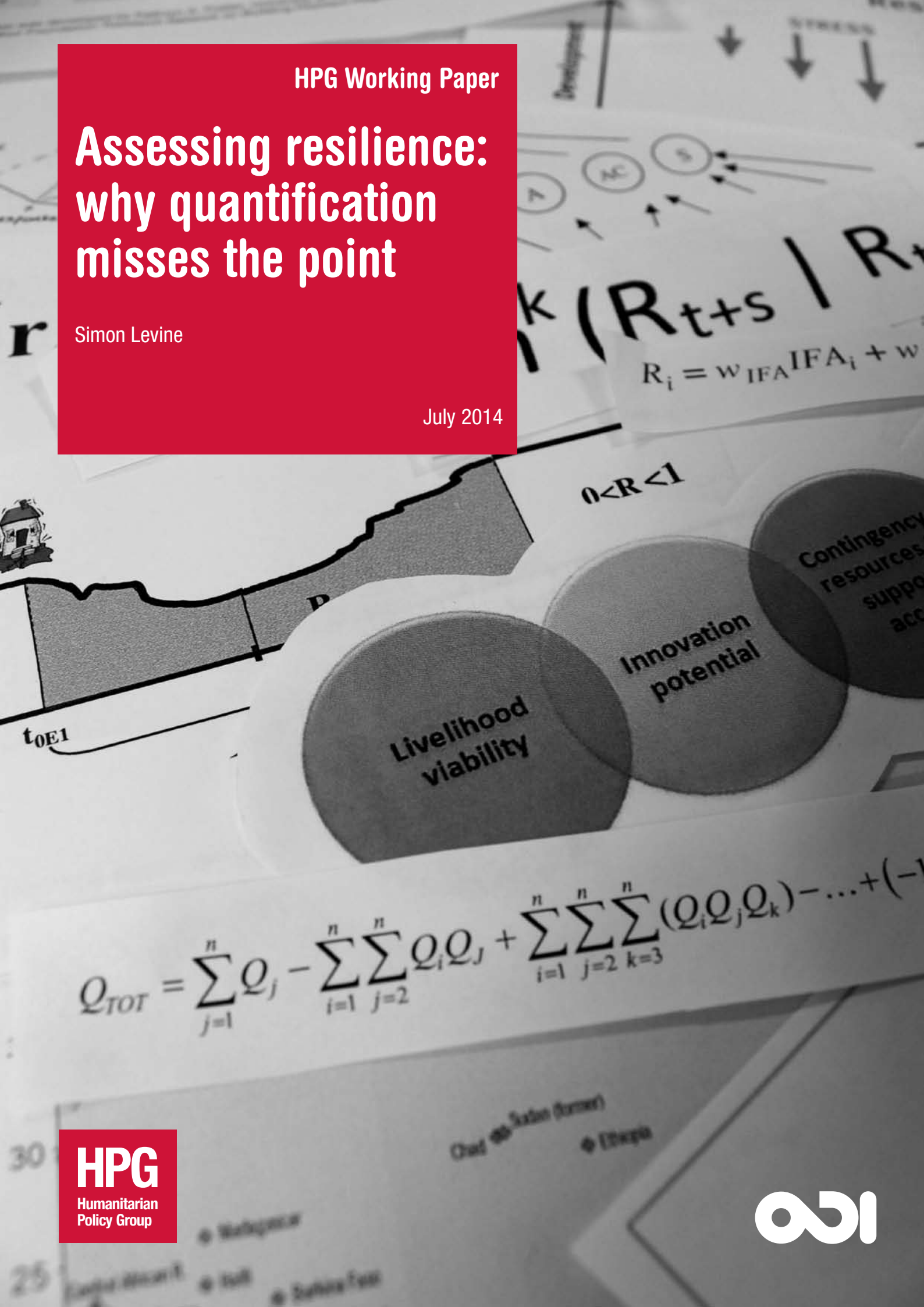
HPG Working Paper

Assessing resilience: why quantification misses the point

Simon Levine

July 2014



HPG
Humanitarian Policy Group

ODI

About the author

Simon Levine is a Research Fellow with the Humanitarian Policy Group (HPG) at the Overseas Development Institute (ODI)

Acknowledgements

The ideas in this paper owe a debt to more people than can be mentioned – thoughts and analysis written in published papers, given in presentations and shared in discussions. Particular thanks to those who gave helpful comments on an earlier draft of this paper, Tim Waites, Anita Shah, Joel Hafvenstein, Simone Di Vicenz, Irina Mosel and Luca Alinovi (who was most constructive in helping to develop a paper that he disagreed with profoundly). As always, thanks to Matthew Foley for helping make the paper more readable.

Humanitarian Policy Group
Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
United Kingdom

Tel. +44 (0) 20 7922 0300
Fax. +44 (0) 20 7922 0399
E-mail: hpgadmin@odi.org.uk
Website: http://www.odi.org.uk/hpg

ISBN: 978 1 909464 83 4

Contents

1 Introduction

The increasing use in recent years of the word 'resilience' as a heading for thinking about sustainable development, food security, linking relief to development, adaptation to climate change and the need to give greater priority to addressing vulnerability has been much discussed. It has proved attractive for many partly because it provides another way of addressing these long-standing and important challenges, and because it appears to offer a way to bring different disciplines and perspectives under a single conceptual umbrella.[1] It may also have such wide appeal because it both points to a practical agenda, through a political critique of the development policies of governments and aid organisations (i.e. 'pay more attention to people's vulnerability to crises'), and opens up new academic space for thinking about old problems, drawing upon such diverse fields as ecology, complexity and system theory and econometrics.

This wide appeal contains within it the seeds of possible confusion and hesitation. What exactly is the relationship between such varied interests in resilience and how, if at all, should they be combined in practice? One of the difficulties in combining these different perspectives has lain in attempts to develop ways of measuring resilience, for use both as a diagnostic tool and as a way of assessing the impact of efforts to support resilience. This paper offers an analysis of these attempts, not through a technical critique of specific proposed tools but by focusing instead on what can be learned from these various endeavours about how resilience is being understood. It seeks to understand how the appeal of the same concept to so many disciplines has resulted in the *reification* of resilience, whereby resilience has come to be seen as a many-sided 'thing', rather than as a way of thinking about many different kinds of problems. This reification is shaping, not only efforts to quantify resilience, but also the way in which efforts are being designed to 'build' it.

The search for a way to measure resilience is being undertaken by academics, NGOs, UN organisations, donors and national and international inter-organisational panels.[2] As such it is a rare case where several professional perspectives are reflected in the range of approaches being developed, including econometric equations, participatory approaches and complex statistical treatment. There are models which calculate the hypothetical economic cost of a crisis and others for quantifying what people have that might help them cope with crisis. Much of the development and humanitarian world is watching on the sidelines, waiting to know which product will receive the authority of consensual approval. It seems that there is, however, broad agreement on one principle: if we can only find the right indicators and the right calculations, we will have gone a long way to working out how best to build – and then quantify – resilience.

Three important concerns lie behind the demand for better resilience metrics: the need to give more attention to vulnerability in development policy and aid; the need for development policy to think more about an uncertain future and people's own agency; and the need to transform the way in which the collection, analysis and use of evidence for decision-making, including quantified evidence, takes place. These three problems have been combined, and tools for measuring resilience are being asked to address them all at the same time. There are good reasons for seeking to quantify the scale of the problems to be addressed, and the outcomes and impacts of interventions and policy changes. Indeed, far too little investment is being made in quantifying the impacts of policies and interventions in development generally, including in areas discussed within the resilience agenda. However, this paper argues that the desire for a *single* tool to address all of these tasks is unlikely to be satisfied because resilience is not a single 'thing'. Helping people and systems (health services, markets) to be more resilient is not a single class of activities: since there may be nothing at all in common between different challenges and

1 The umbrella metaphor has been used by Klein et al. (2003), Masten and Obradovic (2006), Strunz (2011) and many others.

2 For example the universities of Florence, Cornell and Tulane; Mercy Corps, Oxfam and others; FAO, UNICEF, UNISDR, WFP and UNDP; various departments within DFID; the Resilience Measurement Technical Working Group under the Food Security Information Network (established by FAO, WFP and IFPRI); and the Interagency Resilience Learning Group in the UK.

different solutions, progress towards them cannot be measured with a common tool.

This paper critically examines current approaches to developing a measure for resilience. It then proposes a different way forward by showing how the demand for assessing resilience and vulnerability can be met in a variety of ways – none perfect, but none claiming to provide a final definition of resilience or an ultimate measure of what actually matters to the lives of vulnerable people. Vulnerability analysis and impact monitoring can be improved enormously by using the tools that have already been developed within many different sectors where there is broad agreement about what the sectors mean and what needs to be done. Progress will not be better made by developing new methodologies to address the problems mentioned above in a quantified way, focusing on a vague and amorphous concept like resilience. Improving the way resilience is measured should mean changing the institutional emphasis placed on evidence collection and analysis and our understanding of how people cope with difficulties, uncertainty and constraints to their agency. Some of this work is technical, but much of it is organisational, making sure that the best existing tools are actually applied and their findings put to proper use. The attempt to find the perfect resilience index is not so much a difficult quest as a search for a holy grail – an impossible, or even an imaginary, task. It is also a task which is distracting attention from the important changes that are urgently required.

Many definitions of resilience have been proposed. Most seem a little more technically specialised than the everyday meaning of the word. Little disagreement has arisen, mainly because they are seen as a more scientific and academic formulation of what we know we all mean by resilience. This paper is loath to enter a debate about definitions, or to propose a new one for an English word with a written heritage going back nearly 500 years. However, discussions on quantifying resilience rely on, and then set in stone, a very specific and often implicit understanding of what is believed to be important in the definition.

An imprecise word is perfectly acceptable for everyday use, but in order to take measurements it is necessary to establish precise borders to the meaning of what is to be measured. Effectively, a new definition is needed, with more precise edges. This becomes problematic if the discussion is then presented as being about what resilience 'really is'. This risks moving to a new working definition of resilience that is created, not by theoretical analysis or from evidence, but because most people in the sector have passively accepted the parameters which are to be measured. Such a decision may be the most important factor in determining aid spending in this area for the next decade. This makes what are sometimes called 'sector metrics' enormously important. These metrics should be based on our understanding of resilience, not the other way round. Resilience is

Box 1: Understanding resilience with hindsight: the Chicago heatwave of 1995

In July 1995, a heatwave killed 739 people in Chicago. Social scientists studied the geographical patterns of mortality within the city to see which factors made some neighbourhoods more resilient than others.[3] The results were unsurprising: death rates were highest where air-conditioning was absent, poverty was highest and where violent crime was most prevalent. Eight out of the ten areas with the highest death rates followed this description. A closer look, though, revealed a more subtle story. Three out of the ten areas with the *lowest* mortality from the heatwave also had this same profile. The standard, generic models of resilience could not explain why one of the poorest and most violent neighbourhoods had death rates lower than some of the most affluent parts of the city. Comparing two neighbouring areas with similar vulnerability profiles revealed a more sophisticated picture. Mortality was high where a declining population had brought the closure of many shops and businesses; in contrast, where people still had reason to hang out on the street, a strong fabric of social relations had prevented people from dying alone. This was more important in saving lives than having air-conditioning. No one had thought to include these kinds of parameters in their profiling. If future interventions were guided by the crude story told by generic indicators of resilience, future interventions to build resilience to heatwaves would be less effective – and probably much more expensive – than a programme based on understanding these local stories.

3 This summary of the Chicago heatwave is drawn from Eric Klinenberg, 'Adaptation: How Can Cities Be Climate-proofed?', *The New Yorker*, 7 January 2013.

about people's ability to cope with life's difficulties. It should not be equated with the efforts that people make to ensure they are resilient, or with the efforts that others make on their behalf. Resilience is not, as often claimed,[4] a process, though *supporting* people's resilience may be a process. Abilities usually change over time, but this does not make them processes.

Resilience can be used not only about people, but also about many other entities, from health services and banking systems to football teams. This paper restricts the discussion to the resilience of people, whilst acknowledging that the resilience of other 'systems' is of functional importance, because they can help people. The paper does not speak of resilient communities, districts or nations, though many argue for a multilevel understanding of the word.[5] Although it may be correct to call a community resilient, it is not clear exactly what this means or how it relates to the lives and the resilience of the community's members. (A community capable of maintaining its status quo may be good or bad news for some of its members.) Applying resilience to different levels not only complicates (how do you add them together?) but also confuses. It is unclear if resilient communities are conducive to resilient people or vice versa, but either way this should be based on a great deal of empirical evidence which has not yet been gathered.

One of the sources of greatest difficulty for measuring resilience has been that resilience is not a single ability. People are resilient to different degrees depending on the threat or risk being discussed, and in coping with any problem people may rely on many different abilities. None of this is a problem in normal discussions. This paper contends that arguments about the definition of resilience or trying to identify in advance the individual abilities that people may rely on is a poor starting point for thinking about ways of assessing resilience. The more obvious starting point is the different tasks that a measurement tool is being asked to fulfil. The rest of this paper looks at some of the weaknesses and dangers in current attempts to quantify resilience, and offers a way of meeting the policy, programming and administrative demands whilst avoiding at least some of the pitfalls.

4 See Manyena (2006) for a discussion of resilience as a process or an outcome.

5 USAID (www.usaid.gov/resilience) and DFID (2011) define resilience respectively in terms of 'people, households, communities, countries, and systems' and 'countries, communities and households'. Most agencies speak of resilience-building objectives in relation to communities.

2 Current approaches to measuring resilience

This section identifies five different approaches behind current work to develop a measure for resilience. It is concerned with the overall approaches, rather than the specific examples presented. It does not, therefore, offer an assessment of specific strengths and weaknesses of the individual tools from a technical or methodological perspective, except where these illustrate important aspects that are inherent in the approach itself. Indeed, one of the problems of many of the papers on measuring resilience is that they spend far more time justifying the technical details, such as the exact statistical tests used, than they do on relating their formulae to real-life situations. Measuring resilience is one of the few times when the devil is not in the detail, but in the attention to detail.

2.1 Quantification based on functionality

The resilience of technical systems, for example the safety of back-up computer systems for commercial aircraft or the ability of a bank to survive a market collapse, is relatively easy to assess, and has been practiced relatively successfully for a long time. Such models quantify resilience successfully and usefully because:

- they measure a functionality which has an unambiguous and uncontested definition;
- functionality can reasonably be reduced to a few clearly defined variables;
- they study systems which are clearly defined and managed very deliberately to optimise that clearly defined functionality;
- performance is assessed in relation to known and clearly defined hazards; and
- there is a clear and critically important purpose to quantifying the resilience of these systems.

Similar approaches have been successfully used in the development sphere. The functionality of an infrastructure project may be a purely engineering consideration, but it can also be assessed for sustainability and poverty reduction impact, for example by using a (commercial) software-based tool such as ASPIRE.[6] ASPIRE gives a visual score on 20 parameters covering institutional, social, economic and environmental perspectives which have a great deal in common with resilience. The quantification does not attempt to aggregate the scores across the different parameters: the tool is designed instead to show where a project scores strongly or weakly, and allows project managers to monitor improvements through the various phases of design and implementation.

Models have been developed to measure resilience in more complicated systems, for example MCEER at the University of Buffalo, which assesses seismic resilience. Cimellaro et al. (2008) use the model to analyse the resilience of a hospital system (rather than a hospital building) to an earthquake. They reduce performance to a finite number of parameters (including the structural failure of hospital buildings, equipment failure, loss of electricity or water and the time taken to move between hospitals in the area) and a set of mathematical equations linking them. This is a fairly specialised application of the model, but it becomes important when such a model is then applied to the much wider concept of 'community resilience'. Because this is now happening (see below), it is important to go back and examine the conditions which need to apply for such an analysis to be meaningful and useful:

- there have to be reasonable models for quantifying damage to the different parts of the system;
- shocks have to be defined and quantified;
- the costs of repair must be calculated;
- there must be good models for predicting recovery times; and
- there must be good ways of estimating the reactions and resourcefulness of the system's managers.

6 See www.oasys-software.com/aspire.

There must also be a straightforward definition of what needs to be quantified that is clearly related to what is important in real life. In the case of a hospital system, the purpose of the system is for seeing sick or injured patients, and hospital waiting times are a reasonable proxy for performance (on the assumption that, when provided, the quality of care is a constant). Cimellaro et al.'s model is useful because they are analysing a system which is managed, meaning that there is an overall organisation governing how individual decisions are taken, and the behaviour of individuals (whether working within the system or as service users) is highly predictable. However, in real life even hospital functioning is much more complicated than the equations in Cimellaro et al.'s formulae suggest. The tragedy of Stafford Hospital in the UK is a good illustration of what happens when certain, ostensibly reasonable, indicators of 'performance' are treated as fixed criteria for assessing performance and when system management is then influenced by that assessment. Up to 1,000 people are estimated to have died between 2005 and 2008 because the hospital was managed specifically in order to meet certain service parameters, as a result of which basic standards of patient care were ignored.

Although there is no reason to believe that the conditions listed above apply in quantifying 'community resilience', attempts have been made to do this with the development of the PEOPLES methodology (Renschler et al., 2010).[7] PEOPLES is based on the theoretical idea of 'community functionality', which is a single value representing 'an aggregation of all functionalities related to different facilities, lifelines, etc.'. As with MCEER, once this step is taken resilience is then measured by an integration of the functionality function over time, i.e. functionality over time is plotted on a graph, and the area under the line is calculated to measure how much 'functionality' is lost in total over a given period, for example until recovery is complete. Community functionality is deemed to comprise the seven component dimensions from which the acronym is derived. These are further broken down into elements; the environmental/ecosystem dimension, for instance, comprises water quality/quantity, air quality, soil quality, biodiversity, biomass (vegetation) and other natural resources. A lot of detailed work has been invested in choosing the various elements and devising indicators by which they can be measured, and how community resilience at different geographical levels can then be combined. However, this technical attention has ignored the more important questions. What is the justification for choosing these particular elements? What reasons are there for believing that there is a fixed set of elements of resilience in different contexts? If dimensions which are hard to pin down, such as social capital, are important, what justification is given for believing that, just because it would be useful if they could be measured, there is any validity in indicators that purport to do this?

These functional approaches to measuring resilience are created to be precise and give clarity and transparency to arguments for different policy choices. The appeal of reductionist mathematical analysis in giving a semblance of objectivity and of certainty is paradoxically greater the more the object of study lacks certainty, clear definitions or predictability, and therefore the more the certainty and precision are actually illusionary. Even when used for simple managed systems such approaches have limitations. They tend not to use thresholds to define the necessary level of functionality, but consider any loss of function of a critical service or in a critical parameter to be equally to be avoided. (See section 3.1, below, for further discussion of thresholds.) By using mathematical functions such as integration to quantify loss of function, they are unable to distinguish between the ability to resist loss and the ability to recover quickly, since the two dimensions (x and y on a graph of functionality over time) are simply multiplied in calculating the area of 'loss of functionality'. Cimellaro et al. take the perspective of the hospital system and not that of the user. In real life, theoretical functionality and actual performance are often very different, and such measures are only ever as good as their definitions of functionality: critically, they also depend upon political choices which are masked by the technocratic perspective, including decisions about whose perspective (and whose resilience) counts.

2.2 Quantification based on indicators and characteristics

The approaches to measuring resilience which are gaining the most ground within the aid world rely on establishing indicators for sets of characteristics which

7 PEOPLES is the acronym for Population and demographics, Environmental/ecosystem, Organised governmental services, Physical infrastructure, Lifestyle and community competence, Economic development and Social-cultural capital.

Box 2: Definitions and measurement: how hard do you work?

It is very difficult to measure something unless we know exactly what it is that has to be measured. This makes measurement almost impossible unless we can define clearly what is being measuring. Most normal words do not have neat definitions that would make this easy. For example, how can we measure how hard someone works? Managers usually have a pretty good sense of how hard their staff are working, but quantifying this is notoriously difficult. It is even more difficult to compare two people who are doing very different jobs. Physicists can measure work precisely because they give it a simple definition (force x displacement), but this definition has nothing at all to do with what the rest of us mean by work. No confusion is generated because the meanings of the words in normal life and in physics are kept apart – and physicists do not claim that they have defined the 'real' meaning of *our* word.

The parallels with 'resilience' are striking. It is often possible to know (roughly) how resilient people are in different situations – but we struggle if asked to quantify it precisely, and even more so if asked to quantify the difference in resilience between two people living in very different situations. The problem may appear to be solved if 'resilience scientists' gave resilience a simple, mathematical definition, but then their use of the word would no longer be the same as the 'resilience' that we have been using for decades – and in a thousand different situations and in a hundred different ways those differences will matter. Resilience scientists would be able to create tools for making accurate and precise measurements, and for comparing resilience across time and space – but we would not actually know any more what it is that we are measuring.

are purported to determine resilience. This builds on longstanding approaches to developing 'indices' to measure wide-ranging concepts such as poverty (the Human Poverty Index), development (e.g. the Human Development Index (HDI)), hunger (e.g. the Household Food Insecurity Access Scale), 'readiness for change' (KPMG and Oxford Economics, 2013) and water poverty (e.g. Sullivan, 2002). Examples of such approaches for resilience include those developed by the Food and Agriculture Organisation (FAO) and the University of Florence (e.g. Alinovi et al., 2009; Alinovi et al., 2010; Ciani and Romano, 2013), Oxfam (Hughes, 2013) and the Livelihood Vulnerability Index (e.g. Hahn et al., 2009).

The technical differences between the various approaches relate largely to the choice of indicators or characteristics and the ways in which they are weighted or combined. All the models use judgment, rather than empirical evidence or analysis of vulnerability in any given situation, to choose the generic characteristics of resilience that they use.[8] Currently, judgement is the only way to derive such indicators since longitudinal evidence for determining those characteristics does not exist – nor is there consensus on what such evidence would be.

If resilience is deemed to be composed of the same building blocks in all situations, and if people's resilience is deemed to be augmented by an improvement in the score of any of the building blocks, this effectively licences the use of the same generic 'resilience-building activities' in all circumstances, regardless of people's actual situation or vulnerabilities. Even if locally appropriate indicators of asset ownership were chosen, such as ownership of cattle, land or bank accounts, the same essential problems remain: resilience is seen as modular, with any 'module' (e.g. assets) able to substitute for deficiencies in any other (e.g. exploitative local elites), and the solutions to a lack of resilience are thus disconnected from the actual causes of vulnerability faced by any group of people.

There are of course situations where different people rely on different strategies in the face of vulnerability, or different kinds of interventions can each help people to cope with crisis. What is required

8 The methodologies of Venton and Fitzgibbon (2013) and Tulane University (2012) need to be distinguished from the characteristic-based models discussed in this section. They both use participatory approaches, i.e. the judgement of affected people, to establish the dimensions of resilience to be assessed. Such approaches tend to produce context-specific measures which cannot be used for making cross-situational comparisons and which are thus different from the other index-based approaches discussed here. Indeed, the Tulane approach is hardly a quantification of resilience at all, since its utility lies in facilitating an analysis of the different dimensions separately, rather than combining them into a single measure.

is a way of assessing the usefulness or success of these different strategies: a modular approach to quantification does not allow this question even to be posed, much less answered, since the ability to cope is simply defined as the sum of the scores of all the different things that people try to use. The model also assumes that improvements in any component of a resilience score are of equal importance to people's lives. This is clearly untenable. Indeed, improvements in some dimensions may convey no benefit at all if they do not reach a certain threshold. Although it is better to be poor than to be very poor, it is less plausible to think of resilience in this way. If a person is unable to cope, having a half measure of protection will sometimes be as useless as having none at all. People living in a flood plain with only a half-completed dam for protection are not made more resilient if emergency work makes the dam stretch three-quarters of the way across the flood path. Modular approaches to constructing resilience scores are like approaches to flood protection which just count the number of bricks in a dam and ignore everything else about it.

The prospect of these approaches becoming more influential in setting policy, guiding programming and determining resource allocation raises other worries.

- Vulnerability is created by a combination of individual, household, local and national dimensions, which includes the role of power inequality and politics. It is almost impossible to think about combining an analysis of power imbalance with household-level asset holdings in a single measure. So far, national and political issues have been reduced to household-level access to services.
- Modular approaches also do not consider how people's lives are shaped by interactions between different dimensions or components of resilience. Institutions are needed to make assets 'come alive' (Levine et al., 2011) because they determine the benefit that any individual can derive from their assets.[9] In the same way, the value of education depends upon job opportunities. Taken alone, it is impossible to understand the value that any measured variable really has for people's lives.
- Well-known work showing how assets can in some circumstances become liabilities (e.g. Lautze and Raven-Roberts, 2006; Young et al., 2009) has long since refuted any possibility of interpreting indicators unambiguously. Commonly used indicators of social capital which measure the number of organisations of which people are a member are meaningless without consideration of the nature of those organisations, the relationships that exist between people within them, the costs of belonging to them and many other questions. (The fact that external actors often only consider formal organisations that look like the ones they are used to is a separate problem.)

It is important to stress that this is *not* a critique of the usefulness of quantitative research per se in contributing to an understanding of resilience: it is a critique of the use to which the results are being put. Many different kinds of research will be necessary to throw up new ways of thinking about resilience and vulnerability and to help build up evidence over time to answer many specific questions. Such research is not the same as 'measuring resilience'. There is a danger that the generic application of these modular methodologies for assessing people's resilience comes to be used as a basis for setting policy, disbursing funds, designing interventions, for monitoring and evaluation or for assessing value for money.

2.3 Quantification based on food access

If the ability of households to cope with a livelihood shock or stress can be reduced to their ability to meet their basic needs within a given set of access conditions (e.g. prices), then measurement becomes plausible and useful. The household economy approach (HEA) has been used for over 20 years as a tool for assessing this predictively for different population groups in different possible scenarios.[10] Livelihood profiles, created by quantifying all the income, expenditure and food sources of households, together with detailed quantification of their coping, have thrown important empirical light on

9 The example given in Levine et al. (2011) was a local irrigation system where status and wealth were needed in order to claim water rights effectively. It was the institutions that governed the system that determined how much of an asset the irrigation system was for different people.

10 See Holzmann et al. (2008) for an overview of the methodology. Applications of the approach can be found at http://www.heawebsite.org, http://www.hea-sahel.org and http://www.dppc.gov.et/Livelihoods/RegionsLZ.htm.

Box 3: Interpreting the evidence: data do not tell a story on their own

Although many of the proposed characteristics of resilience have intuitive plausibility, a closer examination dispels confidence in accepting any of them uncritically in all situations. One example can illustrate the dangers of analysis-free measurement. Many projects and papers associate a diversity of income sources with greater livelihood resilience,[11] and this is repeated in several quantification models (e.g. Alinovi et al., 2009 and 2010; Hughes, 2013). However, there is a wealth of evidence that a high degree of livelihood diversification is often an indicator of vulnerability or food *in*security – i.e. a reaction to livelihood stress. Simply counting income sources may lead to several fundamental errors:

- Though intuitive, the link between having many options and being resilient is often simply not true. There is a wealth of household data showing that the poor have more income sources than the better off and that these multiply when they have to cope with extra difficulties. There is a cost to spreading risks.
- Where different income sources are vulnerable to the same threat, no extra resilience is gained by diversifying. An indicator about diversification would have to include an analysis of the 'covariate risk',[12] but this poses a huge challenge for both data collection and analysis.
- Resilience comes from having *potential* income sources which can be relied upon if another fails, and not necessarily from *actual* income sources (i.e. in the absence of a problem). But how are potential sources to be counted?
- Distinguishing one 'income source' from another may be simple in theory, but is more complicated when the details of people's lives are examined. Is crop farming a single activity? If not, should it be broken down crop by crop? Should farming on rented land be distinguished from farming on one's own land or on common land if there are different vulnerabilities?

11 For example CARE's 'Toolkit for Integrating Climate Change Adaptation into Development Projects' (v 1.0), Frankenberger et al. (2012) and many others, though Frankenberger et al. also note that such diversification sometimes comes at the cost of remaining poor.

12 That is, the extent to which the different options are likely to suffer at the same time if one problem occurs.

vulnerability and resilience in different places, even showing how policies designed to reduce the risk of people facing crisis (i.e. to make people more resilient) have in fact made them more vulnerable (e.g. FEWS, 2008; Levine et al., 2010).

HEA is one mechanism for looking at resilience in a quantified way. Its practical usefulness is precisely because its scope is restricted, examining only the household economy and limiting its modelling to economic possibilities. It does not attempt to predict institutional or political changes. It is not, and does not claim to be, a quantification of resilience. Its strengths are an important commentary on what is needed from any tool. It offers a conceptual framework that does not attempt to go beyond what it can coherently analyse; a research methodology for collecting the data it will use; and an analytical model for using the data that is clearly linked to the research methodology. It provides clear quantified conclusions that are theoretically coherent, and which lead to clear actionable and quantified recommendations for action.

This does not make HEA a candidate for a tool to quantify resilience, even if the resilience agenda in some parts of the world (principally the Sahel and Horn of Africa) is being driven by the concerns that are the focus of HEA – people's livelihood security and their ability to access food and basic needs in the face of future risks (weather, global prices, demographic pressure). Its limited scope misses much that is important to people's resilience (e.g. agency, adaptive capacity, governance, institutions, protection, people's non-economic goals, gender relations). However, there are many situations where HEA is a useful element within an approach to problem analysis, quantification and impact assessment around issues central to resilience. It would be ironic if one of the only practical tools that offers quantified analysis were undermined by the desire for a single tool that reduces all aspects of resilience to a single measure.

There are interesting parallels and differences with the model of Barret and Constas (2013). They also see resilience (or 'development resilience') as fundamentally about the ability of people to maintain a standard of well-being in the face of difficulties, and they are interested in the different thresholds that can be applied to humanitarian (survival) and development (poverty) challenges. However, they

go beyond food security in their conception of well-being. As a result, they are theoretically closer to what we are interested in in thinking about resilience, but this is achieved at a cost of vagueness in what well-being actually consists of, or how a single number could represent it. They leave this question open, suggesting, for example, that Sen's capabilities could be used for the different components of well-being. However, though the model is useful for illustrating the idea that well-being and resilience can be caught up in poverty traps, it is based on a theoretical construct for which there is no evidence. The model is also likely to be challenged by those who approach livelihoods from a more political economy perspective, rather than mathematical models. Statistical models often, as in this case, assume that change is exogenous, i.e. determined by factors outside the system itself. Many analysts (most notably Lautze and Raven-Roberts (2006) and Young et al. (2005)) have argued that, on the contrary, the lines between external contexts and the internal dynamics of a livelihood strategy are at best blurred, if it makes sense to distinguish them at all.

Models based on food access or well-being have to cope with the challenge that resilience is not a measure of current well-being, but an ability to maintain well-being, and this cannot simply be assessed from current well-being. Any model therefore relies on prediction. Food security models have to assume that the choices people make can be predicted for given future scenarios. Although this is acceptable for the purposes for which HEA was created – a very short-term calculation of whether or not people will cope without help in a predictable crisis – it loses credibility the longer the future horizon and the more the focus of attention is opened up from the specific activities and coping strategies that households already have.

2.4 Quantification based on activities

The case for investing in resilience is being made partly on the grounds of value for money, and this necessitates some form of quantification, which must include putting a monetary value on improvements in resilience levels. This work has to date not involved measuring people's absolute level of resilience, but rather the changes in their level of resilience resulting from particular interventions. Through a mixture of evidence and modelling of the known or surmised impacts of different intervention types, a gross value has been put on the aggregated change (Venton et al., 2012) without disaggregated measurement at individual or household level.

These kinds of calculation have been used to illustrate the benefits of supporting people's ability to cope, rather than waiting until they need life-saving aid – or at least, they have illustrated the financial savings that can be made by spending on one rather than the other. However, this kind of methodology will probably be restricted to two uses: making the highly generic and political case for targeting development policy and aid at those most at risk of falling into crisis, for which headline figures are useful; and making a quick triage of possible interventions to get a first sense of their possible relative economic value. This approach cannot be used for monitoring or evaluation since it assumes a level of effectiveness for an intervention type. More importantly, it cannot be used for decision-making on policy or intervention design because it is not based on an assessment of how resilient or vulnerable different people are, and because it does not attempt to understand why people are resilient or vulnerable. Much of this is related to a third limitation: the use of intervention archetypes as the unit of analysis rests upon an analysis of impacts that is purely technical-economic. Issues of power, politics and institutions cannot be included in a study of intervention archetypes, though these determine the actual impacts in practice in different situations. This limits, but does not necessarily negate, the value of the analysis. Deciding what to do, for whom, where and how must then ultimately be determined by other analysis, and if there are demands for this to be quantified, then other ways of measuring resilience would have to be found.

2.5 Quantification derived from theoretical resilience frameworks

The main purposes of an analytical framework are to indicate the information which is to be collected in order to analyse the subject, and then to indicate how the information is to be put together in analysis. One might expect that the conceptual frameworks being developed to guide and inform thinking on resilience by DFID, Practical Action, USAID/TANGO

and Oxfam, for example, would also guide thinking about how to *quantify* resilience. Yet the main efforts at measuring resilience make no reference to, or use of, these frameworks, even where the organisation which has developed a framework is then engaged in developing a methodology for measurement. Would attempts to measure resilience be improved if they were more deeply rooted in current theoretical thinking on resilience?

There are good reasons for thinking not. Apart from the question whether resilience is really something that can be counted at all, there are too many critical questions to any analysis of how well people will be able to cope with difficulties which are either ignored or masked from view by current conceptual frameworks. This does not imply the need for a new, better, all-inclusive resilience framework: no model can ever meet all our analytical needs. Developing a range of smaller conceptual models to help in thinking about specific issues related to resilience will probably be a more useful investment of time. Current models may have great utility for thinking about certain issues within resilience, but until other difficult questions are adequately dealt with by our models they cannot be thought of as a theoretical basis for a broad measure of resilience. These questions are the subject of the next chapter.

3 Conceptual and methodological challenges in assessing resilience

The limitations with current models for measuring resilience have been ascribed to a lack of data, or they are considered still an unfinished work in progress. This paper argues rather that there are conceptual and methodological challenges in quantifying resilience, and that no methodology has yet given reason to believe that future methodological developments will be able to address them.

3.1 Conceptual challenges

3.1.1 Analysing resilience without thresholds

It should not be possible to think about resilience without considering thresholds. Whichever definition of resilience is used, most people who are concerned about the lives of people affected by crises (or trapped in chronic vulnerability) are worried less by how much people may lose in the event of difficulties than by how much they will be left with. Resilience is in some way about coping: yet, as we have seen, most approaches to resilience ignore this fact, defining it as a function of how far people fall and how fast they recover, but overlooking where they fall to (see for instance DFID, 2011; Bene, 2013; PEOPLES). Neither the speed nor scale of people's recovery is evidence that they had coped throughout the crisis, nor can it be used to measure how badly they suffered during it.

Others have taken the opposite approach, whereby people are placed above or below a defined resilience threshold (e.g. Venton and Fitzgibbon, 2013). In the jargon of quantification, this makes 'resilience' a dummy variable; that is, a variable that can only be either yes or no. It makes it impossible to say how resilient people are because, by definition, they either are or are not resilient. Resilience can then only be measured in terms of probability:[13] the various measured characteristics cannot be spoken of as *constituents* of resilience, but only *predictors* of its likelihood. Such a yes/no definition of resilience does not conform to normal usage, since it certainly makes sense to talk of people being more or less resilient. So, on the one hand we need to use a threshold to talk about resilience, but equally we cannot define resilience exclusively by this threshold. Constas et al. (2014) makes resilience a 'normatively indexed capacity', meaning that it is measured by reference to the threshold of what is acceptable. This chimes well with what is important to us. A parallel is offered by the construct of a poverty line, which even if slightly arbitrary can be useful even though people gradually become more or less poor.

However, the parallel also exposes a difference: being more or less resilient is not simply about being above or below a threshold of well-being – it is also about one's ability to remain there. In other words, poverty thresholds are always set at a certain value of poverty, whereas on these models resilience thresholds are not being set by a level of resilience but by a certain level of current well-being. The fact that there is a degree of vagueness in how thresholds relate to the concept of resilience is not the problem. It is attempts to do away with this vagueness for the purposes of quantification that risk misleading.

3.1.2 Vulnerability traps

People consider different thresholds when determining their life strategies. For some, ensuring survival or security is a goal in itself, and they may have to forego the possibility of a better life in order to ensure this (Wood, 2003). They may need a significant improvement in their situation (e.g. in opportunities, assets, political voice, social organisation, education) before they dare even to attempt to gain a level of resilience where they feel secure as autonomous human beings. Where they survive through dependency, making this attempted jump to resilience may be made impossible by the very strategies required for survival. (Daring to stand up for yourself may result in you losing the support you need

13 Using probit/logit regressions.

to survive.) This has some parallels with the notion of a poverty trap, though the barriers are not only economic, but are also created by political, psychological and cultural forces which are hard to measure – including fear, manipulation and peer pressure. Resilience cannot be adequately understood, let alone measured, without incorporating these dynamics.

3.1.3 Are there minimum conditions for resilience?

Resilience is being promoted as an aid objective for the most difficult places (conflict, fragility, protracted or recurrent crises). Yet much of what has been written about resilience assumes quite ideal conditions as its prerequisite. This makes the analysis inapplicable for those looking for a practical way to help people cope in these difficult situations. This is not merely about the impossibility of achieving full resilience in difficult places: in Bahadur et al. (2010), the list of the most commonly identified requisites for resilience includes diverse economic opportunities, effective governance and a high degree of equity, which are so far from the realities of difficult places that they cannot serve as a guide even for setting a direction for desired change. Are there situations where it is inappropriate even to think in terms of resilience? Most frameworks simply avoid this question. Frankenberger et al. (2012) is a rare exception in arguing that '[in] certain situations – such as those where formal government remains fragile or absent and/or those experiencing ongoing violent conflict … resilience building may be impossible unless and until basic minimum conditions are present'.

This conclusion sits uneasily with what could be called the 'political' case for resilience (e.g. as expressed in Ashdown, 2011), that the duty of development aid is to focus on the vulnerable in difficult places in order to make it less likely that they will fall into serious crisis. It can be argued that it is always possible to increase people's agency, to expand their range of choices and to improve, if only incrementally, their ability to cope (Levine and Mosel, 2014). If this is so, we need ways of approaching resilience that leave room for discussing both the minimum conditions for achieving an acceptable level of resilience and also how to improve resilience in places where life will remain below this acceptable level.

3.1.4 Economic and non-economic resilience

Resilience is frequently addressed mainly in relation to food or livelihood security, not because this is all that resilience means, but because that is the most urgent problem which external agencies feel they can address. However, even in the economic domain of food security, people's economic goals and livelihood strategies are usually influenced by non-economic objectives. Many people bear high economic costs in order to gain or maintain social status or acceptance (Carr, 2013). People vulnerable to crisis are particularly likely to see their economic and non-economic goals as inseparable, e.g. relying on social acceptance for their economic or physical survival (Jaspars and O'Callaghan (2010) found that, in conflicts, issues often considered under the label 'protection' can be indistinguishable from livelihood issues). Livelihood frameworks have struggled to capture this, and so far resilience frameworks have not fared any better. It presents a huge challenge for measurement for two reasons: it is difficult to know how to disentangle what needs to be measured; and even if that is achieved, the parameters identified are often very difficult to quantify in a meaningful way.

3.1.5 Geographic scope

There are two distinct challenges in setting geographic boundaries around the study of resilience. First, many people's livelihoods are trans-national, because of remittances, (seasonal) migration and cross-border trade. An assessment of the resilience of people's livelihoods in parts of Niger and Burkina Faso may need to pay as much attention to shocks which threaten economic opportunities in Ghana and Côte d'Ivoire as it does to shocks closer to home. Second, even people's home activities take place in an economy (and a political economy) that is shaped by international forces, whether regional or global. The impact of globalisation varies enormously in different situations. Following the global food price shock of 2008, food prices in many parts of Ethiopia rose several times more than prices in the world market because of poor market integration (Ulimwengu et al., 2009). In contrast, prices in Uganda rose by much less (Benson et al., 2008) because of a great reliance on staples (cassava, sweet potatoes, green bananas) that cannot easily be exported from a landlocked country. The treatment of such issues must be fitted into any consideration of resilience, but current approaches to measuring resilience do not leave room for analysing them.

3.1.6 The resilience of people, households or communities?

Although aid is interested in people (i.e. individuals), most current attempts to quantify resilience have done so at the household level, and much of the

discussion about resilience speaks of building resilient communities. The focus on households is problematic when considering resilience beyond immediate food access, because intra-household relations are both an enabling and a constraining factor for people's ability to advance their futures, both economic and non-economic. The link between the resilience of people, villages, towns or countries is not straightforward. 'Community resilience' can help or harm marginalised or poorer households and individuals. It cannot be assumed that resilient countries can be built by building resilient communities by building resilient households, nor, as Pfefferbaum et al. (2011) propose, that resilient communities are made up of resilient individuals *plus something extra* ('the ability to transform the environment through deliberate, collective action'). The opposite assumption, that people's resilience can be supported by supporting resilience from the community (or state) down, fares no better. Pain and Kantor (2012) used evidence from rural Afghanistan to offer very practical – but context-specific – help in identifying characteristics at community level that can help explain when support at community level is likely to benefit everyone, and when it is more likely to further entrench elites. Individual resilience is a very complex construct of community, household and individual characteristics, a challenge that has yet to be faced openly by the conceptual frameworks and measurement methodologies on offer.

3.1.7 Risk and resilience

The ambiguous nature of risk in relation to resilience poses several challenges to measuring resilience. There is a danger that risk will be interpreted as being about uncontrollable natural hazards rather than about structural (read: political) vulnerability (e.g. Cannon and Mueller-Mahn, 2010; Harris and Mitchell, 2012), especially because it is much easier to measure 'objective' events such as rainfall than it is to 'measure' the circumstances which deprive some people of access to irrigation. Second, accepting risk is part of the ability to invest in the future, one of the fundamental indicators of agency – and resilience. While measurements must not automatically interpret risk aversion as resilience, it is harder to know what to do about this. Different societies, economies and cultures vary in risk aversion, and there are different cultural attitudes towards the relative merits of 'protecting' people against the consequences of loss compared to encouraging entrepreneurship by rewarding success and risk-taking. Such questions are cultural, not simply 'scientific' or 'objective'. Third, exposure to hazards (e.g. by living close to the coast) may bring benefits which more than outweigh the dangers. Measurement of resilience or vulnerability must be underpinned by a good understanding of the rationale of the people whose lives are being studied.

3.1.8 Resilient like reeds or trees?

Many cultures use the contrasting images of the tree which withstands a storm and the grass or reeds which temporarily accommodate it. Both can be called resilient. There are trade-offs in choosing between these two pathways to resilience, and investing in one will close off opportunities to take the other. Assessing resilience must assess the robustness of the tree (i.e. how much people have been able to progress along their strategy of withstanding shocks) and consider whether, in the future, survival or resilience will be better guaranteed through withstanding or adapting to problems (mimicking the tree or the reeds). This is challenging, because it is rarely possible *without hindsight* to judge which pathway was the right one. Maladaptation undermines resilience, but if the future is uncertain it may not be possible to identify in advance which kinds of adaptation are maladaptation. Resilience frameworks and assessment methodologies must leave space for differences of opinion about the two overall strategies, which cannot happen if resilience is reduced to a single score (or to a spider diagram of five numbers) which does not explicitly distinguish between the two. Such approaches make it impossible to discuss whether assistance is making things better or worse.

3.1.9 Future-looking resilience?

There is a tendency to base the assessment of resilience on parameters which are believed to have contributed to resilience in the past, but the past is not necessarily a guide to the future. Urbanisation, for instance, brings new markets and employment opportunities and displaces old ones; it changes the value of different kinds of land; it may increase the threat of land-grabbing by urban elites; it changes the political balance that determines the distribution of government funds; it changes the size of government revenues; and so on. In the future, resilience may depend on different institutions, different skills and different social connections from those that are being measured based on analysing yesterday. An explicit explanation of the expected scenarios, and an explicit discussion of what parameters are believed to be important in the future, is needed for any assessment of resilience; without this, pronouncements cannot be tested, interpreted or criticised.

3.1.10 Adaptive capacity

People can cope better with an unknown future if they have the ability to deal with change – in the jargon, 'adaptive capacity'. This is recognised by many frameworks and analyses, but measuring it remains elusive. It is common to use parameters such as number of years of schooling, but a person's adaptive capacity cannot simply be a function of how long they sit in school. (This is a case of using the jargon of 'proxy indicator' to mask the fact that we prefer to measure something irrelevant rather than nothing at all.) There is no evidence that any approach to measuring adaptive capacity has any validity – hardly surprising, since it is a composite of psychological, cultural, technical, financial, social and political factors. Supporting adaptive capacity is perhaps the most important thing that can be done to help people: should we necessarily refrain from trying until we know how to measure success in doing so?

3.2 Methodological challenges

Once these conceptual challenges have been acknowledged, some familiar methodological difficulties need more explicit treatment in relation to quantitative research on resilience. There is little value in developing a theoretically perfect model for quantification that cannot be properly applied in real life. There are challenges both in operationalising some characteristics of resilience (e.g. social capital) and in collecting reliable data.[14]

3.2.1 Getting data that is reliable and meaningful

Some models for resilience assume away the challenges of operationalising complicated social dynamics. Assessing social capital has, for example, been reduced to counting the number of organisations that people belong to, and attempts to deal with adaptive capacity were discussed above. Social relations *can* be analysed, but the part they play in building or undermining resilience requires that they are not only counted but also characterised and understood.[15] It is much quicker and easier (in both research and in implementing projects) to focus on the formal dimensions of life, but this is unlikely to be very useful in societies where the informal is paramount. There are also well-known problems around applying quantitative models in situations where it is extremely difficult to obtain reliable data, for instance on income. There are advantages if we are able to treat data mathematically and statistically, but these advantages cannot counter the problem that the methodological demands for collecting data that can be treated statistically (most critically, the exclusion in interviewing of probing questions, and interpretation based on explanation[16]) mean that statistically treatable data is the least likely to mean what the researcher thinks it means. The 1960s description of this misplaced trust in (then newly computerised) statistical analysis, 'GIGO' (garbage in, garbage out), seems to have been forgotten. Discussions on methodological development in quantifying resilience have not taken as their starting point the question of what is practically possible.

3.2.2 Distinguishing the person from the place

There is a difference between how resilient or vulnerable a person (or household) is, the fragility of their livelihood strategies and the viability and supportiveness of the places where they live.[17] There is nonetheless a link between them. Definitions of resilience have often elided this difference. If the link between the resilience of people and how far their context supports or undermines their resilience is not clear, as can happen in quantification techniques which combine parameters relating to people and to places in a single measure, there can be potentially disastrous consequences for policy. An examination of the resilience of different livelihood strategies in Kenya (Alinovi et al., 2010) found that pastoralists were the least resilient of six identified livelihood groups. The policy implication that could be drawn from this is that pastoralists need to be helped out of their vulnerable livelihoods into one of those livelihoods found to be more resilient, such as wage employment, agro-pastoralism or smallholder crop farming. The study did not, though, support such a conclusion. The pastoralists scored low because of their exposure to regular shocks (unreliable rains) and because of poor access to basic services and the marginalisation of the arid lands. However, abandoning the very livelihood strategy that was developed to cope with erratic rains would not solve their marginalisation.

14 Operationalisation means giving a working definition, for the purposes only of the process of measurement. For example, we could measure how happy people are by counting how often they smile in a day or use a survey to see how many 'happy sentences' they say they agree with, though no one would argue that this is what being happy actually means.

15 One example of this is the work of Pain and Kantor (2011), which described seven kinds of relations between people in rural Afghanistan, ranging from reciprocal through patronage to exclusion.

16 When survey methodologies demand that everyone be asked exactly the same questions in exactly the same way (to remove bias), probing and follow-up questions are not allowed.

17 The implications of confusing them has such practical import that some (e.g. Wisner et al., 2004) have argued for the need to use 'vulnerability' only for people, and to use 'fragile' or 'hazardous' for livelihoods or places which are at risk.

4 Towards a way forward

There are signs that the resilience movement may have some success in bringing together thinking from different disciplines to address problems of poverty, vulnerability and risk. This is welcome and long overdue. However, arguing that situations need to be examined by combining many different perspectives is not the same as suggesting that those perspectives have been superseded by a new composite perspective. The perceived 'problem' of quantifying resilience both stems from and drives the tendency to create a distinct resilience sector. In philosophical jargon, we are making the mistake of 'reifying' resilience – that is, treating it not as an abstract idea, but as if it were something concrete. If there is no essence of resilience, no core set of characteristics that always determine it for all people at all times, then we are freed from the perceived need to create a new box in which to locate our efforts to address people's vulnerabilities. There is no longer a need to think of 'resilience-building' as a separate group of activities: whatever is done to help people as a result of thinking about their vulnerability can be called resilience-building, with no implication that the various things that are done have anything in common beyond this. If this is recognised, then the urge to create a new way to measure those efforts also evaporates. If we no longer have the illusion that there is a separate box of resilience activities we are freed from the need to create a new quantification tool for use in measuring resilience.

Giving up the search for a universal measure of resilience is not a negative call to give up on measurement or on more rigorous assessment. Instead, we must focus our efforts on ensuring that we get better at quantifying the things that really matter. What these are will depend on the reason why we want quantified understanding. Progress thus proceeds by considering and addressing the needs quantification is intended to meet, rather than, as has been the case so far, the thing to be quantified. Much of the concern to develop new methodologies for quantifying resilience would be dispelled if existing good practice around analysis, assessment and monitoring were used more frequently in addressing the following needs:

- establishing impact monitoring to inform the management of interventions and policy;
- learning, over time, which interventions and policies are most useful for which populations in which situations;
- choosing between investments in competing policies or interventions;
- understanding better the determinants of resilience to various threats in different situations;
- being accountable to those providing funds for investing in resilience;
- making a political or advocacy case for investment in resilience; and
- making comparative assessments of need in order to target resources.

4.1 Impact monitoring

Understanding the changes which interventions and policies bring to people's lives, and how these changes occurred, is probably the priority area where quantified understanding is needed. The need for rigorous and quantified impact assessment across the humanitarian and development spectrum is gradually, if belatedly, being recognised. Impact monitoring for 'resilience-building interventions' brings no particular methodological implications (because impact monitoring involves comparisons across time, and not across contexts, there is no reason to look for a generic or context-free methodology). As with any other intervention, a monitoring system starts with a problem and situational analysis, which determines which change is intended, why and how it is to be brought about. All interventions should be based on documented analysis which gives the rationale for addressing the intended change, explaining why that is important for improving people's ability to cope or removing a constraint to their development; and a programme theory or theory of change, explaining how the change is expected to play out for different people, and how the intervention will interact with the various economic, political and other forces that ultimately shape people's lives. This analysis will normally be based on analytical frameworks from various disciplines (e.g. livelihoods, nutrition, public health), but the use of multiple frameworks, including some model of resilience, may sometimes be useful.

Box 4: Quantified data in impact assessment: moving beyond the 'qual versus quant' dichotomy

Impact assessment methodologies use statistical treatment of quantified data to meet two separate objectives: substantive (to measure impact) and procedural (as a way of proving that correlations are not random). In the places where resilience is most at risk, data rarely performs this second task well because of the degree of variability inherent in people's lives over time, the importance of other external factors, the number of interacting variables and the difficulty in identifying and operationalising many of these parameters. As a result, it is hard to find statistically significant correlations, to interpret the lack of correlation and even to interpret unambiguously those correlations which do appear (i.e. even if we know what happened, we do not know why or how).

The procedural objective has sometimes undermined the substantive objective – to understand the size and prevalence of changes in different aspects of people's lives. Quantitative studies have become more and more the domain of specialists in the methodology rather than in the substantive questions, and as a result they are often not useful for those who most need to understand them. However, quantitative methodologies maintain their authority because they give the appearance of being comparable and transparent. In fact, this appearance rests on some implausible assumptions, including that data are reliable and of good quality and that a context-free comparison of data allows conclusions to be drawn. Quantitative methodologies establish a high standard for internal validity, but the absence of a situation-specific interpretation of the numbers makes external validity very dubious (i.e. we can't know if the data tells us anything at all about any other situations apart from the ones studied). It is not very useful to invest heavily in impact monitoring that establishes statistical correlations between outcomes and interventions, but does not permit anything more to be said than that.

Six different purposes can be identified for statistical, quantitative research in understanding impact. These relate to what the data are being asked to do (1 and 2 below); methodological implications (3 and 4); and practical considerations (5 and 6).

1. To quantify impact.
2. To prove causal links by correlations, using an explicit causal model.
3. To facilitate retesting of the data analysis.
4. To have comparable methodologies that enable standardisation and cross-study analysis.
5. To permit large sample sizes at a reasonable cost (because survey interviews are much faster than in-depth qualitative interviews).
6. Lower demands for skilled field researchers (use of unqualified enumerators).

As discussed, there are usually better alternatives for the first; achievement has been disappointing, at best, at the second; the third is rarely used; and the fourth is based on dubious demands to strip lessons from their context and to ignore problems of external validity. The remaining justifications, that it is easier and cheaper, are valid, but value for money rests upon being able to deliver a product of good value.

For a forward-looking characteristic like resilience, predictive assessment is particularly important. This requires probabilistic analysis, which appears to give extra justification for relying on statistical techniques. However, prediction entails drawing conclusions from the past (what was measured) and projecting them into another situation (the future). To make such projections, there has to be good reason to believe that two or more situations are alike in essential ways. However, this relies on establishing the external validity of the conclusions drawn from statistical analysis: if questions about context are regarded as being outside the domain of quantitative research, then a kind of 'certainty' is bought, not by tackling the question of whether or not it is reasonable to make a comparison between two situations or make projections into the future, but by ignoring it. If predictions are not then accurate, statistical enquiry does not admit any inherent failings, but instead argues that exogenous, 'unforeseeable' factors meant that the predictions did not apply. Qualitative enquiry, on the other hand, tends to focus more on the external validity of its assessment – though it may risk drawing unfounded conclusions because of insufficient attention to internal validity.

Box 4: *(continued)*

It would be useful if there were tools which separated the two distinct functions of assessing size and proving causal links. There is no substitute for quantification in order to know how much interventions have helped (or harmed) people. However, causal links can be investigated in ways other than by finding correlations. There are many implications of changing the way in which causation (i.e. ascription of change) is established. If it is done through more qualitative enquiry, there will be added costs because gathering information from interviews is more demanding than collecting data through simple survey work. Much smaller samples will be used, raising the risk that representativity will be compromised. Another drawback is that (rightly or wrongly) conclusions are often considered less convincing where they are not backed by statistical analysis.

However, if there is more widespread use of mixed-method approaches that allow quantification, but without relying solely on statistical analysis to establish why changes happen, this will generate better understanding of what is happening. The information is also likely to be much more reliable, because it is probed and tested. It is for researchers and managers in each situation to decide whether this advantage outweighs the disadvantages.

Analysing causal links directly, rather than inferring them from correlations, will demand a range of information-gathering processes, which look for less data (i.e. the value of a variable in abstraction from its context) and more quantified information (i.e. numbers which take their meaning from their context, within a coherent and comprehensible 'story'). More formal quantification may still play specific roles within assessments in mixed-methods approaches.

Any change in 'resilience' will then be assessed not in the abstract, but as a specific and practical removal of constraints or increase in opportunities. Monitoring should ideally assess not only changes in the end state (impact, outcomes), but also as much of the network of causes and effects as possible. This will help avoid an unduly narrow focus limited to the pre-identified intended change.

The one area where monitoring impact about resilience is more complicated than for some other areas is that there may be a need to go beyond what could be called the first level of outcomes (e.g. increased income, better protection of land rights, better access to health care). It may be necessary to understand whether and how these first-level changes contribute to progress towards the larger objective, namely people being better able to cope with change or crisis. This is made difficult when these objectives are articulated in an abstract way ('people will be more resilient'), rather than working out (together with the people concerned, one would hope) what *precisely* was intended – e.g. a particular group of people are less likely to fall into crisis in the event of specific shocks, people have greater ability to invest in their future or people (and services) are able to recover more quickly in the event of a certain kind of crisis. Being specific about the intended objectives will not always make them simple to measure, but it will remove the temptation to search for a generic (i.e. vague) way of measuring them.

4.2 Understanding the determinants of resilience

The study of people's resilience, how they face difficulties and why some people fare better than others cannot be based on a standardised tool based on a pre-determined formula for understanding resilience. Such research is intended to explore what parameters are associated with different outcomes in the face of different problems, or what characteristics are good predictors of resilience: it makes no sense therefore to base it on an assumption that we have already identified them. What makes people resilient is a whole field of enquiry, necessitating very many, specific and researchable questions in different situations. This plethora of research questions will need a diversity of approaches, in particular tools which follow people's lives over time, measuring their ups and downs from many different perspectives. Longitudinal quantitative studies have much to offer if

we are to gain a better understanding of how and why different people cope in different situations: such work though is about exploring resilience, not measuring it.

This is not a new area of study and much work has already been done. Often, though, individual studies have been left in isolation rather than brought into a collaborative effort at building and sharing understanding. This is an institutional and system problem; the solution lies in making better institutional efforts, not in developing standardised methodologies. Communications technology has made sharing perspectives and learning easier than ever. It is not difficult to create open learning hubs around specific countries, populations and problems where people from different perspectives and disciplines can contribute their insights and evidence and can learn from others to create a level of analysis and understanding that is greater than the sum of its parts. This would have huge value, whether or not a consensus ever emerges.

4.3 Learning which interventions and policies are most useful

Several major donor initiatives around resilience include a strong learning component under the responsibility of a dedicated learning partner.[18] Currently there is so little evidence and analysis about impact that the needs are as great in many long-established sectors as for newer preoccupations such as resilience. There is no simple solution, but the best way forward does seem clear and new methodological tools do not have any obvious role. In order to learn what works, we need to be constantly generating evidence (such as that provided by good monitoring of impact in many different situations), arguing about whether we are collecting the right evidence and creating spaces for that evidence to be interpreted in different ways. Resilience is not a new sector, but, as many have written before, it is an area of work where many different disciplines can contribute.

General lessons about resilience may emerge by seeing comparisons or contrasts from one country to another, but cross-situational learning cannot emerge from context-free analysis: it will slowly be gained if patterns can be discerned in the specific and context-grounded analyses of different situations. Quantitative work around these specific situations – not attempts to measure some reified idea of resilience – will be an important part of this learning.

4.4 Choosing between investments in competing policies or interventions

The design of interventions and decisions on resource allocation ought to be based on a consideration of alternatives, which should be judged across a range of criteria:

1. The size of the likely benefit.
2. The number of people likely to be assisted.
3. The relative need of those being assisted.
4. Cost-effectiveness.
5. How sustainable the benefits are likely to be, including support to adaptive capacity.
6. How positive or negative their likely secondary impacts will be.
7. The prioritisation given to the overall objective (by different stakeholders).
8. The probability that the interventions will have their intended benefit.
9. Operational considerations, including capacity and security.

The need for such comparative analysis is sometimes recognised for development interventions,[19] though usually is either absent from any documentation or only performed as an after-thought after a choice has already been made. In humanitarian action, the demand for considering alternatives has begun more recently with an increasing range of acknowledged instruments, and a literature on 'response analysis' is beginning to emerge (Darcy and Hofmann, 2003; Maxwell and Stobaugh, 2012; Maxwell et al., 2013; Marsland et al., 2011). For resilience specifically, a productive discussion should at a minimum encompass the following:

18 Examples include Tufts University, which is playing such a role in USAID's resilience programming in Somalia, Valid International, which is accompanying DFID's multi-year humanitarian funding to assess its role in building resilience, and HPG, which is supporting learning for the DFID-funded Building Resilient Communities in Somalia consortium.

19 For example, World Bank project appraisal documents contain a standard section entitled 'Alternatives considered and reasons for rejection'.

1. A good situational analysis, including analysis of possible future trends, incorporating several theoretical perspectives (economic, gender, vulnerability, adaptive capacity).
2. Clear prioritisation of what needs to change and for whom.
3. An explanation of how the current situation has been created and maintained (by political, social, economic and other forces).
4. Analysis of how the current situation can be changed and how stable any new outcome state will be in the face of those political, social, economic and other forces.
5. An understanding of what kinds of impacts are most likely from different kinds of intervention.
6. An appreciation of how the direct changes which any intervention brings (e.g. in assets, gender relations, freedom of movement) will affect people's range of options, their ability to exercise independent choice and their ability to cope with problems.

Some of these questions should be informed by quantified evidence and analysis – but using a context-specific analysis, which can be based on existing tools. Quantification does not provide a value-free escape from subjectivity, and it must remain the servant, not the master, of judgement, if only because it will always be necessary to balance trade-offs between the competing criteria listed above. Only the last question in the above list requires assessing resilience specifically. The tools and measurements needed for answering this question will be determined by the problems being analysed, such as engineering (e.g. likely damage from seismic events), economics (e.g. change in trade as a result of investment in infrastructure), food security, land rights and gender inequality. The right tools are available, but are not being used or demanded. Addressing this will require political or institutional change (better systems and processes). This is clearly a long-term task, and needs to start now.

4.5 Accountability to those providing funds

There is a growing demand for accountability in aid to move beyond financial accountability (i.e. the transparency with which money is spent) and to include accountability for impact. This accountability can only happen if there is good understanding of the scale and size of change brought to people's lives. However, inadequate attention in the past to impact accountability was not caused by a lack of tools for measurement, and a technical solution (measurement tools) will not on its own be an answer to the bureaucratic, institutional and systemic problems that have left us knowing so little about the impact of decades of interventions and policy changes – whether for 'resilience' or for anything else.

Changing this will not be easy, and the topic is beyond the scope of this paper. It should, though, be clear that accountability cannot be improved by the use of the current tools being advanced for measuring resilience – indeed, if agencies are allowed to report their performance using a resilience index, then any hope of accountability may be lost. Resilience scores can be improved quite simply by identifying the easiest component or dimension of change and focusing entirely on achieving this, regardless of whether or not this is relevant to unblocking the constraints that prevent people from being able to cope in times of difficulty. For example, any distribution of 'assets' can improve a resilience score, even if it is largely irrelevant to improving the quality of people's lives or their resilience. Indices which are based on hidden or arbitrary calculations undermine transparency: agencies ought to hold themselves to account for correctly identifying specific constraints in people's lives, for tackling them to the extent that they claim will be possible and as a result for making the changes to people's lives that they claimed would happen. This can and should be made to happen, for all aid interventions. A political commitment to doing so does not depend on any technical advances in analysing resilience and should be made immediately.

4.6 Making a political or advocacy case for investment in resilience

The case for investing in resilience, rather than waiting for a crisis to occur, needs to be continually argued for. In some quarters the case is best made through numbers, and in particular some kind of cost–benefit analysis, even if cost calculations do not capture well all the benefits brought by this kind of investment. The case is weakest when it relies on arbitrary or generic indicators to measure how big a change in 'resilience'

can be bought for each dollar, for two reasons. First, it is only a matter of time before the case for resilience is undermined by the evidence that vulnerability persists despite resilience-building efforts, and second, the case for 'increasing people's resilience index' is far less persuasive than the case for saving lives in the event of a flood, ensuring that fewer buildings fall down after an earthquake (saving money in rebuilding them) or making sure that fewer children fall into malnutrition as a result of a conflict.

The value for money argument needs to be based on a transparent and agreed understanding of value. Since it is usually cheaper to lift people out of poverty if they live near the capital, in high potential areas or near tarmac roads, simple cost-per-output arguments may tend to justify diverting resources away from those already marginalised – the exact opposite of the rationale behind the call for putting resilience at the heart of the development agenda. The value of any impact must be related to the situation of the people benefiting. If used in a context-free way, an artificial currency of generic resilience units may thus, ironically, undermine the value for money case for investing in resilience.

4.7 Making comparative assessments of need in order to target resources

A comparison of levels of 'resilience-building needs' may seem to be needed in order to allocate resources, but this perceived need to make 'resilience comparisons' comes from thinking of resilience-building as a separate sector, with its own budgets and its own separate set of objectives. There is, though, no justification for such a separatist view of resilience. The case for investing in resilience is an argument for targeting aid differently, and for incorporating an analysis of vulnerability and risk as key considerations in planning in all sectors. Support for economic growth, for example, should not be geared simply to increasing GDP, but must ensure, for example, that unemployed youth benefit from any improvement in economic activity. Support to such economic activity is still an economic programme – but one based on a resilience analysis. There is no reason why it will have anything in common with any other intervention designed using a resilience analysis.

The political case for resilience draws on dissatisfaction with the low degree to which development aid has been targeted at helping the marginalised and those vulnerable to acute crisis. However, many indicators are already used for assessing relative need in different areas – access to business services, food security, marginalisation, access to health care, quality of education – and there are no grounds for thinking that a resilience index is needed in order to improve the situation. As in the other areas discussed above, the need is not for new methodological tools, but for the better use of tools we already have. Indeed, current targeting of resources is so inequitable that it is perverse to demand more sophisticated technical methodologies to do a job that has been ignored for decades. Improving the way resilience is measured should mean changing the institutional emphasis placed on evidence collection and analysis and the attention to long-term understanding of how people cope with difficulties, uncertainty and constraints to their agency. The attempt to find the perfect resilience index is distracting attention from the important changes that are urgently required.

5 Conclusion

We can think of people's resilience as meaning many different things: what possibilities they have, how much freedom they feel they have to follow those possibilities, what risks they face, how well they feel they can cope with these risks, whom they can turn to for help and on what terms. It is impossible to say in advance how resilient people are until they are faced with a test, but even then there are so many dimensions of 'coping' (economic, social, psychological) that there will be differences of opinion on how well they succeeded. None of this presents a problem in normal conversation, but when we try to turn resilience into a scientific definition, trouble inevitably starts. The challenge of trying to measure resilience then appears to be a supreme methodological challenge, partly because two sets of problems are being posed simultaneously without differentiating between them: the normal technical debates about the adequacy of different indicators and how to measure them (with which the industry continues to struggle in many other fields); and the conceptual discussions which are trying to arbitrarily fix precise boundaries to the meaning of the word.

Discussions about the priority objectives of national development policies and aid should not be transferred to a technical discussion around the quantification of resilience. Three particular dangers of doing so should be mentioned here. First, when we try to measure what is important, we make important what it is that we measure. This should not be set without consensus on what needs to be done. Second, this consensus cannot be achieved properly when the arguments for any given set of indicators are presented as technical questions about assessment, because this hides the fact that the arguments for indicators are implicitly imposing a very particular understanding of what resilience entails, and what the role of development policy and aid should be for vulnerable people. There is a sleight of hand here, where the eye is made to follow the magic wand of the technical argument, allowing the rabbit – the political decision about development policy – to be concealed in the hat unseen by the distracted audience. Third, in some cases the measurement approaches go further, leaving no space for analysing the very issues which need to be discussed. Once language has been set, the limits for possible debate follow. Many of the current discussions around the measurement of resilience may reinforce existing tendencies to target spending on predetermined kinds of activities, rather than on interventions derived from an analysis of what would actually help the people we ought to be assisting. That surely matters.

References

Alinovi, L., E. Mane and D. Romano (2009) *Measuring Household Resilience to Food Insecurity: Application to Palestinian Households*. Working Paper. Rome: FAO.

Alinovi, L., M. D'Errico, E. Mane and D. Romano (2010) *Livelihoods Strategies and Household Resilience to Food Insecurity: An Empirical Analysis to Kenya*. Paper prepared for the European Report on Development. Brussels: EU.

Ashdown, P. (chaired by) (2011) *Humanitarian Emergency Response Review*. London: DFID.

Bahadur, A. V., M, Ibrahim and T, Tanner (2010) *The Resilience Renaissance? Unpacking of Resilience for Tackling Climate Change and Disasters*. Strengthening Climate Resilience Discussion Paper 1. Brighton: Institute of Development Studies.

Barrett, C. B. and M. A. Constas (2013) *Toward a Theory of Resilience for International Development Applications*. Draft for comments. New York: Cornell University.

Béné, C. (2013) *Towards a Quantifiable Measure of Resilience*. IDS Working Paper 434. Sussex: IDS.

Benson, T., S. Mugarura and K. Wanda (2008) 'Impacts in Uganda of Rising Global Food Prices: The Role of Diversified Staples and Limited Price Transmission', *Agricultural Economics*, 39(s1).

Cannon, T. and D. Mueller-Mahn (2010) *Vulnerability, Resilience and Development Discourses in Context of Climate Change*, *Natural Hazards*, *55*,

Carr, E. R. (2013) *Livelihoods as Intimate Government: Reframing the Logic of Livelihoods for Development*, *Third World Quarterly*, 34:1.

Ciani, F. and D. Romano (2013) *Testing for Household Resilience to Food Insecurity: Evidence from Nicaragua*. Department of Economics and Management, University of Florence

Cimellaro, G. P., A. M. Reinhorn and M. Bruneau (2010) 'Seismic Resilience of a Hospital System', *Structure and Infrastructure Engineering*, vol. 6, nos 1–2.

Constas, M. and C. Barrett (2013) 'Principles of Resilience Measurement for Food Insecurity: Metrics, Mechanisms, and Implementation Plans'. Paper presented at the Expert Consultation on Resilience Measurement Related to Food Security, Rome, 19–21 February. Rome: FAO.

Constas, M. A., T. R. Frankenberger and J. Hoddinott (2014) *Resilience Measurement Principles. Toward an Agenda for Measurement Design*. Technical Series No. 1, Resilience Measurement Technical Working Group. Rome: FAO and WFP.

Darcy, J. and C.-A. Hofmann (2003) *According to Need? Needs Assessment and Decision-making in the Humanitarian Sector*. HPG Report 15. London: ODI.

DFID (2011) *Defining Disaster Resilience: A DFID Approach Paper*. London: DFID.

FEWS (2008) *Mozambique Livelihood Baseline Profile, Zones 16a (River Shore) & 16b (River Valley)*. Maputo: FEWS.

Frankenberger, T., M. Langworthy, T. Spangler and S. Nelson (2012) *Enhancing Resilience to Food Security Shocks*. White Paper (draft). Tucson, AZ: Tango International.

Frankenberger, T. and S. Nelson (2013) *Background Paper for the Expert Consultation on Resilience Measurement for Food Security*. FSIN, www.fsincop.net.

Hahn, M. B., A. M. Riederer and S. O. Foster (2009) 'The Livelihood Vulnerability Index: A Pragmatic Approach to Assessing Risks from Climate Variability and Change. A Case Study in Mozambique', *Global Environmental Change*, vol. 19, issue 1.

Holzmann, P., with T. Boudreau, J. Holt, M. Lawrence and M. O'Donnell (2008) *The Household Economy Approach: A Guide for Programme Planners and Policy-makers*. London: Save the Children.

Hughes, K. (2013) *A Multidimensional Approach for Measuring Resilience*. Working Paper. Oxford: Oxfam GB.

Inter-agency Resilience Working Group (2012) *The Characteristics of Resilience Building: A Discussion Paper*, http://community.eldis.org/resiliencewg.

Jaspars, S. and S. O'Callaghan (2010) *Challenging Choices: Protection and Livelihoods in Conflict: Case Studies from Darfur, Chechnya, Sri Lanka and the Occupied Palestinian Territories*. HPG Report 31. London: ODI.

Klein, R. J. T., R. Nicholls and F. Thomalla (2003) Resilience to Natural Hazards: How Useful Is This Concept?', *Global Environmental Change Part B: Environmental Hazards*, 5, (1–2).

KPMG and Oxford Economics (2013) *2013 Change Readiness Index*.

Lautze, S. and A. Raven-Roberts (2006) 'Violence and Complex Humanitarian Emergencies: Implications for Livelihoods Models', *Disasters*, 30.

Levine, S. (2010) *A Food Security Analysis of Karamoja: Why Pastoralism Is Not the Problem But the Solution*. Kampala: FAO.

Levine, S., E. Ludi and L. Jones (2011) *Rethinking Support for Adaptive Capacity to Climate Change: The Role of Development Interventions*. London: ODI.

Levine, S. and I. Mosel (2014) *Supporting Resilience in Difficult Places: A Critical Look at Applying the 'Resilience' Concept in Countries Where Crises Are the Norm*. HPG Commissioned Paper. London: ODI.

Manyena, S. B. (2006) 'The Concept of Resilience Revisited', *Disasters*, 30(4).

Marsland, N., S. Mohamed and M. Immink (eds) (2011) *Developing a Response Analysis Framework for Food Security Emergencies*. Discussion Papers. Rome: FAO.

Masten, A. S. and J. Obradovic (2006) 'Competence and Resilience in Development', *Annals of the New York Academy of Sciences*, 1,094.

Maxwell, D., J. W. Parker and H. Stobaugh (2013) 'What Drives Program Choice in Food Security Crises? Examining the "Response Analysis" Question', *World Development*, vol. 49.

Maxwell, D. and H. Stobaugh (2012) *Response Analysis: What Drives Program Choice?*. Medford, MA: Tufts University.

MCEER (2006) *MCEER's Resilience Framework: Resilience Concept Drives Development of New Knowledge, Tools and Technologies*. Buffalo, NY: State University of New York.

Mitchell, T. and K. Harris (2012) *Resilience: A Risk Management Approach. ODI Background Note*. London: ODI.

Pain, A. (2012) *A Conceptual Analysis of Livelihoods and Resilience: Addressing the 'Insecurity of Agency'*. HPG Working Paper. London: ODI.

Pain, A. and P. Kantor (2012) *Local Institutions, Livelihoods and Vulnerability: Lessons from Afghanistan*. HPG Working Paper. London: ODI.

Pfefferbaum, R. L., B. Pfefferbaum and R. L. Van Horn (2011) *Communities Advancing Resilience Toolkit (CART): The CART Integrated System*. Terrorism and Disaster Center at the University of Oklahoma Health Sciences Center.

Renschler, C. et al. (2010) *Framework for Defining and Measuring Resilience at the Community Scale: The PEOPLES Resilience Framework. MCEER Technical Report –MCEER-10-006*. Buffalo, NY: State University of New York.

Strunz, S. (2011) *Is Conceptual Vagueness an Asset? Resilience Research from the Perspective of Philosophy of Science*. Working Paper Series in Economics No. 205. Lüneburg: University of Lüneburg.

Sullivan, C. (2002) 'Calculating a Water Poverty Index', *World Development*, 30.

Tulane University with State University of Haiti (2012) *Haiti Humanitarian Assistance Evaluation from a Resilience Perspective*. New Orleans, LA: Disaster Resilience Leadership Academy, Tulane University.

Ulimwengu, J. M., S. Workneh and Z. Paulos (2009) *Impact of Soaring Prices in Ethiopia: Does Location Matter?*, IFPRI Discussion Paper 846.

Venton, C. C. and C. Fitzgibbon (2013) *Community Based Resilience Analysis (CoBRA) Implementation Guidelines Version 1*. Nairobi: UNDP.

Venton, C. C. et al. (2012) *The Economics of Early Response and Disaster Resilience: Lessons from Kenya and Ethiopia*. Final Report. London: DFID.

Wisner, B., P. Blaikie, T. Cannon and I. Davis (2004) *At Risk: Natural Hazards, People's Vulnerability and Disasters (2nd ed.)*. London and New York: Routledge.

Wood, G. (2003) 'Staying Secure, Staying Poor: The Faustian Bargain', *World Development*, 31 (3).

Young, H. et al. (2009) *Livelihoods, Power and Choice: The Vulnerability of the Northern Rizaygat, Darfur, Sudan*. Medford, MA: Tufts University.

Young, H. et al. (2005) *Darfur: Livelihoods under Siege*. Medford, MA: Tufts University.

HPG

Humanitarian Policy Group

The Humanitarian Policy Group is one of the world's leading teams of independent researchers and information professionals working on humanitarian issues. It is dedicated to improving humanitarian policy and practice through a combination of high-quality analysis, dialogue and debate.

ISBN: 978 1 909464 83 4

Humanitarian Policy Group
Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
United Kingdom

Tel. +44 (0) 20 7922 0300
Fax. +44 (0) 20 7922 0399
E-mail: hpgadmin@odi.org
Website: http://www.odi.org/hpg

ODI