use a fixed and often limited number of criteria by enabling visualisation of missing domains. For example, recent focus group work in coastal Kenya which asked participants ‘what matters for “a good life”?’ revealed several domains that had been missing in previous more conventional assessments based on the livelihoods framework, such as the importance of being free to make decisions about life [to have a desired degree of autonomy] (Abunge et al., 2013). This kind of insight has helped to strengthen acceptance of a broader well-being perspective in poverty assessments.

The first phase of the quality of life research carried out within the Well-being in Developing Countries ESRC Research Group (WeD) which enabled the construction of the WeDQoL (described later) used focus groups and interviews to explore ideas about well-being in particular communities. The methodology attempted to avoid normative accounts by supplementing questions about what it means to live well with ones on personal sources of happiness and happy memories. The qualitative QoL data suggests that the most important areas of life for people in this sample are close relationships (family, ‘natal’ family, and partner), followed by material well-being (income, assets, satisfaction of basic needs, home and community environment, and access to local services), and relationships with the community and the wider world (for example, with relatives living in the city or overseas, or with the regional government). Religion (for example, believing in God or the human potential for enlightenment, conducting acts of worship, and living ethically) appeared to be very important in Bangladesh and Thailand and fairly important in Ethiopia. Similarly, education for themselves and their children was a priority for people in Bangladesh and Ethiopia, although apparently not in Thailand or rural and urban areas of Peru. This illustrates both the universality of dimensions such as material conditions or relationality and the specificity of the domains which populate them.

A final example is *Oxfam’s Humankind Index* whose consultation process asked people what aspects of life that affected ‘their ability to live well in their communities’ and engaged almost 3,000 people in Scotland using focus groups and an online survey (Walker et al., 2012). Oxfam Scotland’s Humankind Index is calculated from scores on a set of well-being ‘sub-domains’ (see Figure 3).

In validating the framework set out in the earlier part of the paper, the first stage will be identifying what matters most to people in particular settings.

Establishing what things contribute to well-being in particular contexts - focus groups

One commonly used approach to this is to invite purposively selected groups to describe a typical person in their community who ‘is doing well’ and another who ‘is not doing well’ to give a sense of the type of characteristics

Figure 3: Oxfam’s Humankind Index (2010/2011)

Sub-domain	Weight	Measure	Score	Change since 2009/10	
Housing	11	52.3	558	-20	↓
Health	11	94.0	1004	11	↑
Neighbourhood / Environment	9	60.0	524	9	↑
Work satisfaction	7	70.8	496	-	
Good relationships	7	13.3	91	1	↑
Safety	6	23.0	134	17	↑
Green spaces	6	44.0	256	3	↑
Secure/suitable work	6	91.1	531	-3	↓
Having enough money	6	47.0	274	-12	↓
Financial security	5	-4.3	-21	29	↑
Culture/hobbies	5	63.5	308	12	↑
Local facilities	4	44.0	171	-4	↓
Skills and education	4	27.0	105	4	↑
Community spirit	4	75.0	291	12	↑
Good transport	4	76.0	295	4	↑
Good services	3	66.0	192	3	↑
Tolerance	3	66.0	192	-	
Feeling good	2	80.4	156	-1	↓
Total			5558	65	

Source: Oxfam, 2013 (p5).

that are valued (and scorned/stigmatised) within that particular society. The group is then asked to discuss what they need to live well in this community and why. This process establishes a list of self-generated criteria against which it is possible to measure well-being in that community. The items that are identified may consist of things that people need to *have* to do well, need to be able to *do*, or need to be able to *be* to be well. This approach enables us to both identify locally relevant items and the terms that people use to describe these items. Critically it begins to afford us insights into what well-being means in particular contexts and it is to be expected that the detail will differ in different social and cultural contexts and be different for people in economies at different stages of development. Since focus groups are a form of deliberative process this also helps to identify issues around which there are differences of view and where there is consensus.

Whilst they take less time than a large number of individual surveys, the use of focus groups runs the risk of allowing either elite or ‘group-think’ views of the community to emerge. In this respect both the organisation and facilitation of focus group discussions of well-being require considerable prior experience of the community and careful forward planning to mitigate against these possibilities. There may also be significant community organisation and strategising in the prioritisation of well-being domains, to maximise the likelihood of achieving specific acquisitions from external (donor) actors (Coulthard, 2006).

A second possible stage in this process of exploring the meaning and dynamics of well-being in particular places involves the use of what can be regarded as a Foresight method. The groups can then be asked to reflect on their list of well-being items and to prioritise these using scenario-style pictures which depict the existing and future/planned activities of specific development projects. Each picture is used to explore the perception of the group as to who would be winners and who losers and in respect of which well-being needs. These discussions further illuminate areas of convergence and divergence between different social groups on the processes whereby well-being is achieved or not and in doing so relate these to the perceived (and anticipated) effects of project interventions. Specifically, they would:

- Discuss people’s actual and hypothetical engagement with projects
- Identify possibilities and opportunities to take their well-being domains into consideration in planned and existing project interventions
- Develop ‘well-being-led scenarios’ e.g. ask groups to discuss changes/amendments to a project: ‘what would this project look like if your (group’s) improved well-being [or locally relevant equivalent] was its main goal?’

This well-being scenarios approach was used effectively in a recent ESPA-funded project,¹¹ which facilitated group discussions around predicted scenarios of coastal development in Kenya. Groups met around a series of pictures (drawn by a local artist to depict a combination of possible events/impacts) and debated their accuracy, what was missing, who would be affected positively and negatively, and how people might react to the anticipated changes.

As described earlier, elite capture is a concern for group methods. Although this can be addressed to some extent through group composition and by building on knowledge of our study communities, the method is strengthened if we triangulate with results from individualised assessments of well-being items and domains.

Eliciting items using individual interviews, including the Global Person Generated Index

The same type of list of items that are important for well-being can be built through a series of one-to-one interviews. This reduces the risk of ‘group think’ or of people withholding information because of social status effects or embarrassment, but it also loses the deliberative dimension that is present in the focus group approach. These interviews can be structured to greater or lesser extent. A fairly loosely structured approach would simply begin by asking the person what is important for their well-being in their community and can then deal with responses in an open-ended way. But these approaches can be given considerable amount of structure as illustrated by the Global Person Generated Index (GPGI). The GPGI is an ‘individualised’ measure that uses a mix of open-ended questions and scoring to establish a person’s satisfaction with the areas of life that are most important to them.¹² The tool, which originated in the health care context, has been successfully validated in research across several countries (Ruta et al., 2004; Britton and Coulthard, 2013; Coulthard et al., 2014). It invites respondents to nominate up to seven areas that they consider important to their lives and explain their importance, and to then score each to indicate their level of satisfaction using a Likert scale. The advantage of applying an individualised method (to a sample of the population) alongside focus group data is that it adds methodological rigour by permitting triangulation to identify elite capture or other biases.

Organising into Domains

There are likely to be many context-specific items in the well-being lists compiled from focus groups and individual interviewing in each study site and the next challenge is organising these into domains. This clustering process can be done either through statistical clustering (e.g. factor analysis) and/or with reference to existing theory and empirical evidence on what have been identified as possible universal well-being domains. The reason for drawing on theory in the clustering is that relying on subjective ‘self-generated’ criteria alone can result in gaps and unintentional exclusion of important aspects of life; domains that are ‘taken for granted’ in particular contexts, that are ‘important but forgotten’ during interview, or are excluded through ‘ignorance’ or particular ‘strategising’ that communities might be engaged in.

For example, the importance of having clean drinking water may be recognised as an essential well-being item in parts of Africa, but be taken for granted and thus not be mentioned in the UK (see Scotland’s Humankind Index, Oxfam, 2013). The approach taken by Oxfam (2013), in their development of Scotland’s Humankind Index, was to carefully prompt communities to reflect and comment on possible ‘taken for granted’ criteria, recognising that people

¹¹ ESPA is the Ecosystem Services for Poverty Alleviation research programme, jointly funded by UK NERC, ESRC and DFID. <http://www.stockholmresilience.org/21/research/research-videos/1-10-2014-exploring-tradeoffs-in-well-being-in-coastal-systems-in-kenya.html>

¹² It is based on the ‘Patient Generated Index’ which is widely used in health care (Ruta et al., 2004).

might not mention those areas of life in which they are satisfied, or which are automatically provided. This might also involve asking people to reflect on what are *means* to a valued outcome, e.g. clean water, and what is the valued outcome (e.g. health, or being sufficiently healthy to earn a living). The provisional domains could then be taken back to the study sites in order to be verified through further discussion in focus groups.

This step of research provides a greater insight into what these domains might mean in local contexts and how they are interpreted by people (for example, in the development of the Oxfam Scotland Humankind Index, workshop participants were given well-being domains and asked to then describe what they meant to each individual (Oxfam, 2013). In order to move towards the development of a practical instrument for development policy and practice it might be necessary to reduce the number of domains to a manageable number, using the individual and group voting methods described in section 8.

6. Assessing how well people are doing: measurement and indicators of achievement in well-being domains

Once either a ‘top-down’ framework or ‘bottom-up’ process has been used to identify what people need for their well-being, and this has been organised into domains, we then have the challenge of assessing how well people are doing in respect of their achievements of the things that matter. There are two ways this can be done: the first is to identify existing data that might provide an indication of how well people are doing in respect of the domains and items, the second is to begin to generate data that is suited to the new framework.

Many of the new and emergent frameworks make use of existing data. The frameworks tend to be complex and require a mix of different kinds of information and as a result data is often drawn from a number of different sources. Thus when the Better Lives framework is applied to OECD countries it uses a mix of data from a range of national and international sources (OECD, 2011b).¹³ In the case of the OECD this includes ‘official’ data from national statistical systems, government departments and from international agencies, but it also acknowledges the need to use ‘unofficial’ statistics (for example, from the Gallup World Poll). While these are regarded as having potential quality problems, they are needed because ‘official statistics (i.e. statistics which are produced by National Statistical Systems) are either lacking or are not comparable across countries for some dimensions of well-being (e.g. subjective well-being, civic engagement and social connections)’ (OECD, 2011b: 8).

Because existing data regimes have not been designed with these frameworks in mind the existing data that are used are often either ‘proxy’ or ‘circumstantial’. Proxy data are used where there is no direct indicator of the actual item which is regarded as important for well-being but where there are credible data for something that is proposed as a proxy for it. For example, the quality of health is included in both the OECD Framework and the Human Development Index and in both cases it is indicated by the proxy of life expectancy.

Circumstantial data refers to that data which indicates the presence of circumstances in which it is believed well-being (in respect of a particular item) can be achieved, but does not tell us directly whether and how the indicator is contributing to well-being. For example, one of the indicators used for the ‘jobs and earnings’ domain in the application of the Better Lives framework in OECD countries, is ‘the share of the working age population who are currently employed in a paid job’. While this is an indicator that tells us about the circumstances in the labour market of the country in question it only tells us of the likelihood of a person having a job. On balance, having a paid job is probably better for well-being than not having one and no doubt having an income from such a job can contribute to well-being, but the indicator tells us little about who has paid jobs and who does not and what the differences are for their well-being.

As can be seen there are number of problematic issues that begin to emerge in respect of the use of existing data. First, it is usually not particularly focused on human well-being; second, even if a proxy is plausible it is usually broad-brush and/or an aggregated number which tells us little about the ‘quality’ of the outcome or about the differences between people in the population; third, given that official statistics take a considerable time to collect, collate and present, it is likely to be ‘old’ data that tells us little about the well-being of the population at the particular time at which the indicators are assembled and analysed.

In order to overcome some of these problems some of the more recent national well-being measurement initiatives are beginning to generate new data that is designed specifically for their adopted national well-being framework. Thus in Mexico the national statistical service (INEGI) has been finding ways to add more specific well-being components to existing official statistical collection instruments, while in the UK the Office of National Statistics has supplemented existing data sources with a newly designed ‘Personal Well-being Annual Population Survey.’¹⁴

The need to find or generate better data points to a further challenge for the broader global adoption of such frameworks, which is that richer countries are also relatively better off in terms of the quantity and quality

13 ‘The Compendium of OECD Well-being Indicators’ (OECD, 2011b) explains and discusses the range of data that has been used to operationalise the Better Lives Framework in particular country studies.

14 <http://www.ons.gov.uk/ons/rel/well-being/measuring-national-well-being/personal-well-being-in-the-uk--2012-13/sb--personal-well-being-in-the-uk--2012-13.html>

of official statistical data being generated. As such they have a good range of data (even if much of it is proxy or circumstantial) to use to give meaning to well-being measurement frameworks. In poorer, less developed countries the quantity and quality of existing statistical data is generally much lower. Moreover richer countries usually have more available capacity and funding with which to generate new bespoke data on a national scale, whereas the funding for and capacity in statistical offices are limited and overstretched in most developing countries.

In the context of the ‘bottom-up’ elicitation process described in the previous section there are two likely data gaps that might need to be addressed in generating original well-being specific data. As we have noted one of the distinctive characteristics of the new initiatives to measure well-being is the acceptance that this will require both objective and subjective data. While some data on people’s objective circumstances may already exist (depending on what level of detail is required) it may not exist for all of the items identified as being important for well-being. Thus any initiative to generate well-being specific data may necessarily involve the generation of objective data on new items that are reported as being important for well-being (for example, objective data on the extent of a person’s civic engagement). Some of these new items may be difficult to measure objectively.

The biggest data gap however is likely to be in respect of subjective well-being data. The recent OECD ‘Guidelines on Subjective Well-being’ (2013) provide a comprehensive review of the conceptual foundations and practical steps to be taken in subjective well-being measurement. Taking note of both the competing (or component) conceptions of subjective well-being (life evaluation, affect and eudaimonia) and also the exigencies of different purposes, scales and resource availability the Guidelines recommend a progressive set of modules. These propose that subjective well-being measurement begin with a core module that focuses on a general and global evaluation of subjective well-being, and then, depending on purpose and available resources move into ever deeper evaluations of different aspects of well-being through five further modules. The fifth of the modules (Module D) deals with ‘Domain Evaluation’ and aims to ‘collect people’s evaluative judgements on how well different aspects of their life are going’ (OECD, 2013b: 263).

The approach to domain satisfaction outlined in this module is still at a very broad-brush level (for example, How satisfied are you with your health?) and this reflects both the purpose of the OECD framework and the level of scale at which it operates (mainly at national scale). As we have discussed above, the purpose that this paper is dedicated to is at a lower level in policy processes and as such requires greater detail from domain level assessments of satisfaction.

The processes of ‘bottom-up’ item elicitation and domain organisation that are described in the preceding

section are intended to identify or validate a small and manageable number of domains, which are accepted as being of value across the different contexts for which the study is intended. These domains contain within them lists of items that people in these specific locations and contexts report as being important for their well-being. For example, under the ‘income and wealth’ domain there may be some specific assets that most people report as being necessary for their well-being in that community (e.g. in a coastal fishing community this may be a fishing boat, while in a farming community this may be a plough). In order to get to the detail of how well people are doing in respect of these items that matter for well-being in the communities under study it is necessary to select what would be an appropriate indicator for that item and domain. Again this can be made more or less context specific, depending on where the study is being conducted and its purpose. If one were studying only coastal fishing communities with a view to contributing to the formulation of new fisheries management rules, then the indicator may be ownership of a particular type of boat; whereas if one was studying both fishing and agricultural communities in a more general exploration of poverty policy options then the indicator might be made universal by asking about ownership of the key productive asset. While the domain and category of item can remain universal, the value gained by asking about perceived achievement or satisfaction in respect of a particular form of productive asset may be important in respect of understanding issues such as differences with a community or population and the dynamics of systems of production within that location.

In order to create a measure it is necessary to develop a scale.¹⁵ The most commonly used indicator for well-being outcomes is the level of ‘satisfaction’ a person reports as feeling in respect of the particular item or within that particular domain (see the measures in Ruut Veenhoven’s World Database of Happiness, which includes Likert and self-anchored scales such as Cantril’s ladder). This is a very specific and well researched methodological field and the detail and sensitivities of this approach to the subjective assessment of outcomes is carefully and extensively discussed in the OECD Guidelines on Measuring Subjective Well-being (2013). This approach has already been trialled in development contexts and Woodcock et al. (2009) report using individual level scores for satisfaction (using a Likert scale where 1 represents ‘the worst you can imagine’ and 5 represents ‘exactly as you would like it to be’) for each well-being domain generated through earlier qualitative research. It is possible to use different types and lengths of scale (e.g. binary – yes, no; Likert with adjectival descriptors that are three point – very satisfied, satisfied, dissatisfied – or even up to seven point, and continuous scales) and each of these has their particular merits and drawbacks.

One of the main challenges to satisfaction-based approaches is that satisfaction levels are influenced by

¹⁵ Scale = the way in which someone rates their own state of well-being on a particular domain.

adaptation (Nussbaum, 2001; Qizilbash, 2006; Sen, 1999: 62): where a person's goals and standards are adapted in relation to changing external circumstances (for example, adapted down to enable them to cope with an increasingly adverse set of conditions, or up in response to improvements). It is worth noting that the process of development is intended to produce improving circumstances and as such upward adaptation is something that should be both expected and confronted in development practice.

This is a point in the process at which the relationship between subjective and objective data becomes critical. One way that the issue of adaptation¹⁶ can be dealt with is to triangulate subjective reports with objective levels of welfare. Having both objective and subjective reports on the same items can yield very valuable information and provide important insight into the dynamics of adaptive preferences. For example, there are numerous examples from countries all round the world where when their situations in particular domains are viewed objectively people are doing very badly but where they nevertheless express satisfaction in respect of their performance in that domain.¹⁷

Crucially, however, this is an area in this approach where qualitative work becomes particularly important and insightful. Statistical techniques can be used to assess the validity of subjective well-being scores but it is important also to use qualitative strategies to explore validity. This highlights the potential of an important feedback loop in such processes of indicator generation and measurement since it is valuable to know whether what is purported as being captured by the measure is actually understood as providing useful information by the people whose well-being is being measured. 'Face validity' can be tested using 'member checking' with interviews and focus groups and cognitive debriefing during the administration of the measure. Construct validity can also be tested by seeing whether the scores reflect existing differences in well-being within the community that are established by other approaches to generating evidence. For a further feedback loop, and to test what might be called 'convergent validity', it is also possible to conduct interviews or focus groups to discuss findings amongst policy-makers and practitioners. These can be used to test the earlier proposition about the extent to which there are any real disconnects between the view of the situation held by policy-makers and practitioners and the scores generated by people for whom they are implementing policy or projects.

Having developed a robust instrument for measuring well-being, analysis it is then a matter of exploring and identifying key commonalities and differences between how well-being is being scored for different items and by

different groups. Results can be analysed by stratifying samples along different lines, for example by age, gender, wealth, ethnicity, or social status/marginalisation.¹⁸

7. Priorities amongst things that matter: trade-offs, weighting and resource allocation

The issue of understanding how the different things that matter relate to each other is important for a range of reasons. In terms of better understanding how different people construct their well-being in different contexts and at different times it is important to understand what people prioritise and what trade-offs they are prepared to make between different aspects of well-being. Both of these sets of insights are important from the policy perspective and relate to the challenge of establishing weightings in respect to the different things that matter.

The issue of how weights get assigned to the different things that people regard as important for their well-being is particularly important when considering how decisions are made about public resource allocation. This is highlighted in Ravallion's (2010) critique of the Alkire-Foster Multi-dimensional Poverty Index (MPI): 'listing of dimensions of poverty is one thing, but assigning weights to each different aspect of poverty, so policy-makers have some guidance on where to put their (limited) resources, is quite another (Ravallion, 2010)' (cited in Melamed, 2011). While the issue of weighting often seems to be dealt with in a detached way as a technical matter, it does in fact raise a profound political issue: it addresses the question 'whose priorities count?'

A common starting point in discussions of human priorities is what is often referred to as Maslow's Hierarchy of Needs. Although it is a misinterpretation of Maslow's 'Theory of Human Motivation', it is often used to assert an implicit weighting schema for poor people and suggests that people first are motivated by physiological needs and only then move up a pyramid to fulfil other more complex higher order needs that include safety, belongingness and love, esteem, and self-actualisation. In his later writing Maslow dissociated himself from the imposition of a simplistic hierarchy and empirical evidence indicates that higher level needs also have a vital role to play in human physiological survival. While it may often be the case that people in impoverished circumstances give high priority to meeting basic needs for food, water and shelter, there are also numerous examples, from countries at all levels of development, where people's choices disprove the notion of a fixed and simple hierarchy. People make trade-offs between meeting different kinds of needs. These decisions cannot be blithely dismissed as irrationality, but rather they involve people reflecting their own needs priorities

¹⁶ Including maladaptation – which is where people just get their subjective assessments wrong.

¹⁷ Sanitation facilities can often provide a good illustration of this – see <http://www.communityledtotalsanitation.org/>

¹⁸ The focus on marginalised groups is inspired by Oxfam's Humankind approach and rationale: 'If we accept that dominant institutions and forms of representation – including, potentially, mainstream survey and census instruments – tend to favour and reflect the interests of dominant groups and individuals, then there is a need to ensure that the research design for a project of this type prioritises seldom heard voices' (Oxfam 2012: 8).

and choices in particular ‘real-world’ social, cultural and economic circumstances. To presume or impose such a hierarchy when formulating and implementing public policy is patronising and morally problematic. As with the identification of what it is that is important for well-being the simplistic imposition of ‘top-down’ assumptions, rather than the ‘bottom-up’ investigation of what peoples’ priorities are runs the risk of misdirecting public policy decisions.

As Rodriguez-Takeuchi (2013) suggests, knowing how people prioritise different aspects of their lives is useful in that an increase in a more important factor will increase overall well-being more than a similar change in a less important factor. The problem is that how people assign weightings to different aspects of their lives varies substantially between individuals, is heavily influenced by specific contexts, and also changes over time according to different circumstances. This may be obscured in well-being research due to the methods of analysis and emphasis on what cuts across groups. In the midst of complexity and uncertainty over what is a priority and why, ‘development practitioners implicitly assign weights based partly on good evidence and analysis, but partly on development fads and political imperatives’ (Melamed, 2011:2). Melamed, et al. (2012) point out that such a lack of transparency over resource allocation decisions means that development processes are not scrutinised by political feedback mechanisms, creating an imbalance between democratic forces in donor countries, where citizens demand accountability for how their taxes are spent, and democracy in recipient countries, where people often lack sufficient voice to determine what resources they receive and whether these deliver the outcomes they prioritise. Concerns have been raised that this lack of transparency will result in a ‘tilting’ of the balance towards the views and priorities of taxpayers in donor countries, and away from views of recipients of development aid as to what is needed (ibid).

A more transparent approach to prioritisation would attach weights to different outcomes that reflect peoples’ priorities in a way that can guide decision-making and evaluate outcomes (Melamed, 2011). This would enable us to combine a multi-dimensional generic outcomes measure, to facilitate comparisons of effectiveness and cost effectiveness across different sorts of interventions, with a weighting method that reflects the priorities of different groups affected by these interventions. As an illustration, Rodriguez-Takeuchi (2013) uses the example of ‘good health’ and ‘having voice in community decisions’. She posits that many people would assume that having good health matters more than having a voice in community decision-making, and that these different weightings lead to a greater contribution of health to overall well-being, relative to voice. Such weightings of importance could direct development strategies to allocate greater attention (and budget) to support the most important priorities of a majority of people, leading to a greater increase in society’s well-being (compared with a similar change in a

less important factor) (see also Russell and Hubley, 2005). This is true, but also suggests an almost utilitarian focus on the greatest happiness for the greatest number, although there is no reason why weights couldn’t be used to make a small but highly deprived group considerably better off. Similarly problematic is the assumption behind how people weight things. In the example given here, ‘health’ and ‘voice in community decisions’ are different things – the first being more of an end and the second more of a means. A person may well prioritise voice over health if they thought that having a voice would lead to better decisions and be a means to a better outcome such as improved health.

In health studies, the aim seems to be to arrive at a common value across a population by aggregating individuals’ values or by deriving them directly at the aggregate level. However, it is unlikely that there will be unanimity on weights, as is evidenced by a wide range of research. For example, drawing on WeD survey data from Thailand which asked people to assign a ‘necessity’ rating to 51 well-being domains, Woodcock et al. (2009) illustrate the lack of consensus regarding the perceived necessity of different domains (the only item that all 369 respondents rated as being very necessary was water) (Woodcock et al., 2009 – see Box 1 for more detail). As Sen (1999) and Alkire (2012) argue, it is because of this inherent and unavoidable variety in how people prioritise different well-being domains that any multiple dimensional measure should be robust to a *range of plausible weights*. The method illustrated in Box 1 (Woodcock et al., 2009) used self-generated well-being domains to develop a ‘Thai Individualized Goal Attainment’ (TIGA) score, which combined individual perspectives on importance (necessity) with a person’s reported level of satisfaction with each domain. As the authors’ argue:

‘as universal items are not equally valued, respondents should nominate the areas they consider important and assess their performance against their own standards.’
(ibid:168)

The information on importance, including how this varied within a given location or between different groups, illuminated how people trade off different important areas of life in a given scenario, e.g. a new economic opportunity that requires migration. Asking people to weight things using a constrained scoring method should elicit both types of information at the same time so developing a robust methodology for this would be an important and distinctive contribution, building on rather than repeating large-scale consultations with people experiencing poverty worldwide. A key question for us will be how much can we reasonably aggregate, even if in doing so we lose some nuance about differences between people’s views, given the wider objective of getting policy-makers to use this information more systematically.

Box 1: Weighting in the WeD Quality of Life measure (methodology summarised in Woodcock et al., 2009)

The WeD Quality of Life measure (WeDQoL Weighted Goal Attainment Scale or WeDQoL-goals) has two components that measure i) how necessary an item is to a respondent's well-being and ii) how satisfied they are with this item. The first component is used to weight the second so the final score represents satisfaction with items that are valued by the respondent. In Thailand (Woodcock et al., 2009) the WeDQoL-goals consisted of 51 items and were interview-administered to 369 people, aged between 15 to 89 (mean age 45.7, sd 18.0). Respondents rated the perceived necessity for well-being of the 51 goals using a three-point scale (0 to 2, where 0 represents 'unnecessary', and 2 'very necessary'). They then rated their satisfaction in achievement of the same goals. The scores for necessity were used to weight the scores for satisfaction so goals that were 'not necessary' were excluded when calculating goal satisfaction, while those that were 'very necessary' were weighted more highly. The factor analysis of WeDQoL-goals identified three factors which have been labelled by the researchers as 'community/social/health' (twenty-three item, alpha 0.90), 'nuclear family' (six item, alpha 0.82), and 'house and home' (fifteen item, alpha 0.80). The 'house and home' factor can be sub-divided into two sub-scales for a more fine-grained analysis: 'basic house and home' (eleven item, alpha 0.80) and 'luxury' (four item, alpha 0.61) (see table 2). The factor scores enable a more reliable representation of the importance a respondent accords to a domain of life (for example, their family) than would be provided by a single item. The WeDQoL-goals were applied in all four countries (see Yamamoto et al., 2008 for Peru; other reports available on request) using a common format with additional country-specific items.

Approaches for weighting

In eliciting priorities there are four main types of approach, many of which have been used in developing countries:

- Weighting items according to importance (Woodcock et al., 2009, or part of an individualised measure of (health-related) quality of life, Ruta et al., 2004)
- 'Participatory numbers' (Holland, 2013), which provide accessible and transparent methods for eliciting and debating people's priorities (this might include matrix scoring, ranking, or pairwise ranking)
- Experimental methods used by behavioural economists in developing countries (Cardenas and Carpenter, 2008; van Rijisbergen and d'Exelle, 2011)
- Values elicitation within health care (Mullen, 1999) through deliberative and experimental methods

In assessing these, Green et al. (2000) propose the following criteria: practicality (i.e. acceptability to respondents), reliability, empirical validity (i.e. whether it can predict actual choices) and theoretical validity. We will keep these in mind as we outline the approaches listed below, alongside a further criterion of whether what they produce is simple and useful, or can be made so.

Weighting

As discussed above, weighting items using people's preferences increases face validity and the acceptability of the measure to different publics. However, opinion is divided as to whether it increases the accuracy of measurement because satisfaction and importance are highly correlated and knowing how different domains are valued does not increase understanding of judgements of satisfaction with life as a whole (reviewed in Guardiola et al., 2013). Individualised measures such as the Patient Generated Index (discussed earlier) and SEIQOL were applied in different forms and with different groups in the 1990s and 2000s, however neither seems to have influenced

mainstream Patient Reported Outcome Measures (PROMS), possibly due to concerns about whether they were sufficiently robust measures (e.g. Camfield and Ruta, 2007) and the difficulties of interpersonal comparison.

Returning to the example of the Humankind Index, once the sub-domains listed earlier had been agreed, people were asked to score them according to importance using a voting system. The voting was conducted in street stalls and workshops, and formed part of a survey. Weightings were produced by analysing the votes from the various events. The outcome was a set of weights for the well-being sub-domains, which the Scottish public believe are important for living well in their communities. Interestingly, some of the well-documented conundrums of self-reported satisfaction were evident here (as is acknowledged by the authors), particularly in the sub-domain of health. In a country which has some areas that have notable lifestyle and health problems, 88% of respondents self-assessed their state of health as being either good or very good. Whilst other data could have been drawn upon (such as life expectancy), in this case it would have reduced the comparability of data across years. Each measure (a single statistic) for each sub-domain was multiplied by the given weight to arrive at a final 'score' per domain. The sum of these scores creates the overall Oxfam Humankind Index Score, which can be used to analyse changes over time as a signifier of progress in Scotland's prosperity (Walker et al., 2012).

Participatory numbers

Participatory numbers offers a new model for including people in generating and using the numbers that shape their experience of development (Holland, 2013). It draws on tools used within participatory needs assessment and priority setting (Chambers, 2007) such as matrix scoring or ranking. These can be adapted to constrain people's choices in a way that more closely resembles techniques used in health economics. For example, people are given a

limited number of units to use to indicate priority, requiring them to trade off one dimension against another, or an overall ranking is generated by asking people to compare one attribute with another and say which they value most (pair-wise ranking). While rankings are ordinal data so cannot be analysed in the same way as scores, which are conventionally treated as interval data, they can be made equivalent. Barahona et al. (2003) discuss this and other methodological issues in outputs from the Reading-based 'Methodological Framework Integrating Qualitative and Quantitative Approaches for Socio-Economic Survey Work' programme which forms the methodological foundation for much of the participatory numbers work.

Experimental methods

Much of the experimental work of behavioural economics has focused on topics that potentially influence, but are distinct from well-being such as cooperation, trust, risk, patience, and fairness. However, the methodological lessons learned from their application in developing countries have implications for our project. For example, Cardenas and Carpenter (2008:329-331) note – drawing on examples from the Andes, Andhra Pradesh and the USA – that behaviour in experiments correlates with numeracy, is influenced by the normative context (e.g. the place of supernatural forces), is affected by the perceived credibility of the experimenter and the salience of what is at stake, and is particularly vulnerable to recruitment biases, including peer effects.

Van Rijsbergen and D'Exelle (2011) offer one example of the potential of experimental methods in their use of discrete choice experiments, which combines elements of pairwise ranking and vignettes, to understand use of delivery care in Tanzania by different population groups. They compare results from observational studies with stated preferences elicited through choice experiments where rural women were asked directly about their preferences for different services.¹⁹ Women were asked to participate in a ranking exercise of hypothetical obstetric care providers using attributes such as health care worker attitude which women might take into consideration in deciding where to have their next child. By combining the different attributes they obtained 64 possible scenarios which they reduced to eight, recognising that this was too large for an individual to rank. Participants then compared, e.g. a doctor who smiled and listened carefully, but was remote, expensive and unreliable with, e.g. a nurse who did not smile and listen carefully, was also remote and unreliable, but was free, and ranked the different scenarios in order of preference. They were asked to do this in relation to both a normal delivery and an emergency one (scenarios were on individual cards

which also had pictures) and the results were analysed using conjoint analysis. Conjoint analysis enables the value attached to an attribute to be inferred from people's choices between options containing more or less of the attribute in question, though this is not a transparent process. The experiment provided information on women's preferences and showed how they were shaped by factors such as age, socio-economic status and previous experience. Methodological lessons relate to the value of pictures and 'anchoring vignettes', which are used to establish intra-personal comparability in survey responses (King et al., 2004), and the need to limit the number of scenarios assessed (i.e. the complexity of the task).

Vignettes have also been used by Kapteyn et al. (2007) to compare life satisfaction between respondents in the US and Netherlands and by Beegle et al. (2012) to look at subjective welfare in nearly 5,000 households in Tajikistan. Respondents were asked to place themselves and four vignette households on a subjective welfare ladder with six rungs; they placed themselves first, then the households, and then themselves again, at which point a quarter of the sample changed the ranking of their households, indicating the importance of information and deliberation in making a judgement. Beegle et al. (2012) illustrate the potential of including simple valuation methods within surveys to access the views of a larger and more representative sample (a historical problem with experimental methods). Vignettes could also be used as 'marker states' that characterise a typical good, medium or poor outcome for a given domain (Hopkins and King, 2010) and could be developed in particular settings using methods such as Krishna's (2009) 'stages of progress'.

Values elicitation within health care

As van Rijsbergen and D'Exelle (2011) observe, most of the work on preference elicitation has been done within healthcare to engage different publics in priority setting (this might relate to decisions as fundamental as the values of the health service or as specific as the location of a particular service). Mullen's (1999) review lists the different approaches in order of intensity of engagement, for example, events/ phone-ins, questionnaires, representative focus groups, key informant interviews and citizens' juries. Within these approaches and particularly within more structured approaches she identifies simple unconstrained choices to assess intensity of preferences (e.g. voting, for example: selecting the four most importance service developments from a list of nine) and simple constrained choices to compare different options (e.g. health economic techniques such as time or person trade-off, standard gamble and willingness to pay)²⁰ (ibid, table 2). She notes

¹⁹ This approach has also been used in sectors outside health care, for example to evaluate water policies (Hope, 2006) or preferences for firewood in rural Guatemala (van Kempen et al., 2009).

²⁰ These involve inferring preferences from willingness to pay for an outcome, willingness to trade off time or individuals, and willingness to gamble. A difficulty with these stated preference methods is the existence of framing effects, and confounding effects arising from preferences over the thing being traded (time preference in the case of time trade off, and attitudes to risk in the case of standard gamble). Discrete Choice Experiments (a form of pairwise ranking) and Visual Analogues Scales (VAS) where a person rates a hypothetical health state on a scale of 0 to 100 are now more commonly used to elicit weights.

that the most effective evaluation exercises are usually carried out in two stages where the attributes/criteria are valued first, and then the options are valued against these criteria, although in practice this is rarely done and may be a more difficult task given that it is hard to evaluate the importance of, for example, responsiveness, outside a specific context.

The methods listed above face different challenges relating to their specific attributes; for example, constrained choices using money are inevitably influenced by respondents' own purchasing power, although this is not the case with points or tokens. Common problems relate to inter-personal comparisons of utility which require a common frame of reference (see Arrow's theorem) and aggregation of individual preferences in a way that preserves their transitivity and ensures that preference orderings are not affected by the introduction of irrelevant alternatives. For this reason Mullen (1999) proposes presenting the full distribution of results using graphical methods rather than attempting to construct a single aggregated measure whose results would be affected by the method of aggregation. We would concur with the importance of presenting a range of possible weights derived from different methods/informants and discussing the differences between them, as is also proposed in Melamed et al. (2012:23) in relation to 'decision conferencing'. Mullen (1999) warns that 'dangers can arise from the use of computer software, which permits researchers to employ methodologies which they do not fully understand...the implications of' (p.231). (Also see Decanq et al. (2011) and Santos and Ura's (2008) demonstrations of the sensitivity of poverty classifications to the type of weight used.)

Mullen's review concludes with a set of questions to ask of any technique (ibid, table 3) which are equally applicable to our proposed methodology, for example whether respondents are permitted to answer in a way that is both meaningful to them and to the problem. Other examples of techniques for eliciting weights are reviewed in Rodriguez-Takeuchi (2013) who adds to the health economic techniques mentioned earlier, Discrete Choice Experiments/Pairwise ranking and Rating Scales. Swing weights, which are used in Multi Criteria Decision Analysis (Hammond et al., 1999; Dodgson, 2009), allow choices to be made in different scenarios, as with the earlier maternal health example, which enables exploration of what people choose when their preferred option cannot be realised.

Disability Adjusted Life Years (DALYs)

One way in which a standard metric is created within global health is through the 'Disability Adjusted Life Year', which was developed by the Global Burden of Disease study (GBD, Murray and Lopez, 1996, first edition). DALYs combine epidemiological data with a complex series of weights and discounts, including weights for age (life years are assigned different value at different ages), disability (life years are assigned different value according to the health state, on a scale from 1, death to 0, perfect

health), time discounts, and a hypothetical life expectancy of 80 years for men and 82.5 years for women. The process and the outcomes have been widely critiqued on a number of grounds, as we discuss below.

DALYs are typically calculated using the universal weights of the GBD which were elicited from a panel of health experts in Geneva using a rating scale. Twenty-two of the health states, the 'indicator conditions', were assessed with a 'Person Trade-off technique' where the experts were asked how many lives saved of people in a particular health state (e.g. quadriplegia) they would consider equivalent to saving the lives of 100 healthy people. While from a development perspective the decision to use universal weights for health conditions seems to contain a number of assumptions (e.g. relating to the importance of context), the developer of DALYs argued that it should not make a difference to the burdensomeness of the health condition, whether it exists in a person in New Jersey or New Delhi. Even though their lived experiences and opportunities for participation may differ, their health state is the same and should have the same value. Reidpath et al. (in Bickenbach, 2008:9) argue conversely that the average disability weight for paraplegia of 0.671,

'misrepresents the reality of the life of a rural dwelling Cameroonian with paraplegia compared with an urban dwelling Australian with the same condition. People with paraplegia in Cameroon lead a highly stigmatized life of almost complete social disengagement; there is no social and environmental infrastructure to support their disability and moving around their environment unassisted is impossible. In contrast, people with paraplegia in Australia experience a much easier life; the support infrastructure is considerably better, and significantly, the provision of infrastructure is mandated and enforced through social systems such as building codes'.

In response to this critique, there have been some attempts to set local weights (e.g. in Zimbabwe, Chapman et al., 2006), however, these have been less common than might be expected outside the Global North (e.g. Australia, Canada, the Netherlands). Perhaps as result, diseases typical of poor populations are often assigned by GBD a relatively small burden because the values of 'wealthy experts' in privileged countries with responsive health care systems have determined the disability weight of those conditions. We could find little work on the acceptability of different evaluation methods in developing countries; however, Salomon and Murray (2004) suggest that among public health professionals internationally there are no differences in weightings when PTO, TTO, SG or VAS methods are used.

Some of the problems with the DALY approach outlined above were addressed in the 2008 modification of the method where evaluators were asked a series of discrete choice questions to indicate which of a pair of health states

they would regard as worse. The team also included the views of community members and health workers from the Philippines, Tanzania and India in the evaluation (although not necessarily people with disabilities) rather than relying exclusively on experts.

8. Conclusion

In this paper we have reviewed the growing field of new initiatives for measuring progress. The limitations of income-focused measures of development and progress are now well understood. The paper offers a contribution to the challenge set by the Stiglitz-Sen-Fitoussie Commission Final Report to shift from measuring progress in terms of production to measuring progress in terms of improvements in human well-being. While much of the current effort to develop new metrics are focused at the level of the nation state and are for global comparison, this paper focuses particularly on how these new measurements might be applied to make development policy and practice more effective at the grassroots level of policy and practice. The paper explores a number of elements of the emerging consensus in how and what needs to be measured. There are two related major points of consensus: the first is that human well-being must be understood as multidimensional and the second is that its assessment or measurement must involve consideration of both objective well-being and subjective well-being.

The paper notes that the well-being frameworks that guide the various new measurement initiatives can be either ‘top-down’ or ‘bottom-up’ in character, but most of the current efforts are ‘top-down’, arising from one or other conceptual, philosophical or ideological position. Although the idea of participatory development has been around for some time now there have been few systematic attempts to build a framework for understanding well-being that is founded in ‘bottom-up’ consultations about what matters for the well-being of the people to whom these measures are supposed to refer. That said, it would be wrong to over-emphasise the gap between ‘top-down’ and ‘bottom-up’ frameworks and most frameworks of either kind come out with broadly similar general framings. A brief review of a number of major frameworks that are currently informing well-being measurement initiatives suggests that, although

there is a great confusion of language and terminology, they usually identify three general and universal dimensions of well-being: material well-being – dealing with the material aspects of life; quality of life – dealing with a broader set of considerations in life that affect our more general quality of life; and relational well-being – the patterns of relationships that either produce improvements in future well-being or detract from it.

The paper briefly addresses some of the potential confusions that arise out of the interchangeable use of the terms ‘dimension’ and ‘domain’, but based on an adaptation of the OECD Better Lives Framework recommends the adoption of that for universal dimensions and then breaking each of these down into a set of domains. In order for a new measurement regime to work effectively there needs to be conceptual consistency between the applications of the framework at different levels of the human system and also when it is being used for different and diverse purposes. The set of three top-line dimensions must be considered universal and there can be agreement on a set of universal domains – although the number of these has yet to be definitively agreed and not all domains may be necessary for consideration for the purpose of different measurement exercises. Within domains there may be some items that are universal and some that are context specific. Similarly, but depending on purpose, policy need and the scale of the measurement task, indicators may either be universal or more context-specific.

Given that there is a widely held view that there are often disconnects between what development policy-makers and practitioners think is important for the well-being of poor people and what poor people themselves think is important, the potential for the ongoing misdirection of development effort is considerable. The remainder of the paper explored the different methodologies that might be available for: the systematic identification of what is important to people for them to live their lives well; assessing how well people are doing in their achievements in respect of the things that they regard as important for them to live well; and establishing understandings of how the different things that are important for well-being relate to each other and how people prioritise and make trade-offs between them.

References

- Abunge C, Coulthard S, Daw T.M. (2013) 'Connecting Marine Ecosystem Services to Human Well-being: Insights from Participatory Well-being Assessment in Kenya', *Ambio* 42(8): 1010-21.
- Alexandrova, A. (2005) 'Subjective Well-Being and Kahneman's "Objective Happiness"' *Journal of Happiness Studies* 6: 301-324.
- Alkire, S (2012) 'Value judgements in multidimensional poverty measurement design: Summary of the workshop and closing remarks'. *Maitreyee* 21. Southborough, MA: Human Development and Capability Association.
- Alkire, S. and Foster, J.E. (2011) 'Counting and Multidimensional Poverty Measurement' *Journal of Public Economics* 95(7-8): 476-487.
- Anand, P. Hunter, G. Carter, I. Dowding, K. Guala, F. and van Hees, M. (2009) 'The Development of Capability Indicators', *Journal of Human Development and Capabilities* 10(1): 125-152.
- Axford, N., Jodrell, D., and Hobbs, T. (2014) 'Objective or Subjective Well-Being?', in Ben-Arieh et al. (eds.), *Handbook of Child Well-Being*. Netherlands: Springer.
- Barahona, C., & Levy, S. (2003) 'How to Generate Statistics and Influence Policy using Participatory Methods in Research: Reflections on Work in Malawi 1999-2002'. *Working Paper* 212. Brighton: Institute of Development Studies.
- Beegle, K., Himelein, K., Ravallion, M. (2012) 'Frame of reference bias in subjective welfare', *Journal of Economic Behavior & Organization* 81(2): 556-70.
- Beetham, D. (1991) *The Legitimation of Power*. Palgrave Macmillan: London.
- Bickenbach, J. (2008) 'Measuring Health: the disability critique revisited'. Unpublished (downloaded 11/9/14 http://peh.harvard.edu/events/2008/global_burden_disease/day_1/Bickenbach_paper.pdf).
- Booth, D. (2005) 'Missing Links in the Politics of Development: Learning from the PRSP Experiment'. *Working Paper* 256, ODI: London.
- Britton, E., and Coulthard, S. (2013) 'Assessing the Social Well-being of Northern Ireland's Fishing Society using a three-dimensional approach'. *Marine Policy* 37: 28-36. doi: 10.1016/j.marpol.2012.04.011.
- Camfield, L. and Ruta, D. (2007) 'Translation is not enough: using the Global Person Generated Index (GPGI) to assess individual quality of life in Bangladesh, Thailand, and Ethiopia'. *Quality of Life Research* 16 (6): 1039-1051.
- Camfield, L. and McGregor, J.A. (2009) 'Editor's Introduction: Quality of Life and International Development Policy and Practice.' *Applied Research in Quality of Life* 4(2): 128-134.
- Camilo Cardenas, J. and Carpenter, J. (2008) 'Behavioural Development Economics: Lessons from Field Labs in the Developing World', *The Journal of Development Studies* 44(3): 311-338.
- Catley, A., Burns, J., Abebe, D. and O. Suji (2008) *Participatory Impact Assessment. A Guide for Practitioners*. Feinstein International Centre, Boston.
- Chambers, R. (1997) *Whose Reality Counts? Putting the Last First*. London: Intermediate Technology Publications.
- Chambers, R. and Holland J. (eds.) (2007) 'Who Counts? The Power of Participatory Statistics'. London: Publisher Practical Action Publishing.
- Chambers, R. (2005) 'Ideas for Development Earthscan'. London.
- Chambers, R. (1997) 'Whose Reality Counts? Putting the Last First'. London: Intermediate Technology Publications.
- Chapman, G., Hansen, K.S., Jelsma, J., Ndhlovu, C., Piotti, B., Byskov, J. and Vos, T. (2006) 'The burden of disease in Zimbabwe in 1997 as measured by disability-adjusted life years lost', *Tropical medicine & international health* 11(5): 660-71.
- Coulthard, S. (2006) 'Developing a people centred approach to the coastal management of Pulicat Lake, a threatened Indian lagoon'. PhD thesis, University of Bath.
- Coulthard, S., Sandaruwan, K.L., Paranamana, N. and D. Koralgama (2014) 'Taking a well-being approach to fisheries research: insights from a Sri Lankan fishing village and relevance for sustainable fisheries' in Camfield, L., *Methodological challenges and new approaches to research in international development*. London: Palgrave Macmillan.
- Cummins, R. (1998) 'The Second Approximation to an International Standard for Life Satisfaction', *Social Indicators Research* 43: 307-334. Netherlands: Kluwer Academic Publishers.
- Decancq, K., Ootegem, L.V., Verhofstadt, E. (2011) 'What if we voted on the weights of a multidimensional well-being index? An illustration with Flemish data'. *CSB working paper* 11/10.
- Diener, E., Emmons, R. A., Larsen, R. J. and Griffin, S. (1985) 'The Satisfaction with Life Scale', *Journal of Personality Assessment* 49: 71-75.
- Diener, E., Oishi, S. and Lucas, R. (2009) 'Subjective well-being: The science of happiness and life satisfaction' in Lopez, S.J., and C.R. Snyder (eds.) *The Oxford Handbook of Positive Psychology (2nd Ed)*. Oxford: Oxford University Press.
- Dodgson, J. (2009) 'Multi-criteria analysis: a manual'. London: Department for Communities and Local Government.
- Doyal, L. and Gough, I. (1991) 'A theory of human need'. Basingstoke: Macmillan.
- Durand, M. (2013) 'Plenary Address, 4th OECD Global Forum, New Delhi, 2012. In Proceedings of the 4th World Forum on Statistics, Knowledge and Policy'. Paris: OECD.
- Eyben, R. (2013) 'Uncovering the politics of "evidence" and "results". A framing paper for development practitioners.' (<http://bigpushforward.net/wp-content/uploads/2011/01/Uncovering-the-Politics-of-Evidence-and-Results-by-Rosalind-Eyben.pdf>).
- Gough, I. and McGregor, J.A. (2004) 'Guest editor's Introduction: Human Well-being: Communicating Between the Universal and the Local', *Global Social Policy* 4(3), 275-276.
- Gough, I. and McGregor, J.A. (2007) (eds.) *Well-being in Developing Countries: From Theory to Research*. Cambridge: Cambridge University Press.
- Green, C., Brazier, J. and Deverill, M. (2000) 'Valuing Health-Related Quality of Life A Review of Health State Valuation Techniques', *Pharmacoeconomics* 17 (2): 151-165.
- Guardiola, J., Andre's J. P-T. (2013) 'Building Weighted-Domain Composite Indices of Life Satisfaction with Data Envelopment Analysis', *Social Indicators Research* (in press) DOI 10.1007/s11205-013-0346-3.

- Gupte, J. and te Lintelo, D. (2015) 'Informal Work and Wellbeing in Urban South Asia'. End of Project Report to DFID South Asia Research Hub. Brighton: Institute of Development Studies.
- Hall, J. and Rickard, L. (2013) 'People, Progress and Participation: How Initiatives Measuring Social Progress Yield Benefits Beyond Better Metrics', *Global Choices* 1. Gütersloh, Germany: Bertelsmann Stiftung.
- Hammond, J. S., Keeney, R. L. and Raiffa, H. (1999) *Smart Choices: A Practical Guide to Making Better Decisions*. Boston, MA: Harvard University Press.
- Helliwell, J., Layard, R., and Sachs, J. (2013) *The World Happiness Report* (2nd). New York: UNSDNS.
- Hill, M. (2005) *The Public Policy Process*. London: Pearson.
- Holland, J. (2013) *Who Counts?: The Power of Participatory Statistics*. Rugby, UK: Practical Action Publishing.
- Hope, R.A. (2006) 'Evaluating water policy scenarios against the priorities of the rural poor', *World Development* 34(1): 167-179.
- Hopkins, Daniel, and Gary King (2010) 'Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability'. *Public Opinion Quarterly* 1-22.
- Kapteyn, A., Smith, J.P. and van Soest, A. (2007) 'Vignettes and self-reports of work disability in the U.S. and the Netherlands', *American Economic Review* 97(1): 461-473.
- King, G., Murray, C.J.L., Salomon, J.A. and Tandon, A. (2004) 'Enhancing the validity and cross-cultural comparability of measurement in survey research', *American Political Science Review* 98 (February): 191-205 (<http://gking.harvard.edu/files/abs/vign-abs.shtml>).
- Krishna, A. (2009) 'Subjective assessments, participatory methods and poverty dynamics: The stages of progress method' in Addison, T., Hulme, D. and Kanbur, R. (eds.) *Poverty dynamics - interdisciplinary perspectives*. New York: Oxford University Press.
- McGregor, J.A., Boarini, R., and Kolev, A. (2014) 'Measuring well-being and progress in countries at different stages of development: Towards a more universal conceptual framework'. *OECD Working Paper* 325. Paris: OECD.
- McGregor, J.A., Camfield, L and A. Woodcock (2009) 'Needs, Wants and Goals: well-being, Quality of Life and Public Policy'. *Applied Research Quality Life* 4: 135-154.
- McGregor, J.A. (2004) 'Researching Well-being: Communicating between the Needs of Policy Makers and the Needs of People' *Global Social Policy* 4(3): 337-358.
- Melamed, C. (2011) 'Does development give poor people what they want?' *ODI Background note* (<http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6606.pdf>). London: Overseas Development Institute.
- Melamed, C., Devlin, N and J. Appleby (2012) 'Valuing development: could approaches to measuring outcomes in health help make development more accountable'. *ODI discussion paper*. London: Overseas Development Institute.
- Mullen, P. (1999) 'Public involvement in health care priority setting: an overview of methods for eliciting values', *Health Expectations* 2: 222-234.
- Murray C.J.L., Lopez, A.D. (eds.) (1996a) 'The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020'. *Global Burden of disease and Injury Series* 1. Cambridge, MA: Harvard University Press.
- Nambiar, D. (2012) 'Participatory approaches to development. A review of the literature'. *ODI working paper*. London: Overseas Development Institute.
- Narayan, D. et al. (2000) *Voices of the Poor: Can Anyone Hear Us?* Oxford: Oxford University Press.
- Nussbaum, M. (2001) *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.
- OECD (2011a) *How's Life?: Measuring well-being*. Paris: OECD.
- OECD (2011b) *Better Lives Initiative: Compendium of OECD Well-being Indicators*. Paris: OECD.
- OECD (2013a) *Proceedings: 4th World Forum on Statistics, Knowledge and Policy*. Paris: OECD.
- OECD (2013b) *Guidelines for Measuring Subjective Well-being*. Paris: OECD.
- OECD (2013c) *Measuring well-being and progress*. Paris: OECD.
- Oxfam (2013) *The Oxfam Humankind Index: The New Measure of Scotland's Prosperity, Second Results*.
- Qizilbash, M. (2006) 'Well-being, Adaptation and Human Limitations', *Philosophy* 59 (Supplement): 83-109.
- Ravallion, M. (2010) 'Mashup Indices of Development'. *World Bank Policy Research Working Paper* 5432. Washington, DC: World Bank.
- Rodriguez-Takeuchi, L.K (2013) 'Incorporating people's values in development: Weighting alternatives'. *ODI Methodological Background Note* (in press).
- Rojas, M. (2004) 'The Complexity of Well-being A Life Satisfaction Conception and a Domains-of-Life Approach'. Paper for the International Workshop on Researching Well-being in Developing Countries. Hanse Institute for Advanced Study Delmenhorst, near Bremen, Germany 2-4th July 2004.
- Rojas, M. (2008) 'Experienced Poverty and Income Poverty in Mexico: A Subjective Well-Being Approach', *World Development* 36(6): 1078-1093.
- Russell, L.B. and Hubley, A.M. (2005) 'Importance ratings and weighting: Old concerns and new perspectives', *International Journal of Testing* 5(2): 105-130.
- Ruta, D.A., Camfield, L., and Martin, F. (2004) 'Assessing individual quality of life in developing countries: piloting a global PGI in Ethiopia and Bangladesh', *Quality of Life Research* 13: 1545.
- Salomon, J.A and Murray, C.J.L. (2004) 'A multi-method approach to measuring health-state valuations', *Health Economics* 13(3): 281-90.
- Samman, E. (2007) 'Psychological and subjective well-being: a proposal for internationally comparable indicators', *Oxford Development Studies* 35(4): 459-486.
- Santos, M.E and Ura, K. (2008) 'Multidimensional Poverty in Bhutan: Estimates and Policy Implications', *OPHI Working Paper* 13. Oxford: Oxford Poverty and Human Development Initiative.
- Schmidt, S., Bullinger, M. (2007) 'Cross-cultural quality of life assessment approaches and experiences from the health care field' in Gough, I. & McGregor, A., *Well-being in Developing Countries*. Cambridge: Cambridge University Press.
- Sen, A. (1999) *Development as Freedom*. Oxford: Oxford University Press, 1999.
- Stiglitz J.E., Sen, A. and Fitoussi J-P. (2010) *Mismeasuring our lives: why GDP doesn't add up*. New York: The New Press.

-
- Stiglitz, J.E., Sen, A. and Fitoussi, J.-P. (2009) 'Report by the Commission on the Measurement of Economic Performance and Social Progress' Sustainable Development Commission on the Measurement of Economic Performance and Social Progress.
- Thomas, J. W. and Grindle, M. (1990) 'After the Decision: Implementing Policy Reforms', *World Development* 18(8): 1163-1181.
- UN (2012) 'Rio+20 Policy Brief #6: Human well-being for a planet under pressure'. Policy brief commissioned by the international conference Planet Under Pressure: New Knowledge Towards Solutions (www.planetunderpressure2012.net).
- Van Kempen, L. (2009) 'The 'Downside' of Women Empowerment in India: An Experimental Inquiry into the Role of Expectations', *Soc Indic Res*, 94: 465-482.
- van Rijsbergen, B., D'Exelle, B. (2011) 'Use of Delivery Care in Tanzania: The Importance of Poverty, Empowerment and Emergency'. *DEV working paper series*. UEA, UK.
- Vibert, F. (2011) 'Regulation in an age of austerity: Reframing international regulatory policies'. *LSE Working Paper*.
- Walker, P., Michaelson, J., Strauss, K and K. Trebeck. (2012) 'Technical Annex, Oxfam Humankind Index for Scotland – Background Methodology, Consultation and Results' (http://policy-practice.oxfam.org.uk/~media/Files/policy_and_practice/poverty_in_uk/HKI/HKI%20consultation%20methods%20results%20April%202012.ashx).
- Ware, J. E., Sherbourne, C. D. (1992) 'The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection', *Medical Care* 30: 473-483.
- White, S. and Ellison, M. (2007) 'Well-being, Livelihoods and Resources in Social Practice' in Gough, I and J. A. McGregor (eds.), *Well-being in Developing Countries: From Theory to Research*. Cambridge: Cambridge University Press.
- Woodcock, A., Camfield, L. and J.A.McGregor (2009) 'Validation of the WeDQoL-Goals-Thailand Measure: Culture-Specific Individualised Quality of Life', *Soc. Indic. Res.* 94: 135-171.
- Yamamoto, J., Feijoo, A.R., Lazarte, A. (2008) 'Subjective well-being: an alternative approach' in Copestake, J.G. (ed.), *Well-being and development in Peru: global and local views confronted*. New York: Palgrave Macmillan.

Overseas Development Institute

203 Blackfriars Road
London SE1 8NJ

The Institute is limited by guarantee
Registered in England and Wales
Registration no. 661818
Charity no. 228248

Contact us

developmentprogress@odi.org.uk
T: 020 7922 0300

Sign up for our e-newsletter

developmentprogress.org/
sign-our-newsletter

Follow us on Twitter

twitter.com/dev_progress

Project Notes

Outline the project approach and provide
background and methodological information.

Disclaimer

The views presented in this paper are those of the
author(s) and do not necessarily represent the views
of ODI.

© Overseas Development Institute 2015. Readers
are encouraged to quote or reproduce material for
non-commercial use. For online use, please link to
the original resource on the Development Progress
website. As copyright holder, ODI requests due
acknowledgement and a copy of the publication.

This Project Note was funded by the Bill & Melinda Gates
Foundation as part of 'Development Progress', a four-year
research project which aims to better understand, measure
and communicate what has worked in development and
why. The findings and conclusions contained within are
those of the author and do not necessarily reflect the
positions or policies of the Bill & Melinda Gates Foundation.

Acknowledgements

The authors would like to acknowledge with
gratitude comments on this draft provided by
Nancy Devlin. Figures were designed by Patrick
Ward.

Authors' affiliations

Allister McGregor
Professor of Political Economy, Department of Politics,
University of Sheffield (j.a.mcgregor@sheffield.ac.uk)

Sarah Coulthard

Senior Lecturer, Department of Social Sciences & Languages,
Northumbria University (sarah.coulthard@northumbria.ac.uk)

Laura Camfield

Senior Lecturer, School of International Development,
University of East Anglia (l.camfield@uea.ac.uk)