

Relevant Criteria for Selecting Indicators

A Proposal

David A. Clark^{*}

... indicators are precisely that... we cannot expect them to be a complete representation of society. They are simply an indication. The nature of that indication will depend on the choices made with regard to definitions and with regard to data. Different indicators highlight different features of social problems, and suggest different priorities for policy intervention (Atkinson et al, 2002, pp.19-20).

1. Introduction

Developing an equality measurement framework for the UK – or indeed any other country – involves making a number of critical decisions. These include choosing relevant dimensions for assessing inequality and disadvantage across minority groups (including women, children and the elderly); and selecting appropriate indicators for gauging performance in these dimensions.

It is now widely accepted that poverty, inequality and well-being are complex multidimensional phenomena. Multiple lists of human capabilities, needs and rights have emerged over the last couple of decades (spurred partly by the seminal contributions of Martha Nussbaum, 1990, 1995). At the same time several notable attempts to develop methodologies, procedures and rules for compiling lists have emerged (most notably, Nussbaum, 1995, 2000; Alkire, 2002, 2006; Clark, 2002, 2003; Robeyns, 2003; Clark and Qizilbash, 2008; Burchardt and Vizard, 2007, 2008; Ibrahim, 2008). Less attention, however, has generally been devoted to the selection

^{*} Background paper for the Equality and Human Rights Commission (EHRC) project on 'Developing the Equality Measurement Framework: Selecting the Indicators', Centre for the Analysis of Social Exclusion, London School of Economics and Political Science, 8 September 2008. Not for quotation or circulation. I have benefited from useful conversations with Tania Burchardt, David Hulme, Giovanni Razzu, Mozaffar Qizilbash, Polly Vizard and colleagues at CASE. The usual disclaimers apply. Email: David_A_Clark@hotmail.com

of relevant indicators for assessing advantage within given dimensions, which is the focus of this brief.

2. Background and Methodological Issues

The goal is to select relevant indicators for measuring aspects of inequality and disadvantage in the UK across ten pre-defined domains of well-being, which have emerged from the human rights approach and deliberative consultations with the public and individuals and groups at risk of discrimination and disadvantage (see Burchardt and Vizard, 2007, 2008; EHRC, 2008). Instead of combining indicators into a single composite measure of disadvantage (which conceals at least as much as it reveals about inequality¹), an attempt will be made to identify up to thirty spotlight indicators (approximately three indicators per domain) that highlight key aspects of inequality and disadvantage.

A number of methodological challenges are involved in selecting relevant indicators, which include whether to measure aspects that (1) are intrinsic or instrumental; (2) are context specific or universal; (3) relate to individuals or broader social units; (4) focus on levels of achievement or dynamics; (5) help establish causality; and (6) are defined from the bottom up rather than by experts. Further discussion of these issues can be found in Ibrahim and Alkire (2007).

3. Relevant Criteria

Relatively little research has been devoted to developing criteria for selecting relevant indicators for monitoring well-being within or across domains. Some particularly noteworthy contributions include Liu (1974), Atkinson et al (2002), World Bank (2004), Ibrahim and Alkire (2007) and OPHI (2007). On the other hand

relatively *ad-hoc* discussions of the merits of various indicators (which are less helpful for our purposes) are fairly common in the literature (e.g. Hicks and Streeten, 1979; Qizilbash, 1997, pp. 214-215, Chiappero Martinetti, 2000, pp.210-212; Comim, 2008).

Some possible criteria for selecting indicators for measuring poverty, inequality and well-being are summarised in Annexes A-H. Any criteria for selecting relevant indicators must take account of the nature and purpose of the evaluative exercise: namely, in the present context, investigating inequality and disadvantage across disadvantaged groups in the UK (see, for example, EHRC, 2008). The remainder of this section draws on existing literature (Annexes A-H)², discussions with colleagues, and brainstorming at CASE in an attempt to compile a detailed list of relevant principles for choosing between indicators:

- 1) *Relevance for assessing equality and human rights.* Indicators should relate to the lives of people experiencing discrimination and disadvantage and should help us track social problems (A1, B1, G11).³ In gauging relevance, we might consider if the indicator in question is recognised as meaningful by stakeholders, acceptable to the general public and individuals and groups at particular risk of discrimination and disadvantage, understandable in the sense that it is likely to produce results that seem reasonable, and has a clear normative interpretation so that movement in one direction represents an unambiguous improvement (B1; H1).
- 2) *Relevance in terms of coverage versus salience.* Another criterion might be relevance for assessing the nature and extent of inequality *within* specific domains (like B7, B8, C7a, C7b, D1, D4, E3). In this respect, it is important to

recognise that most dimensions of inequality and disadvantage are themselves multi-faceted ('the capability to be alive' might be an exception).⁴ Thus, choosing the best overall proxy indicator – or combination of indicators – to reflect performance across sub-dimensions is a challenge. For this project, the relevant criterion has more to do with *salience* or selecting indicators that highlight specific aspects of inequality which are especially important for one or more groups (more like A1; B1 or criterion 1 above).⁵ This should be kept in mind as there is an important difference between selecting indicators that highlight specific problems in a given domain and choosing indicators that help provide a balanced summary of inequality within that domain.⁶

- 3) *Public participation, legitimacy and ownership.* The selection of indicators should be scrutinised and endorsed by the public in general and by individuals and groups at risk of discrimination and disadvantage in particular (D2). While there should be a general consensus about the importance of indicators, an effort should be made to incorporate differences of opinion and to facilitate socially or culturally specific needs. In short, the drive for consensus should not become a straightjacket. These concerns reflect the belief that any reasonable list of indicators should be endorsed from the bottom up. There is also a strong case for consulting interest groups, users and other stakeholders.⁷ Apart from promoting legitimacy and ownership of the final product, consultation is a means of ensuring that selected indicators are relevant and salient (criteria 1 and 2 above).
- 4) *Disaggregation of statistics by population sub-groups.* Indicators, where possible and meaningful, should be additively decomposable by population sub-groups. In other words, it should in principle be possible to disaggregate

statistics at least by gender, transgender, disability, religion and belief, ethnicity, sexual orientation, age and social class for the purpose of analysis. In the context of inequality and disadvantage, it is necessary to consider a particularly wide range of social divisions (e.g. occupation, education level) some of which may be quite fine (e.g. family type, immigration status).

- 5) *Comparability across space and over time.* As a bare minimum, selected indicators must be available across Britain. For our purposes it is important to be able to compare indicators across England, Scotland and Wales and highly desirable to breakdown indicators by region and local area.⁸ In an ideal world selected indicators would also be comparable internationally.⁹ A great deal of emphasis is typically placed on the capacity to compare indicators across countries (A2, B4, C6), although this is less of a priority for our purposes. Moreover, an indicator should be timely (in the sense that it provides up-to-date information) and revisable in terms of data and underlying concepts (B5; D4).
- 6) *Ideally indicators should relate to individuals.* Whenever possible, indicators should relate to individuals rather than households. Among other things, this helps facilitate the analysis of intra-household inequalities (which are particularly relevant for studying some forms of discrimination), avoids the methodological problems associated with counting households of different sizes/ compositions, and improves the scope for making comparisons (as we are comparing indicators that relate to individuals only rather than indicators that relate to individuals and households). More fundamentally, the focus on individuals is more consistent with the capability approach and human rights perspectives, which place people at the centre of analysis by focusing on the

substantive freedoms or rights of value (e.g. Sen, 1999; Gaay-Fortman, 2006, p. 261). As Atkinson et al (2002) argues the fundamental concern is with the position of a country's citizens, which 'points to counting persons' (p. 29).¹⁰

7) *Indicators should emphasise results in terms of outcomes, processes or autonomy.* Indicators for measuring disadvantage and inequality among individuals and groups should focus on results rather than inputs (A3; F2).¹¹

This is because inequality and disadvantage should ultimately be judged in terms of what people can or cannot 'do' and 'be' instead of the assets and resources they can command (c.f. Sen, 1985; 1999) or the institutional and policy context in which they live (e.g. DFID, 1999-2001; Ellis, 2006).¹² In this respect, such indicators might relate to inequality of outcome (i.e. disparities in terms of achievement), inequality of process (unequal treatment by other people or institutions) or inequality of autonomy (i.e., in terms independence, choice or control) (see EHRC, 2008).¹³

8) *Whenever possible and appropriate indicators should be dynamic rather than static.* In other words, it is useful to select indicators that allow us to identify change or lack of change in an individual's capabilities over time (A4, H2 vs. H6; Atkinson et al, 2002, pp. 31-32). Such an approach allows us to move beyond static comparisons of poverty trends (how many overall became richer or poorer) to a more nuanced approach that considers poverty dynamics (which people or groups moved into or out of poverty over time and which people or groups remain trapped in persistent poverty). This permits a much richer analysis of the causes of poverty and disadvantage as well as corresponding policy options (CPRC, 2004, Hulme, 2006; Clark and Hulme, 2008). Indicators may also be forward looking in the sense that they allow for

the fact that people are disadvantaged not just because they are deprived in some sense (e.g. unemployed), but because they lack effective opportunities to overcome their deprivation (e.g. find employment) in the future (Atkinson et al., 2002, p. 32; Sen, 1999).

9) *Relevance for Public Policy*. Selected indicators should be sensitive to effective policy interventions (B3, C6, E4, H3-H5), especially in the area of equality and human rights. Having said this, care must be taken not to exclude highly relevant indicators of inequality on the grounds that they are not especially sensitive to policy interventions (perhaps, because the results of successful interventions only show up in official statistics over relatively long periods of time – an example might be life expectancy among disadvantaged groups, especially if these groups are starting from relatively high base rates). The design and selection of indicators should also try to minimize possible incentives to improve scores through artificial policy changes or the manipulation of statistics (B3; H6).

10) *Accuracy, reliability, and validity*. Ideally selected indicators should have been previously tested and found to be adequate for research purposes (A4). In particular, indicators should be accurate, reliable and valid (A4, B2, C5, D4, G8 and G12, etc, H9), although most existing criteria for selecting indicators do not explicitly define these terms (B2 is an exception). For our purposes key statistical properties of indicators are defined as follows: (i) *accuracy* in terms of sample size – the available data supports analyses of relevant population sub-groups; (ii) *reliability* in terms of responses to questions and test/ re-test – it has been shown that a specific respondent provides consistent answers to the same question; and (iii) *validity* – in terms of inherent meaning – the indicator

11) *Selected indicators should not impose too large a burden on government agencies or the general public.* Proposals for new indicators should take account of the cost implications (G3, G10, H7) and should be proportionate to the needs of users (B6). In addition, they should not place an excessive burden on respondents (G9).

The following issues are likely to be the subject of ongoing debate. Should any of the criteria be revised or removed? What, if anything, should be added to our criteria? Are some principles more important than others? Is it helpful to distinguish between principles that apply to single indicators (B1-B6), principles that apply to the balance of indicators within a domain and principles that apply to the portfolio of indicators as a whole (B7-B9)?

Finally, what should be done if indicators satisfy some principles but not others? Should we permit or even encourage trade-offs in some areas? The issue of how to deal with potential clashes between different criteria is fundamental, but has not been acknowledged or discussed in any of the literature reviewed (e.g. Atkinson et al., 2002; World Bank, 2004; Ibrahim and Alkire, 2007).

5. Application

These principles are refined and prioritised in line with EHRC's requirements and are synthesised into a more practical and manageable checklist for application purposes below.¹⁴ This involves distinguishing between essential and desirable criteria for selecting indicators on the one hand and principles that apply to the

selection of single indicators, principles that apply to the balance of indicators within a specific domain and criteria that apply to the whole portfolio of indicators on the other. The following criteria draw inspiration from EHRC (2008) and have been adjusted following feedback¹⁵ from the Government Equality Office's Policy and Analysis Group:

Criteria that apply to single indicators

ESSENTIAL CRITERIA

1. Relevance for assessing a specific form of disadvantage or inequality in a specified domain, which is likely to be particularly important for one or more groups of people¹⁶ (criteria 1 and 2 above);
2. The indicator in question qualifies as relevant and legitimate in the sense that it is endorsed (and rated highly vis-à-vis other potential indicators) by relevant stakeholders (criterion 3 above)¹⁷;
3. The indicator can be disaggregated by population sub-group – the most relevant for our purposes are gender, ethnicity, disability, age, sexual orientation, religion and belief, and social class (criterion 4 above);
4. For specific indicators, measures are available that permit monitoring across all three countries that constitute Britain¹⁸ (i.e. England, Scotland and Wales), although the sources and technical specification of some measures may differ (criterion 4 above). In this respect, some indicators will be strictly comparable across Britain, while others are only broadly or loosely comparable;
5. The indicator captures a result in terms of an outcome (achievement) or process (discrimination or other forms of disadvantage such as lack of dignity or respect) (criterion 7 above);

6. The indicator is subject to the standard statistical requirements of accuracy, reliability and validity (criterion 10 above);

DESIRABLE CRITERIA

7. The indicator can be disaggregated by additional population sub-groups such as family type, asylum and refugees status (in addition to the groups agreed with EHRC) (criterion 4 above);
8. The indicator can be disaggregated by regions and local areas (criterion 5 above);
9. The indicator is comparable over time in the sense that existing time series data is available and the statistic in question is collected reasonably frequently for monitoring purposes¹⁹ (criterion 5 above);
10. The indicator is comparable internationally – especially with other EU states (criterion 5 above);
11. Where possible, the indicator should relate to individuals rather than households or broader social units (criterion 6 above);
12. When appropriate, the indicator is *dynamic* in the sense that it helps identify change or lack of change in valuable capabilities over time – this involves being able to track individuals over time and distinguish between those who become disadvantaged, those who manage to escape disadvantage and those who are persistently disadvantaged over time (criterion 8 above);
13. When appropriate, the indicator is sensitive to effective policy interventions without being readily susceptible to manipulation (criterion 9 above);
14. Proposals for new indicators should take account of the cost implications and should be proportional to the needs of stakeholders (criterion 11 above). New

indicators should only be proposed in cases where reasonably close alternatives or suitable proxy indicators are not available;

Criteria that apply to the balance of indicators within each domain

ESSENTIAL CRITERIA

15. The selected indicators for each domain should highlight the most important aspects of disadvantage and inequality in that domain for each and every group of people (see criterion 2 above);
16. The selected indicators for each domain should include measures that either focus on the whole distribution (e.g. the Gini coefficient) or different parts of the same distribution (e.g. the proportion of people on low incomes, middle incomes and high incomes) (resonates with criterion 2 above).

Criteria that apply to the whole portfolio of indicators

ESSENTIAL CRITERION

17. The portfolio of indicators should be easy to communicate, interpret and comprehend (B9, H8; see also G6, G13, G14);

DESIRABLE CRITERION

18. The portfolio of indicators should be balanced across domains (B7, B8);

It is possible that there might be significant trade-offs between principles 15 and 18 (as implied in our discussion of criterion 2 above). If this is the case, background papers should probably flag any key aspect(s) of disadvantage and inequality that proposed indicators are not especially well placed to identify and track.

Most of the existing approaches apply their criteria for selecting indicators in fairly implicit or ad-hoc ways (e.g. Ibrahim and Alkire, 2007). A more robust and transparent methodology for applying our criteria is required. This involves developing a matrix that will systematically rank each indicator against our criteria (see the two selection matrixes for essential and desirable criteria available separately). Such a matrix will form the backbone of the ten domain papers we plan to write.

The following steps are anticipated:

- Map each sub-domain (e.g. ‘the capability to be alive’) to possible indicators (e.g. life expectancy, infant mortality rate, suicide rate);
- Evaluate each indicator against ‘essential’ criteria;
- Add indicators that satisfy essential criteria to the long list of potential indicators;
- Add indicators that fail to satisfy essential criteria to the rejection list;
- Evaluate each of the indicators on the long list against ‘desirable’ criteria;
- Add indicators that satisfy desirable criteria to the ‘short list’ (we may want to consider some kind of scoring mechanism);
- Reflect critically on the indicators included on the short list and long list and revise in line with intuition and common sense (E5). (Any such adjustments should be transparent and explicitly justified). This will involve drawing on principles that apply to the balance of indicators within a domain (criteria 15 and 16) and principles that apply to the portfolio of indicators as whole (criteria 17 and 18);

- Revise the short list and long list in line with internal discussion and stakeholder consultation.

Notes

¹ This is for the same reasons discussed by Amartya Sen in his critique of the human development index: The human development index ‘is a quick and imperfect glance at human lives... the breadth of the human development approach must not be confused with the slender specificity of the Human Development Index’ (Sen, 2006, p. 257).

² An extremely rapid literature review has been conducted using Google Scholar, Econ Lit and the Social Science abstracts.

³ A1 and B1 refer to Annex A, point 1 and Annex B, point 1 respectively, and so on.

⁴ In other words, they can be sub-divided into distinct components as the structure and composition of most existing list – including our own – clearly demonstrate (see Burchardt and Vizard, 2008, appendix 3).

⁵ Clark and Qizilbash’s (2001) provide an extensive list of possible indicators for each dimension of well-being. These include many potential indicators for which data is not collected. This document could be used as a checklist for considering new indicators.

⁶ In this respect one potentially usefully type of information (which concerns the extent and nature of inequality) is being traded for another (which involves highlighting the most urgent or pressing forms of inequality and disadvantage).

⁷ See Clark and Alkire (2008) on the relative merits of public deliberation vis-à-vis expert opinion.

⁸ In some cases it may also be desirable to distinguish between urban and rural areas, breakdown statistics by town and village or disaggregate indicators according to income, consumption or asset ownership levels (see World Bank, 2004).

⁹ It may be especially helpful for the UK government to be able to compare performance with other EU member states.

¹⁰ While these are strong arguments for focusing on indicators that relate to individuals, it is possible that in some cases or circumstances people may want to define themselves in terms of groups. In such cases, we would not necessarily want to preclude the use of broader indicators. The planned consultation events may shed further light on this issue.

¹¹ Clark (2002) briefly reflects on some of the difficulties associated with distinguishing between means and ends.

¹² Of course, it goes without saying that these are important for instrumental reasons and deserve monitoring (Clark, 2005; World Bank, 2004).

¹³ The first two categories are broadly equivalent to Sen's (1999, p.17) distinction between the opportunity and process aspects of freedom. EHRC has not given our project a mandate to investigate suitable indicators of autonomy. However, in practice there is likely to be overlap between some of the relevant indicators for the opportunity, process and autonomy aspects of freedom.

¹⁴ It goes without saying that potential indicators can be subjective as well as objective. However, measures based on questions that deal with attitudes, opinions or aspirations may not qualify as outcome, process or autonomy indicators.

¹⁵ In addition helpful written comments were received from Giovanni Razzu and Liz Speed.

¹⁶ This will be based on a preliminary assessment by experts at CASE, which will feed into, and be revised in light of, our consultations with stakeholders.

¹⁷ There is a strong case for broadening the agreed group of stakeholders to include the general public and representatives from additional minority groups (e.g. gypsies, asylum seekers, etc). As discussed above, to qualify as relevant an indicator ought to be endorsed from the bottom up.

¹⁸ Indicators deemed to be locally relevant notwithstanding, as it is not desirable to exclude country specific indicators on the grounds that they cannot be compared across Britain.

¹⁹ In terms of frequency, an indicator should be collected at least every three years to comply with EHRC's reporting requirements.

References

- Alkire, S. (2002), *Valuing Freedoms: Sen's Capability Approach and Poverty Reduction*, Oxford University Press, Oxford.
- Alkire, S. (2006), 'Public Debate and Value Construction in Sen's Capability Approach' in A. Kaufman (ed.) *Capabilities Equality: Basic Issues and Problems*, Routledge, London, pp.133-154.
- Atkinson, T., Cantillon, B., Marlier, E., Nolan, B. (2002), *Social Indicators: The EU and Social Inclusion*, Oxford University Press, Oxford.
- Bayless, M. and Bayless, S. (1982), 'Current Quality of Life Indicators: Some Theoretical and Methodological Concerns', *American Journal of Economics and Sociology*, 41(4), pp. 421-437.
- Burchard, T. and Vizard, P. (2007), 'Developing A Capability List: Final Recommendations of the Equalities Review Steering Group on Measurement', *CASE Working Paper 121*, Centre for the Analysis of Social Exclusion, London School of Economics and Political Sciences, April.
- Burchardt, T. and Vizard, P. (2008), 'Developing an Equality Measurement Framework: Developing A List of Substantive Freedoms for Children, and Refining The Existing List for Adults', Final Project Report, London School of Economics and Political Science, May 2008.
- Chiappero Martinetti, E. (2000), 'A Multidimensional Assessment of Well-Being Based on Sen's Functioning Theory', *Rivista Int Di Scienze Sociali*, 108, pp. 207-231.
- Clark, D. A. (2002), *Visions of Development: A Study of Human Values*, Edward Elgar, Cheltenham.

- Clark, D. A. (2003), 'Concepts and Perceptions of Well-Being: Some Evidence from South Africa', *Oxford Development Studies*, 31(2), pp. 173-196.
- Clark, D. A. (2005), 'Sen's Capability Approach and the Many Spaces of Human Well-Being', *Journal of Development Studies*, 41(8), pp. 1339-1368.
- Clark, D. A. and Alkire, S. (2008), 'Weights as Value Judgments: An Assessment of Expert Opinion and Participatory Approaches', *OPHI Working Paper 19*, Oxford Poverty and Human Development Initiative, Queen Elizabeth House, University of Oxford, Oxford (forthcoming September 2008).
- Clark, D. A. and Hulme, D. (2008), 'Poverty, Time and Vagueness: Integrating the Core Poverty and Chronic Poverty Approaches', mimeographed, University of Manchester, UK.
- Clark, D. A. and Qizilbash, M. (2001), *Questionnaire Schedule: The Essentials of Life*, Southern Africa Labour and Development Research Unit, University of Cape Town, South Africa.
- Clark, D.A. and Qizilbash, M. (2008), 'Core Poverty, Vagueness and Adaptation: A New Methodology and Some Results for South Africa', *Journal of Development Studies*, 44(4), pp. 519-544.
- Comim, F. (2008), 'Measuring Capabilities' in F. Comim, M. Qizilbash and S. Alkire (eds), *The Capability Approach: Concepts, Measures and Applications*, Cambridge University Press, Cambridge.
- CPRC (2004), *Chronic Poverty Report 2004/05*, Chronic Poverty Research Centre, University of Manchester, UK.
- DFID (1999-2001), *Sustainable Livelihoods Guidance Sheets*, UK Department for International Development (DFID), London. Available online at http://www.livelihoods.org/info/info_guidancesheets.html#1 (accessed 8 September 2008).
- EHRC (2008), 'The Equality Measurement Framework', Equality and Human Rights Commission (EHRC) Briefing Note, unpublished draft, forthcoming.
- Ellis, F. (2006), 'Livelihoods Approach' in D. A. Clark (ed.), *The Elgar Companion to Development Studies*, Edward Elgar, Cheltenham, pp. 346-349.
- Eurostat (2005), *European Statistics Code of Practice: For the National and Community Statistics Authorities*, available online at http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2273,1,2273_47141302&_dad=portal&_schema=PORTAL
- Gaay-Fortman, B. de (2006), 'Human Rights' in D. A. Clark (ed.), *The Elgar Companion to Development Studies*, Edward Elgar, Cheltenham, pp.260-266.
- Hicks, N. and Streeten, P. (1979), 'Indicators of development: The search for a basic needs yardstick', *World Development*, 7, pp. 567-580.
- Hulme, D. (2006), 'Chronic poverty', in D. A. Clark (ed.), *The Elgar Companion to Development Studies*, Edward Elgar, Cheltenham, pp. 61-67.
- Ibrahim, S. (2008), *The Self-help Initiatives of the Poor: The Road to Sustainable Poverty Reduction in Egypt*, PhD Thesis, University of Cambridge, Cambridge, UK.
- Ibrahim, S. and Alkire, S. (2007), 'Agency and Empowerment: A proposal for internationally comparable indicators', Workshop on Missing Dimensions of Poverty Data, University of Oxford, 29-30 May 2007.
- Liu, B-C. (1974), 'Quality of Life Indicators: A Preliminary Investigation', *Social Indicators Research*, 1, pp. 187-208.

- OPHI (2007), 'Missing Dimensions of Poverty Data: Background Information on the Indicators and Survey Modules', Oxford Poverty and Human Development Initiative, University of Oxford, October 2007.
- Moller, V. and Schlemmer, L. (1989), 'South African Quality of Life: A Research Note', *Social Indicators Research*, 21, pp. 279-291.
- Nussbaum, M.C. (1990), 'Aristotelian Social Democracy', in B. Douglas, G. Mara and H. Richardson (eds), *Liberalism and the Good*, Routledge, New York, pp. 203-52.
- Nussbaum, M.C. (1995), 'Human Capabilities, Female Human Beings', in M.C. Nussbaum and J. Glover (eds), *Women, Culture and Development*, Oxford: Clarendon Press, pp. 61-104.
- Nussbaum, M.C. (2000), *Women and Human Development: the Capabilities Approach*, Cambridge University Press, Cambridge.
- Qizilbash, M. (1997), 'Pluralism and Well-Being Indices', *World Development*, 25(12), pp. 2009-2026.
- Robeyns, I. (2003), 'Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities', *Feminist Economics*, 9(2-3), pp. 61-92.
- Sharpe, A. (1999), 'A Survey of Indicators of Economic and Social Well-being', Background paper prepared for Canadian Policy Research Networks, July 1999. http://epe.lac-bac.gc.ca/100/200/300/cprn/english/sie_e.pdf
- Sen, A. K. (1985), *Commodities and Capabilities*, Elsevier Science Publishers, Amsterdam, North-Holland.
- Sen, A. K. (1992), *Inequality Re-Examined*, Oxford University Press, Oxford.
- Sen, A. K. (1999), *Development As Freedom*, Oxford University Press, Oxford.
- Sen, A. K. (2006), 'Human Development Index' in D. A. Clark (ed.), *The Elgar Companion to Development Studies*, Edward Elgar, Cheltenham, pp.256-260.
- World Bank (2004), 'Poverty Monitoring Guidance Note 1: Selecting Indicators', Poverty Reduction Group, Poverty Reduction and Economic Management Network, World Bank. Available online at <http://www.worldbank.org/poverty>, under Topics, Poverty Monitoring, Selecting Indicators.

Annex A: Oxford's Criteria for Selecting Indicators

Ibrahim and Alkire's Criteria for Selecting Indicators to Measure Empowerment

- (1) The chosen indicator should be relevant to the lives of the poor and the areas in which they suffer from a power deficit
- (2) The chosen indicators need to be *internationally comparable*
- (3) The indicators need to assess not only the instrumental but also the *intrinsic* aspects of empowerment
- (4) As empowerment is a process, it is essential to select indicators that would be able to identify changes in agency and empowerment over time
- (5) The choice of the indicators' short list draws on experience with particular indicators. That is, these indicators have [been] previously tested and found to be adequate for research purposes and their shortcomings have been identified. Indicators need to be scrutinised on standard conditions of accuracy, validity and reliability.

Source: Ibrahim and Alkire (2007), pp. 18-19.

Note: Essentially the same as OPHI's criteria for selecting indicators (see below) with one addition (relevance, criterion 1).

OPHI's Criteria for Selecting Indicators to Measure Poverty

- (1) The indicators need to be *internationally comparable*
- (2) The indicators [should] assess not only the instrumental but also the intrinsic aspects of [each] of the dimensions
- (3) It [is] essential to select indicators that [can] identify changes in dimensions over time;
- (4) The choice of indicators [should] draw on experience with particular indicators to date, i.e. how frequently these indicators have been previously fielded and found to be 'adequate' measures for research purposes.

Source: OPHI (2007), p. 2.

Annex B: Atkinson's Criteria for Selecting Social Indicators

Principles that apply to single indicators:

1. *An indicator should identify the essence of the problem and have a clear and accepted normative interpretation. An indicator should be*
 - recognised as meaningful by all users
 - acceptable to the general public
 - understandable and produce results that seem reasonable; and
 - typically have a clear normative interpretation so that movement in one direction represents an improvement.
2. *An indicator should be robust and statistically validated.*
 - An indicators should be measurable in a way that commands general support
 - The data employed should be statistically reliable and avoid arbitrary adjustments
 - Indicators should be validated with reference to other evidence and should be statistically reliable over time.
3. *An indicator should be responsive to effective policy interventions but not subject to manipulation. Indicators should*
 - reflect successful policy interventions
 - avoid temptation to improve scores through artificial policy changes
4. *An indicator should be measurable in a sufficiently comparable way across [the EU], and comparable as far as practicable with the standards applied internationally by the UN and the OECD.*
 - States should be encouraged to develop statistical information to improve the degree of comparability
 - Indicators that are sensitive to differences in social structures across states should be 'equitable' between countries with regard to differing sizes of rural populations and hence differing degrees of production for home consumption.
5. *An indicator should be timely and susceptible to revision. An indicator should*
 - provide up-to-date information (as far as possible)
 - be revisable in terms of data and underlying concepts where advances are made in terms of understanding and where there are changes in policy concerns.
6. *The measurement of an indicator should not impose too large a burden on member states, on enterprise, or on the [EU's] citizens.*
 - The design of social indicators should, wherever possible, make use of available statistics
 - Where new information is needed, then as far as feasible it should be obtained using existing instruments, for example by adding questions to existing surveys

Principles that apply to the whole portfolio of indicators:

7. *The portfolio of indicators should be balanced across different dimensions.*

- Costs in terms of lost transparency from having too extensive a range of indicators.
- Risk of losing credibility, if member states can simply pick and choose.

8. *The indicators should be mutually consistent and the weight of single indicators in the portfolio should be proportionate.*

- [Atkinson does not explicitly define mutually consistency]
- Interpretation of the set of indicators is greatly eased where individual components have degrees of importance that, while not necessarily equal, are not grossly different.

9. *The portfolio of indicators should be as transparent and accessible as possible to the citizens of the European Union.*

- Indicators should be easy to read and understand.
- Dissemination of the results of indicators, and of information about the method of their construction is important.

Source: Extracted from Atkinson et al (2002), pp. 20-25.

Annex C: ISQOL's Criteria for Evaluating Societal Indicators

- (1) The QOL index have a clear practical purpose, i.e., a public policy purpose;
- (2) The QOL index be grounded in well-established theory;
- (3) The QOL index be reported as a single number, but should be able to be broken down into components, similar to the index of leading economic indicators;
- (4) The QOL index be based on time series to allow the periodic monitoring and control;
- (5) The composite QOL index should be reliable, valid, and sensitive as should be the components making up the composite index;
- (6) The measure should help public policy makers develop and assess programs at the individual level (e.g., physicians and counselors helping individuals in need), the family or household level (e.g., social workers helping families in need), community level (e.g., town governments developing policies and programs that can enhance community QOL), state (or province) level (e.g., state bodies developing policies and programs that can assist residents of the entire state or province), the country level (e.g., national agencies developing policies and programs that can assist citizens of that country), and the international level (e.g., international agencies developing policies and programs that can assist the world citizen and the planet at large).
- (7) The domains covered should have the following properties:
 - (a) In total, the domains must encompass the totality of life experience.
 - (b) Each domain must encompass a substantial but discrete portion of the QOL construct.
 - (c) Each domain must be able to be measured in both objective and subjective dimensions.
 - (d) Each domain within a generic QOL instrument must have relevance for all people.
 - (e) If a specific domain is proposed for a non-generic instrument (e.g., independent living skills) it must be demonstrated to contribute unique variance to the QOL construct beyond the generic domains for the target group.
 - (f) Domains must be potentially neutral, positive or negative in their contribution to the QOL construct. Thus all aspects of disease states and functional status cannot be domains since, in their most positive state where they are absent or maximized, respectively, their contribution to the QOL construct cannot be more than neutral.
 - (g) Domains differ from the dimensions of personality (e.g., extraversion, self-esteem), cognitive processes (e.g., cognitive dissonance) and affect (e.g., joy) in that they cannot be measured objectively.
 - (h) The subjective dimension of each domain has both a cognitive and an affective component. They are measured by questions concerning "satisfaction."

Source: Sharpe (1999), pp. 42-43.

Annex D: Ben-Chieh Liu's criteria for selecting QOL variables

Quality of life may be measured and evaluated on the basis of a series of representative variables which sometimes are called quality of life factors. Ideally, those factors should possess the following characteristics:

- (1) They should be sufficiently universal so that the underlying principles would apply to the majority of people in the U.S. today.
- (2) They should be easily understood and there should be a general consensus with respect to their selection.
- (3) They should be flexible enough to encompass any life style among individuals at different places and at different points in time.
- (4) They should be adaptable to changing social, economic, political and physical conditions in a dynamic society and open to verification according to recognized scientific approaches.

Source: Liu (1974), pp. 188-189.

Annex E: Moller and Schlemmer's criteria for selecting QoL indicators

The selection criteria included:

- 1) split sample reliability,
- 2) association with global measures,
- 3) comprehensive coverage of major domains of life,
- 4) relevance to social policy, and
- 5) in the last instance an element of common sense.

The index features global as well as specific concern satisfaction measures. A further recommendation was that quality of life should be assessed as a profile rather than an aggregate index of the 33 measures (Moller and Schlemmer, 1989, p. 287).

Annex F: Morris's criteria for selecting QoL variables

- (1) variables should not reflect specific values, i.e. they should be value free;
- (2) variables should measure results not inputs;
- (3) variables should reflect the distribution of social outputs; and
- (4) variables should be simple

Source: Bayless and Bayless (1982), p. 426.

Appendix G: The European Statistics Code of Practice

The European Statistics Code of Practice is based on 15 principles. Governance authorities and statistical authorities in the European Union commit themselves to adhering to the principles fixed in this code covering the institutional environment, statistical processes and outputs.

[While this is intended as a code of practice for government agencies rather than criteria for selecting relevant indicators, there is arguably a strong case for incorporating many of these principles in relevant selection criteria. Principles 1, 3-12 and 14-15 may be especially relevant.]

1. *Principle 1: Professional Independence.* The professional independence of statistical authorities from other policy, regulatory or administrative departments and bodies, as well as from private sector operators, ensures the credibility of European statistics.
2. *Principle 2: Mandate for Data Collection.* Statistical authorities must have a clear legal mandate to collect information for European statistical purposes. Administrations, enterprises and households, and the public at large may be compelled by law to allow access to or deliver data for European statistical purposes at the request of statistical authorities.
3. *Principle 3: Adequacy of Resources.* The resources available to statistical authorities must be sufficient to meet European statistics requirements.
4. *Principle 4: Quality Commitment.* All ESS members commit themselves to work and cooperate according to the principles fixed in the ‘Quality declaration of the European statistical system’.
5. *Principle 5: Statistical Confidentiality.* The privacy of data providers (households, enterprises, administrations and other respondents), the confidentiality of the information they provide and its use only for statistical purposes must be absolutely guaranteed.
6. *Principle 6: Impartiality and Objectivity.* Statistical authorities must produce and disseminate European statistics respecting scientific independence and in an objective, professional and transparent manner in which all users are treated equitably.
7. *Principle 7: Sound Methodology.* Sound methodology must underpin quality statistics. This requires adequate tools, procedures and expertise.
8. *Principle 8: Appropriate Statistical Procedures.* Appropriate statistical procedures, implemented from data collection to data validation, must underpin quality statistics.

9. *Principle 9: Non-Excessive Burden on Respondents.* The reporting burden should be proportionate to the needs of the users and should not be excessive for respondents. The statistical authority monitors the response burden and sets targets for its reduction over time.
10. *Principle 10: Cost Effectiveness.* Resources must be effectively used.
11. *Principle 11: Relevance.* European statistics must meet the needs of users.
12. *Principle 12: Accuracy and Reliability.* European statistics must accurately and reliably portray reality.
13. *Principle 13: Timeliness and Punctuality.* European statistics must be disseminated in a timely and punctual manner.
14. *Principle 14: Coherence and Comparability.* European statistics should be consistent internally, over time and comparable between regions and countries; it should be possible to combine and make joint use of related data from different sources.
15. *Principle 15: Accessibility and Clarity.* European statistics should be presented in a clear and understandable form, disseminated in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.

Source: Eurostat (2005).

Annex H: World Bank on the Qualities of Indicators

Good indicators share a number of features. These features can be used as a checklist when deciding which indicator to use. Indicators should:

1. *Be direct, unambiguous measure of progress*, i.e. more (or less) is invariably better. For instance, immunization coverage is less ambiguous than household expenditure on health, because an increase in such expenditure could be a good thing – if it means that households have more resources to get healthcare – or a bad thing – if it means that disease incidence has increased.
2. *Vary across group, areas, and over time*. For instance, child malnutrition is more likely to vary quickly over time than life expectancy.
3. *Have a direct link with interventions*. For instance, vehicle-operating costs depend on road quality but also on many other factors, such as international petrol prices. It might therefore not be a good indicator for progress in the roads sector.
4. *Be relevant for policy making*, i.e. at the level of disaggregation relevant for decision-making. For instance, if expenditures are decided at the regional level once a year, the indicator should be disaggregated at the regional level, and monitored on a yearly basis (see World Bank, 2004 for more on disaggregation).
5. *Be consistent with the decision-making cycle*. For instance, use indicators at intervals which match the decision making process, i.e. that can be available in time for budget discussions.
6. *Not be easily manipulated or blown off course by unrelated developments*. Some indicators can be very sensitive to external or exogenous factors and should be avoided. Others can be easily manipulated (e.g. where there is self-reporting, or where incentive structures are such that one might be tempted to under or over-estimate the result). For instance, some administrative data such as data on enrolment in schools or number of visits in health centers are sometimes used to allocate subsidies and budgets, and local authorities might have an incentive to over-report in order to obtain higher grants.
7. *Be easy to measure and not too costly to measure*. For instance, the number of deaths is typically easily recorded, while the number of cases of specific diseases is harder to track accurately. In addition, when large surveys are necessary, monitoring is more costly and requires more time.
8. *Be easy to understand*. For instance, poverty incidence is easier to understand and to communicate than poverty depth.
9. *Be reliable*. For instance, scientific, objective indicators are more reliable than indicators that depend on the interpretation of the user. This is related to the above discussion on ‘manipulation’.

10. *Consistent with data availability and data collection capacity.* This ensures that indicators will be measurable at the times and level selected (see World Bank, 2004 on the steps to choose indicators).

Source: World Bank (2004), pp. 3-4.