Revealing values in a complex environmental program: a scaling up of values-based indicators

Dimity Podger[a], Elona Hoover [a, b], Gemma Burford [a], Tomas Hak [c] & Marie K. Harder [a,b]*

[a] Values and Sustainability Research Group, University of Brighton, Brighton, BN2 4GJ, United Kingdom

[b] Department of Environmental Science and Engineering, Fudan University, 220 Handan Rd., Shanghai, 200433, PR China

[c] Charles University Environment Center (CUEC), José Martího 2/407, 16200, Prague 6, Czech Republic

* Corresponding author.

1.   Introduction

In the time since the Earth Summit held in Rio de Janeiro in 1992, a field of endeavor has developed that focuses on measuring the effectiveness of international, national, regional and community sustainability initiatives through a wide range of indicators (e.g. Bell and Morse, 2008; Hardi and Zdan, 1997; Meadows, 1998; Reid et al., 2006; UN Division for Sustainable Development. Department of Policy Co-ordination and Sustainable Development, 1995, 1996;

UN Division for Sustainable Development. UN Department of Economic and Social Affairs, 2000). The myriad of indicator sets developed have tended to provide measures of environmental, social and economic phenomena, and more recently measures of wellbeing, quality of life, and happiness (European Union, 2015; Fraser et al., 2006; Meadows, 1998; Stiglitz et al., 2009).

It is recognized that each indicator set implicitly embodies a particular set of *values* – a term that can be understood both in the sense of people's judgments about what is important in life (i.e. what they value), and in the sense of ethical principles or standards that guide human behavior (Oxford English Dictionary, 2013). The values embedded in those indicator sets generally reflect those of the evaluating body, and through their measurement and policy utilization, can inadvertently reinforce, encourage and even direct particular sustainability outcomes (Reid et al., 2006, p. 14), while overlooking (and potentially marginalizing) others (Gudmundsson, 2003; McCool and Stankey, 2004; Rametsteiner et al., 2011; Rosenström, 2006). However, these mainstream indicator initiatives do not explicitly explore the role that the ethical values of the actively engaged individuals, communities or institutions play in efforts to address sustainability issues – values such as compassion, integrity, justice and respect (Burford et al., 2013a), which can be applied on the one hand to interpersonal relationships and on the other to human interactions with the wider community of life. This is despite the emphasis that various sustainable development documents, particularly those of the United Nations around the Decade of Education for Sustainable Development, place on the importance of addressing those values embedded in social, economic and political affairs that have 'put the world on an "unsustainable" path' (UNESCO, 1997, p. 8). Both the Earth Charter (Corcoran et al., 2005; ECI Secretariat, 2010) and the United Nations Millennium Declaration (United Nations General Assembly, 2000)

have explicitly called for an 'ethical framework' for sustainability, and specified particular values that could be included in it.

There is a common view in sustainability discourses that it is crucial for everyone to "get our thinking right: to see the interrelations among these problems [of sustainable development] and recognize the fundamental need to develop a new" (UNESCO, 1997) (see also Crompton et al., 2010; De Leo, 2012; Fuad-Luke, 2009; Tilbury and Wortman, 2004; Walker, 2006). One perspective, espoused by authors with socially critical orientations, frames social injustice, inequality and inequity as root causes of unsustainable societies and argues that these problems must be addressed as a matter of urgency in order to ensure a better quality of life for everyone without destroying the natural environment (Robottom and Hart, 1993; Tilbury and Cooke, 2005a). In parallel, it has been argued by ecological philosopher David Abrams (Abrams, 1996) that a root cause of unsustainability is the widespread human failure to connect deeply with the 'more-than-human' community, and to engage with its members as perceiving subjects rather than as objects for human manipulation. We do not take up a position, in this paper, about the respective merits of these arguments or the relationship(s) between them. Rather, we point out that within both arguments is a recognition of the roles that values can play in acting as obstacles, or as motivators, to the task of societal and organizational transformation towards sustainability. There are others who suggest that in order to make progress beyond a narrow focus on specific environmental or social problems, it is necessary to shift attention to worldviews, which frame what is or is not seen as important at individual, organizational, institutional and societal levels (Fien et al., 2001; Sterling, 2001, 2007), and our approach may well have overlap. However, our position is that before trying to analyze complex connections between different domains of values implicit in sustainability, organizations and wider society,

the more specific topic of making values based achievements more tangible and measureable within organizations deserves full attention.

In this paper, we describe a research project which was co-designed with four civil society organizations (CSOs) who work to promote and embed awareness of sustainability concepts – a broad form of Education for Sustainable Development (ESD) They shared a common view that consideration of ethical values was a necessary (even if not sufficient) precondition for their work, and expressed a common frustration at their inability to articulate the outcomes that mattered most to them, much less to monitor and evaluate their progress in relation to those core values.

To address this CSO-based issue, researchers from two universities co-initiated the EU-funded ESDinds[1] research project (2009-2011) with a primary research question focused on whether it was possible to develop indicators and assessment tools to evaluate achievements related to core ethical/spiritual values within CSOs promoting ESD (RQ1).   It is important to note that this research did not presuppose or set out to reveal links between the ethical/spiritual values of the CSOs and their ESD achievements, but only to 'make tangible' any values-linked dimensions of their (various) achievements, so that the chain of ethical values throughout their work could itself be made tangible, find measures, and be monitored. When this was achieved, with the approach described in the first part of this paper, a secondary research question emerged: would the set of ethical values-based indicators designed with the four initial partner CSOs be relevant, comprehensible, and useful for evaluation in new organizations that had not

been involved in developing them? This was not initially expected, but was surprisingly borne out and demonstrated in several such field studies (Burford et al., 2013b; Harder et al., 2014b; Podger et al., 2013). In light of this unexpected finding of some transferability of the ESDinds set of `indicators', a third research question emerged (RQ3): could the approach be scaled up, for example, for use in a large organization which had multi-layered activities at several levels, e.g. national, regional and community levels?

In this paper we present a study of that third question, which shows that the ESDinds approach was indeed able to be scaled up and used for a multi-level organization/ project. That result suggested that more care should be taken to document and present the initial process which produced such a useful and novel set of transferable indicators, as their origin and genesis might be key to deeper understandings needed for later considerations. We thus begin our paper by documenting a description of how the main ESDinds 'indicator' set was developed to answer RQ1, including its theoretical grounding, research design and methodology. We briefly summarize the findings for RQ2 from the small-scale field studies (Burford et al., 2013a; Burford et al., 2013b; Harder et al., 2014b; Podger et al., 2013) and then go on to show how the approach was scaled up for systematic values-based evaluation of a national environmental program in Mexico, thus addressing the third research question outlined above (RQ3), and reflect on the implications of this work for wider use in sustainability arenas.

2. Theoretical Grounding

Before proceeding to describe the specifics of the research design and methodology, this section outlines the theoretical grounding of the work, including its positioning in relation to recent literature on the definition, assessment and measurement of sustainability and the development

of 'soft' indicators. We also describe the overall epistemological positioning of the paper within a paradigm of emancipatory research, and the consequences for our understanding of terms such as 'validity' and 'rigour'.

2.1 Defining and measuring sustainability: an impossible task?

.In the last decade, progress has been made with sustainable development indicators to capture more 'soft' characteristics, and to build on those concepts for values-based achievements.

The term 'sustainable development' is most famously defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [32]. The primary indicator of 'development' itself was taken as GDP (gross domestic product), a measure of monetary flow within a national economy (Bell and Morse, 2011), and an assessment of 'sustainability', in relation to resource management, was attained by comparing one indicator - current resource yield – with an assumed benchmark of 'maximum sustainable yield' (Grainger, 2012).

In the intervening decades, however, it has increasingly been recognized that attempts to define sustainability can never be value-neutral (Lele and Norgaard, 1996). The importance of social justice as an element of sustainability, and the 'three-pillar' model or 'triple bottom line' emphasizing environmental, economic and social dimensions, is now widely accepted (Pope et al., 2004); yet many authors have also referred to a less tangible 'fourth pillar' or 'missing dimension' of sustainability (Burford et al., 2013a; Dahl, 2012; Littig and Griessler, 2005), variously characterized as 'cultural' (Hawkes, 2001; Kagan, 2011; UNESCO, 2013; Woodley,

2006), `political/institutional' (European Commission: EUROSTAT; Spangenberg, 2002; Spangenberg et al., 2002; UN Division for Sustainable Development. UN Department of Economic and Social Affairs, 2000), or `religious-spiritual' (Clugston, 2011; Hedlund-de Witt, 2011; Interreligious Statement Towards Rio+20, 2012),   Indeed, the challenges of trying to define sustainability are now well documented in both the environmental management and design fields, to the extent that it has frequently been described as a 'wicked problem': a problem which is confusing, ill-formulated and impossible to solve with scientific and technical approaches alone, and which involves multiple stakeholders with conflicting values (Buchanan, 1995; Fuad-Luke, 2009; Grainger, 2012; Wang, 2002).

In this climate, as noted by McCool and Stankey (McCool and Stankey, 2004, p. 295), public debate about what should be 'sustained' is often preceded or even pre-empted by attempts to develop indicators: "the search for indicators is guided more by what can be measured (a technical issue) than by what should be measured (a normative issue)." Sustainability indicator initiatives typically focus on phenomena that are tangible, measurable and comparable across geographic and cultural spectra, leading to the development of such tools as the Ecological Footprint (Hak et al., 2007) and highly aggregated indices, such as the Environmental Sustainability Index (Bell and Morse, 2008; ESI, 2005; Global Leaders of Tomorrow Environment Task Force, 2005). Yet alongside the growth of this "consensual global knowledge which erases difference and allows the most powerful to determine what is 'known' " (Hulme, 2010, p. 563), there has also been a growing interest in recognizing and valorizing *other* knowledges, often through the integration of 'top-down' (expert-led) and 'bottom-up' (participatory or community-based) approaches to the development of sustainability indicators (Bond and Morrison-Saunders, 2011; Fraser et al., 2006; Reed et al., 2005; Reed and Dougill,

2002; Reed et al., 2008; Reed et al., 2006). Diverse sustainability assessment initiatives, from buildings (Bragança et al., 2010) and manufacturing (Othman et al., 2010) to natural resource management (Reed et al., 2006) and city-level sustainability assessment (RFSC, 2013), have incorporated 'soft' or qualitative indicators that arguably reflect very different values and worldviews from those underpinning the standard economic model.

Grainger (Grainger, 2012) suggests that if attempts to accommodate multiple stakeholder perspectives are not simply to result in an accumulation of more and more indicators, a new paradigm of indicator development will be required, although he does not set out a clear vision of what this might entail. It is evident that that the quest for a universal and all-encompassing definition of sustainability is ultimately in vain, and that the idea of 'objectivity' or 'neutrality' in sustainability assessment is a myth which serves to conceal the normative assumptions of powerful institutions (Bell and Morse, 2008; Bell and Morse, 2011; Bond and Morrison-Saunders, 2011; Bormann, 2007; Burford et al., 2013a). We would argue, therefore, that the new paradigm of indicator development must make space for both *intersubjectivity* (the construction of a shared understanding through authentic dialogue within in the context of a collaborative practical activity, grounded in shared experience (Talamo and Pozzi, 2011)) and what we term *multi-subjectivity*, i.e. the accommodation of diverse perspectives in the context of a collaborative practical activity *without* attempting to synthesize them, harmonize them or build a shared understanding. We would argue that the values-based approach outlined in this paper is one way (although by no means the *only* way) of integrating intersubjectivity and multi-subjectivity within an indicator development process, and as such, potentially offers important lessons for the field of sustainability assessment. However, our first focus is to use the approach

to make values-based aspects of CSO work tangible: extensions to measures for less tangible aspects of sustainability may follow later.

2.2. Definitions

In this section, we set out the definitions of *indicators* and *values* that were used within the project, as well as introducing the new term *proto-indicators*.

2.2.1. 'Indicators' and 'proto-indicators'

Indicators are understood as proxy measures that are used to monitor the progress of projects, in conjunction with other processes which aid decision-making (Reid et al., 2006). The initial definition used in the project was that of Gallopin (Gallopin, 1997), describing useful indicators as 'variables that summarize or otherwise simplify relevant information, make visible or perceptible phenomena of interest, and quantify, measure, and communicate relevant information.' The phenomena of interest for ESDinds were those values considered by, or found by, CSOs to be important to them in their work: in this sense, we were developing indicators for the ethical principles, practices, and outcomes that are the operational expression of ethical values (Anello, 2006, p. 19). Our understanding of indicators evolved further over the course of the project, however, following the 2011 publication by Hinkel (Hinkel, 2011) that defines an indicator as the *function* linking an observable or measurable variable with a theoretical variable that cannot be directly measured – e.g. 'when air pollution increases, the percentage of tree surface covered by lichen species X decreases'. This led us to introduce the term 'proto-indicators' to refer to "statements or criteria which can guide the identification of measurable

indicators for intangible values, through local conceptualization and operationalization"(Harder et al., 2014a, p. 125).

Through this refinement of the definition of indicators, the process of developing what we have termed 'values-based indicators' consists of two distinct stages: (1) the development of a reference list of proto-indicators with generic wording (e.g. 'People feel that they are encouraged to fulfil their responsibilities'); and (2) the definition of specific measurable variables at the local (i.e. project or organization) level, which can be described as 'indicators' of the enactment of values (e.g. 'Youth members feel encouraged to fulfil their responsibilities within their project') .

2.2.2. 'Values' and `validity'

Values are generally considered to be intangible and therefore unable to be weighed, measured or counted directly (Schlater and Sontag, 1994). However, Handy (Handy, 1970) argues that behaviors and practices associated with values can be observed and measured. The definition of 'values' used by Rokeach (Rokeach, 1973) brings together these two aspects of a value, of both meaning and practice, and was initially used as a guide for designing the data collection tools for the project. Rokeach (ibid, p. 5) writes: 'A value is an enduring belief that a specific mode of conduct or end-state of existence is personally or socially preferable to an opposite or converse mode of conduct or end-state of existence. A value system is an enduring organization of beliefs concerning preferable modes of conduct or end-states of existence along a continuum of relative importance.'

Judgment can be made as to what is considered ethical, and whether certain values constitute enablers of sustainability or barriers to it in particular contexts. ESDinds, to the extent possible, focused on identifying those values that partner CSOs themselves considered to be *ethical* values, which some might call `human', `spiritual' or `moral' values. Ethical values are understood as 'what is believed to be good and of primary importance to human civilisation, and are often articulated as ideals…inform[ing] judgement by defining right from wrong, and good behaviour from bad' (Anello, 2006, p. 19). ESDinds also holds that understandings of morality, and of what is 'right' and 'wrong' with respect to sustainability, are necessarily locally defined by their context, as individuals, communities and institutions engage with the ideas of sustainability. Therefore, this paper attempts to examine the phenomena of values not only through the development of specific (yet customizable) indicators, but also by designing an indicator development process that others may adapt and use in their own contexts.

We were aware from the outset that values and their measurement have been an important area of research in social psychology for more than half a century, and many different theoretical frameworks have been created – from the original 'Study of Values' developed by Allport, Vernon and Lindzey (Allport et al., 1951), through the seminal work of Milton Rokeach (Rokeach, 1973, 1979) and Schwartz (Schwartz, 1992, 1994, 2007; Schwartz and Bilsky, 1987, 1990; Schwartz et al., 1999; Schwartz et al., 2001), to recent environmentally-oriented frameworks distinguishing 'biospheric values' from egocentric and altruistic values (de Groot and Steg, 2007; Stern and Dietz, 1994). Yet while drawing its initial working definition of values from Rokeach (Rokeach, 1973), the ESDinds project did not take any specific framework as its theoretical starting point. Rather, the project was grounded in a paradigm of emancipatory research, defined as research that promotes in the participating practitioners "a critical

consciousness which exhibits itself in political as well as practical action to promote change"
(Grundy, 1987). Among the methodological principles of emancipatory research are *reciprocity
and mutual education*, *dialogue*, *critical praxis*, and *catalytic validity* (Lather, 1986b). These
principles informed a collaborative, dialogic, exploratory, and reflexive approach to the data
collection and analysis methods, not only for values identification, but also for the development
of the reference list of proto-indicators (Podger et al., 2013). We were to find later that this
approach was key to developing effective ways for values to be 'elicited' and crystallised in a
valid manner – a result which we found has many important uses in its own right. In terms of
evaluation it was a vital step on the way to defining what could be measured, i.e. the proto-
indicators, but in fact the problem of first crystallising unarticulated shared values is non-trivial.
Although not a focus of this paper, it is central to other studies (Podger et al., 2013).

The principle of catalytic validity, also described by Messick (Messick, 1989) as
'consequential validity' (i.e. the capability to bring about desired consequences), was particularly
central to the project (Harder et al., 2014a). The overarching goal of the research was to create
values-based indicators and accompanying assessment tools that would be locally valid and
useful, to the extent of being radically transformative, for the partner CSOs (Harder et al., 2014b;
Podger et al., 2013).

3. Overall Research Design and Methodology

Within the emancipatory paradigm outlined above, there were three important primary
elements that characterised the research design of the ESDinds project.

Firstly, the research direction was set and adapted through a collaborative partnership
between the two research institutions and four CSOs, combining technical expertise in indicator

development with practical experience in applying ethical values. A formal process of consultation and review was established in the design, with four meetings of all partners scheduled over the life of the two year project. The research design directed the balance of power and control over the outcomes of the research towards the CSO partners who held the decision-making power utilised at each six-month milestone meeting, to determine research priorities for the coming phase. This design ensured that the output of the project was useful and meaningful to them.

Secondly, the research focus was addressed through three phases of empirically grounded research. The first of these phases involved identifying a broad 'pool' of important ethical values from and for the CSO partners and deriving from them a pilot set of associated proto-indicators to assess their presence. The second and third phases both focused on testing and developing the pilot proto-indicators in field studies, in parallel with developing locally appropriate measureable indicators and assessment tools, and refining the indicator development process itself for wider use and dissemination to other CSOs[2]. Each of the three phases had five general stages of process designed within them:

(i)   consultation among the core group of partners setting the overall direction;

(ii)   consultation among the university teams clarifying the research focus and approach;

(iii)   data collection involving case study research;

(iv)   data analysis by the university teams, in collaboration with the CSOs (with varying degrees of involvement), and drawing on academic literature; and

---

[2] A further element of the research, implemented in parallel with the third phase, involved engagement of over 100 additional CSOs through an online platform to test the indicators, the assessment tools, and the process of identifying and describing values through indicators.

(v)   technical review by the CSOs for decision-making for the next phase.

The design exemplified a `bottom up and top down' approach to proto-indicator development (Fraser et al., 2006), integrating technical expertise and practical experience. This was translated into ongoing refinement of both the proto-indicators themselves and the indicator development process, through a combination of fieldwork on the ground and reflective dialogues within the core group.

Figure 1 depicts the major phases of the research implemented over the two years.

INSERT FIGURE 1 HERE

Finally, case study methodology (Stake, 2000) was adopted. Case study methodology embraces in-depth qualitative methods of inquiry, and allows for the critically reflective dialogue necessary to investigate the complex phenomena of values, as well as participant observation, and document analysis. Within the case study approach, the project emphasized a diversity sampling approach, using a variety of strategies (according to the size and nature of the partner CSOs) and sampling multiple sites within a given organisation where appropriate. This is congruent with literature outlining best practice for the development of community-level indicators, which emphasizes the importance of engaging broad participation throughout the process (Hardi and Zdan, 1997; Palmer and Conlin, 1997; Smolko et al., 2006).

3.1   Case Study Methodology and Data Collection

The goal of the first phase of ESDinds was to understand the values espoused and active for each separate CSO at the organizational level, project level, and at the level of members or participants. The principles of critical praxis oriented inquiry, outlined by Lather (Lather, 1986a, b), supported a dialogic orientation to the various data collection methods, which included interactive semi-structured key informant interviews (as critical conversations between interviewer and interviewee that allowed for the retrospective analysis of practical experiences) as well as group interviews and consensus building, informal conversations and focus groups (Taylor and Bogdan, 1998). Observation and document collection were also important data collection techniques (Creswell, 2013). For each CSO, the data was triangulated, with different methods used to address the research focus. This multi-method approach facilitated the collection of diverse perspectives about which values were considered important by different members. Where there was diversity in intra-organisational values, we did not engage in any debate, but rather incorporated all the differing perspectives within the dataset on an equal footing (i.e. multi-subjectivity), with a view to increasing the diversity of the resulting proto-indicators. Actual conflict in values was not found in these cases.

The approach to engaging broad participation differed from one CSO partner to the next. One organization was small enough for universal participation by staff and volunteers, while in another, a well-timed international conference provided an ideal opportunity to broaden data collection. The size and geographic spread of the two remaining partner CSOs limited the breadth of participation, although in both of these organizations, key informant interviews were achieved with the majority of current and former project staff.

The questions used for semi-structured key informant interviews, group interviews and surveys also varied slightly between CSOs. In some contexts, the language of 'values' was

already being used explicitly by staff, consultants or partners, which allowed for conversations around direct questions such as: *Which values are important in your organization/project? Why are they important? What are some examples of how these values are practiced or implemented in your organization/project? How would you know whether these values were present in your work, or in the outcomes of your work?* In contexts where 'values' had not been discussed overtly, a less direct approach was taken, e.g. focusing on the overall approach of the organization and on a specific project that respondents saw as particularly successful or worthwhile, as an entry to a conversation about values. In addition to the formal interviews, informal conversations were also used to clarify the meanings that interviewees attached to values, and to draw out new proto-indicators.

3.2   Data Analysis

The aims of the data analysis for each CSO were to: (i) identify the core values perceived by the CSOs to be important in the success/progress of the projects or organizations, or the kinds of value outcomes hoped for; (ii) understand the spectrum of meanings and expressions associated with these values by interviewees; and (iii) identify or derive possible proto-indicators and assessment tools that could evaluate the presence of values in the goals, processes and impacts of the CSOs.

An initial code book, populated with schemes and classifications from literature on values (e.g. Hitlin and Piliavin, 2004; Peterson and Seligman, 2004; Rokeach, 1973; Rokeach, 1979; Schwartz and Bilsky, 1990) and indicators (e.g. Sollart, 2005; Tilbury and Janousek, 2006; UNESCO, 2007) was developed to coordinate a collaborative analysis of the large dataset by the five-member research team, using the qualitative software package *Atlas.ti*. As each researcher

analyzed their specific CSO dataset, she or he added further subcategories as required (such as new values found in the data) and circulated the new code book to the other team members. The coding categories primarily used by the research team were Value Type, classifying quotations by a value (e.g. 'integrity', 'justice' or 'respect'), and Indicator Type, classifying quotations according to different types of indicators (e.g. baseline, context, process, learning, output, outcome, impact and performance indicators) (UNESCO, 2007). The latter were not ultimately found to be useful and were later disregarded except to stimulate separate discussions about their relevance.

A content analysis approach was used to develop potential proto-indicators that were either described directly in the text or could be derived (e.g. where interviewees or documents alluded to ways in which a value could be identified in practice). To the extent possible, the original wording contained in the dataset was used to develop the potential proto-indicators. All such statements were coded with a value label, and if applicable the indicator type and level. Finally, following the initial stage of analysis, the precise wording of some of the proto-indicators was modified slightly so that they were (at least in principle) generic enough to apply to other CSOs (e.g. 'my team' changed to 'people').

The initial stage of analysis of data produced a list of 125 values, most of which were linked to multiple proto-indicators in the data. In order to identify a smaller pool of ethical values considered significant and important by the partner CSOs, and manageable for the field work, a basic quantitative analysis was then made to identify the most commonly occurring values within the data set of each individual CSO. This was based primarily on frequency of occurrence across that data set, and secondarily on density within a data source. It was found that the top five values of each CSO had considerable collective overlap, so, based on a principle of fairness, the

top value from each of the four CSOs was selected, plus another that was common to a majority of the CSOs: *unity in diversity*, *trust/trustworthiness*, *justice*, *empowerment*, and *integrity*. Each CSO had at least 3 of their top 5 values represented. The research team prepared 'meaning statements' for these significant values from the data, to clearly communicate their intended meaning from their context. An extensive list of potential indicators was then compiled from these values (examples are given in Table 1 below). In addition, for each value selected, the research team conducted an in-depth cross-disciplinary review of relevant literature in order to identify themes within each value construct and identify potential assessment tools. The proto-indicators developed reflected a wide range of themes and multiple dimensions of meaning and practice of each value.

INSERT TABLE 1 HERE

It should be noted that, as the primary research was focused on 'intangibles' needing the development of indicators, and as there were already several established systems for measuring traditional financial and environmental outputs of projects, any values focused on those were not purposely elicited. The raw data set contained 'ethical' values-related concepts.

3.3   Validation and Technical Review

The preliminary results from each CSO case study – consisting of the pool of ethical values found to be important for CSO activities and sustainability, the contextual statement of meanings and practices, associated proto-indicators, and initial ideas of possible measurement methods - were presented to the Core Group of partners in advance of a formal review meeting. This

meeting, concluding the first phase, included a workshop that involved all four CSO partners in collaboratively prioritizing and rewording the proto-indicators (with the research team listening) to generate a collective list that they regarded as potentially useful across a range of CSOs. In filtering the data down to such a small number of values, it was clearly unrealistic to expect that the resulting draft indicator set would be representative of everything that was important to project partners. It is important to note, however, that it was already apparent that a number of potential indicators related to more than one value label. In accordance with the principles of critical praxis and catalytic validity, ample opportunity was given to allow the CSOs to review and comment on the 'fitness for purpose' of the five value-labels of *Trust/Trustworthiness, Integrity, Unity in Diversity, Empowerment* and *Justice* and their associated proto-indicators. It emerged, at this point, that two of the CSO partners felt that the important element of care towards the environment (or other life forms) was missing from the lists, and that for them the domain of `ethical values' could not be adequately represented without also considering that dimension, e.g. *Care and Respect for the Community of Life*, which is itself a core principle of the Earth Charter. Accordingly, the CSOs as a group called for a revision of the list, which necessitated a second phase of data analysis, aimed specifically at developing potential proto-indicators relating to this value. We acknowledge that it would have been methodologically more robust to have these in mind during the elicitation process, and that future work on values elicitation will need to start from a broader perspective, considering not only interpersonal relationships but also human-biosphere relationships from the start. Based on the comments from CSO partners and discussions during the Core Group Meeting, researchers developed a first list of proto-indicators (N=76) for the first phase of field trials.

Figure 2 below provides an overview of all the steps outlined above, which collectively comprised the first phase of the ESDinds project.

INSERT FIGURE 2 HERE

3.4. Small-Scale Field Trials

The second research question RQ2 was explored in the second and third phases of the ESDinds project, where small-scale field trials of the values-based proto-indicators were conducted within several different CSOs on the ground (Burford et al., 2013b). Although not the focus of this paper, we note for reference that these field trials validated the overall concept of a values-focused evaluation and led to the development of a reference list of (N=166) proto-indicators that were found to be broadly relevant and useful across different types of CSO, thus answering the second research question, RQ2. The approach was also found to be suitable for locally valid assessments, although slight changes to the wording were often necessary to confirm local ownership of the statements – what we called 'flexible wording'.

In addition, the early trials generated significant learning about 'values' and 'indicators' – notably that the proto-indicators in the list were not only perceived as useful for indicating the presence of the six values with which they had originally been associated, but could also be used 'stand-alone' as a 'pool' or 'menu' for constructing other values that were already espoused by the CSOs, e.g. in mission statements (Burford et al., 2013b; Podger et al., 2013). A subsequent systematic study (within the context of an evaluation of an online course about sustainability) illustrated that it was also possible to use locally-elicited values as a starting point rather than the

original six, and still identify a large number of proto-indicators from the 'menu' list that were deemed by users to be relevant (Burford et al., 2013a). At this point it became evident that this CSO-informed approach, developed in the field, of forming a 'menu' of prioritized proto-indicators was surprisingly powerful and had the potential for much wider application. The 'menu' of proto-indicators seemed, between them, to cover many of the 'important things' for CSOs. However, it was unclear whether the approach could be suitable for large-scale, multi-layered programs, where potentially quite different values might be active in different spheres. Furthermore, the interpretation of each might be very different in each sphere, e.g. aspects of equality to a manager compared to a grass roots worker. This, then formed our third research question (RQ3) and is a focus of this paper. The study described in Section 4 was designed to explore this question.

4. Systematising Values-Focused Evaluation: An Illustrative Example

In this section, we report on a larger and more systematic field trial of values-based indicators through participatory evaluation, within a nation-wide environmental organisation in Mexico, set up with a view to answering research question RQ3 above. We were particularly interested in finding out whether proto-indicators could be localized sufficiently to maintain relevance to the different stakeholders (from grassroots projects, right up to the level of senior management) yet still retain some kind of comparability across the various spheres.

4.1. Context

The evaluation was developed for a youth leadership project delivered by a national Mexican NGO. The context of the project needs to be understood in terms of five spheres of analysis: (a) the national NGO (RM); (b) its 'Better People, Better Forests' program (MpMb); (c) 'Youth Leadership for Sustainable Forests' (YLSF), a specific project within the MpMb program; (d) the grassroots projects carried out by the YLSF youth leaders; and (e) the personal lives of the youth leaders.

(a) **Organisational sphere:** 'RM' is a Mexican civil society organisation whose vision is to develop temperate and tropical forests that provide abundant resources and ecological services, where people can live with dignity and in harmony with their environment. The organisation was founded after extensive forest fires in 1998, when the Secretariat for the Environment and Natural Resources invited Mexican companies to participate in funding reforestation programmes and projects in affected areas. Following a successful initiative in which a group of private sector companies matched all donations by their employees, resulting in the planting of several million trees across eight areas of natural protection, RM was created in 2002 as a permanent structure to give continuity to these projects.

(b) **MpMb program sphere:** 'Mejores Personas Mejores Bosques' or *Better People Better Forests* (MpMb) is one of four programs within RM. Through the MpMb program, the organisation works with rural leaders to initiate local changes for community development and better local organisation, which facilitates participatory actions in favour of the forests. (Note that the MpMb program level was not explicitly evaluated in this study).

(c) YLSF project sphere: 'Liderazgo Juvenil para la Sustentabilidad Forestal' or *Youth Leadership for Sustainable Forests* (hereafter YLSF) is a project within the MpMb program. YLSF's essential strategy is to stimulate alliances between youth leaders from strategic communities across the country to develop a network where, through training for collective action, necessary tools and capacity are developed to enable the sustainable development of Mexico's forests. The YLSF network is formed of 44 youth 'agents of change', in 17 regions of Mexico that are seen as strategic areas from the perspective of sustainable forests (across 11 states).

(d) Grassroots project sphere: An essential aspect of the YLSF project is that the participating youth leaders are not expected merely to meet and share experiences, but also to lead grassroots environmental projects in their local communities.

(e) Personal level: Participating youth leaders also engaged in related activities in their personal lives that supported or were influenced by their membership in the YLSF network.

4.2. Evaluation purpose

In accordance with the emancipatory paradigm of ESDinds as described above, and established principles of participatory evaluation (Cousins and Chouinard, 2012; Cousins and Whitmore, 1998; Crishna, 2007; Estrella and Gaventa, 1997; Springett, 2003), the evaluation of the YLSF project was not structured in a 'top-down' way (i.e. senior management selecting a set of values and associated indicators and using those to evaluate the grassroots projects). Rather, its fundamental purpose was to identify which values existed throughout the network and which

actions demonstrated the presence of values in youth leaders, as 'agents' of the *Better People, Better Forests* programme. Values important to the program were identified from the youth participating in the YLSF project as well as those values which employees and project coordinators in RM were hoping to embed and encourage within the program. In practice, this meant that senior managers were open to the possibility of overturning conventional power structures: by allowing what was already being 'lived' at the grassroots to be articulated and valued, they effectively granted the youth leaders authority to re-envision and redefine the formal values statement for the entire project. The elicitation and evaluation process also enabled both the youth leaders and core RM staff to identify potential misconceptions and situations where values were espoused by the organisation, but not recognised in practice, and vice-versa.

4.3. Evaluation methods

The evaluation fieldwork consisted of field visits in six states in Mexico; a workshop with RM staff; an intensive workshop with the National Youth Council of the YLSF project; and a final workshop with RM operational staff involved in YLSF and the MpMb program. In total, 95% of the youth leaders involved in YLSF were involved in the evaluation process, and 100% of relevant staff.

Indicator selection and development: YLSF project managers decided early on to evaluate as many proto-indicator as relevant in order to minimise the bias of selection from senior staff. Thus, a survey instrument was developed using 163 of the 166 *ESDinds* proto-indicators, together with three new proto-indicators that were added in response to

the situation of insecurity being experienced in Mexico at the time of the evaluation. Four potential spheres were identified for evaluating each proto-indicator, namely RM itself (the wider organisation), YLSF as a project, specific projects led by the youth, and their personal experience. (It was decided not to evaluate the MpMb programme at the same time as this included many of the same staff than the RM and YLSF spheres.) Some, but not all, proto-indicators were identified as relevant for multiple spheres. Follow this, each proto-indicator was modified to develop a survey instrument with a total of 203 specific indicators. This allowed for a comparison of results between spheres.

Data collection: Each field visit included a presentation of the evaluation and its purpose, RM's expectations, and an open reflection on each participant's historical involvement in YLSF and the story behind their projects. This initial space for reflection was also used to facilitate a discussion around values. The process was based on establishing an open dialogue about the youth leaders' personal and collective experiences, in which all contributions were appreciated and carefully documented by the research team, resulting in a list of locally elicited 'values'. After that, the participating youth leaders (n=44) were asked to respond to the complete survey based on the 203 indicators. This was delivered in a participatory manner: the researchers asked the questions aloud to the different groups, checking that they had been well understood and then inviting the participants to respond on a three-point scale, either through the use of different physical spaces (moving to stand on a particular colour, in a spiral constructed on the ground from three coloured scarves) or by using body positions (e.g. arms up or crouching down) to denote their response (Harder et al., 2014b; Podger et al., 2013).

Figure 3 shows the different representations of each of the three scale points. In order to break up the four-hour process, the survey was administered in three phases, interspersed with games and changing group dynamics. In addition to the work with youth leaders, the survey questions were also used with staff from RM involved in strategic oversight, management and delivery of the MpMb program and YLSF project.

INSERT FIGURE 3 HERE

Data analysis: As the aim of the process was to understand and communicate the presence of values in YLSF, the last phase of the fieldwork consisted in a process of identifying and clarifying values and linking those values to the appropriate indicators, as determined locally. This step thus enabled the data (locally determined indicators) to be analysed in terms of values (locally defined and identified). In order to complete this final phase of the evaluation, two participatory workshops were organised as follows:

- A workshop with the national youth representatives of the YLSF project: initial semantic groupings were identified from the complete list of all the values that had been mentioned by youth leaders and RM staff. This analysis had the twofold aim of identifying values that appeared to relate to each other (e.g. *communication* linked to *dialogue, coexistence, clarity,* etc.), and highlighting those values that were most commonly mentioned across the whole dataset.   The visual metaphor of a tree was used to organise the values, with different meanings attached to the various parts – branches, trunk, roots, sap, etc (Figure 4, Table 2). This facilitated the identification

of eight core values or 'mother values': one to represent each of the four clusters of closely related values (*Love*, *Trust*, *Identity* and *Solidarity*), and a further four to represent the most mentioned values (*Communication*, *Autonomy*, *Knowledge* and *Responsibility*).

INSERT FIGURE 4 HERE

INSERT TABLE 2 HERE

- A workshop with the MpMb operational team in RM and youth leaders from YLSF programme: each of the 203 indicators was associated with one (and only one) of the eight 'mother values'.   Some examples are included in Table 3.

INSERT TABLE 3 HERE

After the workshops, the evaluation data were analysed in relation to the eight 'mother values'. In this analysis, a colour coding system was used, with strongly positive responses ('a lot') coded as green, neutral or slightly positive responses ('a little') as amber, and negative responses ('not at all') as red. The percentage of strongly positive (green) responses was calculated for each indicator, and then aggregated into an overall figure for the value as a whole. In order to help the organisation to work strategically from the evaluation findings, the researchers – through consultation with key staff members within RM – established an indicative target of 75% 'green' (strongly positive

responses) for each value. Results were calculated and compared for each of the four subsets of indicators, corresponding to different spheres of analysis: RM as a whole, the YLSF project, the grassroots projects, and the youth leaders' personal lives.

4.4. Evaluation findings

Both the qualitative and quantitative results from the evaluation process gave a detailed and multidimensional analysis of the presence of values in the four different spheres identified for the project. The use of the indicators as a reference point for defining values in action made it possible to compare results from the survey across spheres, but also between youth leaders and RM staff. Thus, researchers were able to compare results in terms of the 'presence of values', and identified specific indicators that showed areas of good practice or where improvements could be made, within a respectful process characterised by critical and reflexive dialogue.

An important parallel to draw was between responses from youth leaders to values in action in their projects as opposed to with in the YLSF programme. In particular, the analysis showled that the youth did not feel they had enough representation and decision-making power within the YLSF project or felt that the project and its nation-wide activities had established policies of actions for reducing its environmental impact, when this was always a priority in their local projects, and a criteria against which they were evaluated to gain support.

Results from the analysis also allowed researchers to compare responses from RM staff and youth leaders. Two indicators related to the value of Solidarity were identified as strongly enacted by both staff and youth leaders: 'Local projects and other organisations are open to

working together because they respect each other' and 'We consider our work as a form of service'. The analysis also highlighted areas where there were important differences in perception, that may need addressing. For instance, 58% of staff respondents said that conflicts were never resolved through dialogue, whereas youth always resolved conflicts in their projects through dialogue. The impact of this difference was evident in the challenges youth were facing in the project to develop dialogue across different projects and with the national program, and highlighted an opportunity for influencing the organisational culture of YLSF project from a grassroots level rather than from RM. Another indicator linked to the value of Communication showed a marked difference between the youth and RM staff in how the YLSF network is perceieved and the ability of new youth members to join. This revealed a frustration from youth who opening invited new individuals to form part of their local initiatves, but found it difficult to include these new members in the national network.

Finally, the full analysis of the indicators according to the different spheres allow the researchers to identify how the enactment of values potentially travels from one sphere to another, whether in top-down programmes and implmentation of the YLSF project, or through the involvement of youth leaders in the continuous development of the programme. In this sense, the analysis showed that staff managing and coordinating the YLSF programme could learn from the youth leaders terms of enacting indiators linked to the values of Communication and Love. The results also suggested that many values-based practices within the context of local projects led by youth could be integrated into the national YLSF project, as well as influence youth participating in those projects more. This was specifically the case for indicators relating to the values of Responsibility and Confidence. The values that showed strong enactment across all spheres were Solidarity and Knowledge. A summary of scores is given in Table 4.

INSERT TABLE 4 HERE

The evaluation showed a marked difference between YLSF and grassroots projects for the values of *Trust, Responsibility* and *Communication*. In each case, these values were perceived to be much more present in the projects than in the YLSF programme. Many of the indicators showing less presence of the three values were related to issues of ethics and norms within the YSLF programme, and the absence of processes to encourage broad participation in creating a common vision. The indicators assisted in the identification of specific actions that could be taken to improve the situation, such as involving the youth actively in developing norms and ethical codes within YSLF, communicating more effectively to the youth what different YSLF activities are about, and co-developing strategies to reduce and monitor environmental impact.

5. Discussion

5.1 Scaling-up and generalisability of values-based indicators

The RM example illustrates the feasibility of systematic evaluation with values-based indicators as advocated in our earlier work, based on aggregating indicators for specific named values and using a benchmarking approach to identify areas of concern or of particular achievement (Burford et al., 2013a, p. 3049). This constitutes a significant methodological advance in relation to earlier work on values-based indicators, and may pave the way for scaling up values-focused evaluation to the city level, or even beyond. Below we discuss the scaling-up based on the menu of N=166 proto-indicators developed through the ESDinds project. Based on

those findings, we then discuss the broader potential for scaling-up using the conceptual methodology and approach of values-based indicator development as presented in this paper.

The proto-indicators developed in the ESDinds project were adapted to evaluate values-based aspects of a project from the perspective of youth leaders and employees in four different relevant spheres of action. Furthermore, these were integrated into a locally elicited and defined values framework (the values tree), developed through a ground-up qualitative process. This enabled internal comparison of survey results within the organisation and allowed researchers and project stakeholders to gain insights into differing and common values, and how these are enacted and perceived in different ways across the project spheres. The use of values-based indicators as well as integrating qualitative observations allowed for a meaningful comparison of responses and analysis of differences.

It is essential to remember that the evaluation findings do not constitute objective 'measures' of values in any sense, and that it would be inappropriate to directly compare them with findings obtained from a different organisation. They should be understood, not as 'quantitative findings' in their own right, but rather as a vital component of a *mixed methods* evaluation in which qualitative and quantitative data are mutually complementary and draw their meanings from each other. However, the initial starting point of the pooled menu of N=166 proto-indicators appeared to provide strong and useful foundations for this scaling up. If the context of the evaluation was defined as relevant for multiple organisations, the findings could be broadly compared across organisations for specified items in the set/ menu of N=166 proto-indicators.

The RM case study also brings to the forefront important cross-cutting issues, which may be generally useful for others who are examining the role of ethical values and values-based indicators in sustainability efforts. These lessons are relevant not only in relation to the 'menu' of indicators developed through the ESDinds project and applied at RM, but also in respect of a wider process that could enable CSOs (either alone, or ideally in close collaboration with researchers) to identify values and develop indicators within their own contexts.

As social systems vary and are dynamic rather than static (Reid et al., 2006, p. 7), with complex dimensions and interrelationships including cultural, economic, environmental, political, etc., values indicators typically need to be developed in a clearly defined context. The RM evaluation made use of the ESDinds proto-indicators with very little modification, this may be because both RM and the original ESDinds project partners are CSOs with a mandate to promote sustainability and informal ESD. Thus, while the values-based indicators generated through ESDinds may be directly applicable to some CSOs, in other cases their relevance and usefulness may be more limited.

We have been surprised, however, by the widespread acceptance and perceived relevance of the draft indicators in contexts beyond those for which they were initially developed (Burford et al., 2013a). Some aspects of the research design may be important in explaining this, for instance the fact that proto-indicators were often associated with more than one value 'label' at the point of data analysis, even if they were related to only one value at earlier stages of development. Furthermore, the case study methodology provided extremely rich data; this meant that the proto-indicators developed corresponded to diverse and multiple dimensions of each value initially selected. According to current statistics, the web platform created to disseminate the indicators has received over 8300 hits from 138 countries, and its online community of

interest has now reached 162 members, of which 36% (n = 54) have marked and/or customised indicators that they find relevant.   As in 2013, just over a third of the total membership is from the "non-profit, charitable or humanitarian" sector, with a further 31% coming from the private sector (companies or social enterprises) and 18% from academic or educational institutions. The "Other" category includes faith-based organizations, public sector organizations, families, informal community groups and individuals (authors' unpublished data).

In reflecting on the generalizability of the work presented here, it is important to note that the case study described above illustrates only one of several possible responses to the ESDinds indicator set: accepting it in its entirety, or with small modifications, as the foundation of an evaluation. Another possibility, described in earlier work (Harder et al., 2014b), is that user organizations may choose to select only one or a few indicators for evaluation, or to use them as templates or 'prompts' for the creation of alternative indicators that are more in accordance with their needs. In the event that a particular context might be so different that insufficient of the indicators are useful , what the methodology described above developed in the ESDinds project can offer is the indicator development process itself, which may be used to generate entirely new sets of indicators. The process has already been replicated in the rather different context of higher education with an Austrian university, resulting in some indicators that closely resemble those developed during the original ESDinds project and many others that reflect quite different values (Hoover et al., 2015). A key point to understanding these different successful ways of using the approach is that values must first be crystallized into proto-indicators, which must then be localized appropriately, before being measured in any sense, so the interim step of facilitating

that crystallization is very important. We have focused on this in other studies (Harder et al., 2014b; Podger et al., 2013) and will continue to carry out research into these fascinating aspects.

5.2   Limitations of the study

We acknowledge that the assignation of indicators to values is necessarily intersubjective (Burford et al., 2013a, pp. 3044-3046), and that the connections between specific indicators and 'mother values' may be different for other stakeholder groups – possibly even within RM itself. Second, the internal consistency (Cronbach's alpha) of the 'measures' (Cronbach, 1951) was not calculated in this study, although this would have been possible with a larger sample size - which means that each 'value' cannot necessarily be assumed to be a single construct, and outlying indicators may be skewing the results.

The test-retest reliability of the responses was not examined because of time and human resource limitations, rendering it possible that some of them may have reflected the youth leaders' mood on the testing day or their recollections of recent events rather than their 'real' assessment of the indicator. Although the facilitators spent time generating trust, and 'blind' responses were used on personal and politically sensitive topics, social desirability bias (Arnold and Feldman, 1981) might have affected the results (although one might reasonably expect this to be consistent across all the indicators for a given respondent, if indeed it was present).

5. Conclusions

The ESDinds research project has opened ways of accessing, in a tangible way, the values domain which underpins the approaches of many organisations working in sustainability and ESD. Although originally co-developed with only four CSOs, the values-based indicators and the

localisable approach developed with them have also been shown to be useful with a large national environmental program which has several layers of activity. In that case, the approach was used to identify, make tangible, and make rough measures of, values-based dimensions at organisation, project and personal level; especially those previously lacking articulation at the grassroots level, which is something that has often been seen as impossible. In the case of a program in Mexico with national and local levels of work, the approach has revealed variations of strengths of values-in-action in different spheres which allowed the organisation to plan how to strengthen its operationalization of its values-based principles. We would envisage that, with additional work, it might be possible to adapt this approach to develop reliable and internally consistent 'measures' for values as they are defined and understood in local contexts, and thus allowing even more clear comparisons across programs or projects or levels; or changes in time. This, in turn, could permit internal comparisons, e.g. of different projects within the MpMb program, or examining changes in a project over time.

This study has thus highlighted the unexpected transferability of a proto-indicator set developed initially with a very limited sample of organizations, and derived from a small number of only six Values which (in the light of resource constraints) had been selected from a much larger pool. The reasons for this transferability are, as yet, unclear, but it appears that the constituent proto-indicators can be used to locally construct or define other Values. The usefulness of the ESDinds approach for systematic evaluation within a complex, multi-layered context has highlighted the need for further work to look more closely at its origins, and to test the boundaries of its generalizability. A future research agenda might examine, for example, whether this approach is potentially transferable beyond the voluntary sector to other contexts in

which people collaborate to achieve shared goals (e.g. businesses, schools, universities, health care settings, or multi-stakeholder research partnerships). The ability of the approach to facilitate groups to crystallise their values-in-action is a characteristic in demand e.g. for ESD, philanthropic organisations and some businesses, and should thus be studied further.

Lastly, this work has confirmed that values-based achievements can be made tangible and 'measurable' in some sense, even when within large and complex organisations. The original motivation for the foundational work on values-based measures was so that CSOs involved in environmental or sustainable development work could articulate and communicate clearly to their funders and the wider public how the values that they deemed important were demonstrated in their work. A few claimed that these values were key to their success, but the link between values and success could not be clearly shown if the values themselves were intangible. With the approach presented here, it is now feasible in many cases to investigate such links, and we see that as a future line of research that we and others may follow.

References

Abrams, D., 1996. The spell of the sensuous: perception and language in a more-than-human world. Pantheon, New York.

Allport, G.W., Vernon, P.E., Lindzey, G., 1951. A study of values. Houghton Mifflin, Boston, MA.

Anello, E., 2006. Ethical Infrastructure for Good Governance in the Public Pharmeceutical Sector: Working draft for field testing and revision, November 2006. World Health Organization, Department of Medicines Policy and Standards (PSM), Geneva, Switzerland.

Arnold, H.J., Feldman, D.C., 1981. Social desirability response bias in self-report choice situations. Academy of Management Journal 24, 377-385.

Bell, S., Morse, S., 2008. Sustainability Indicators: Measuring the Immeasurable? Earthscan, London, UK.

Bell, S., Morse, S., 2011. Sustainable development indicators: the tyranny of methodology revisited. Consilience: the journal of sustainable development 6, 222-239.

Bond, A.J., Morrison-Saunders, A., 2011. Re-evaluating sustainability assessment: aligning the vision and the practice. . Environmental Impact Assessment Review 31, 1-7.

Bormann, I., 2007. Criteria and indicators as negotiated knowledge and the challenge of transfer. Educational Research for Policy and Practice 6, 1-14.

Bragança, L., Mateus, R., Koukkari, H., 2010. Building sustainability assessment. Sustainability 2, 2010-2023.

Buchanan, R., 1995. Wicked problems in design thinking, in: Margolin, V., Buchanan, R. (Eds.), The Idea of Design. The MIT Press, Cambridge, MA, p. 14.

Burford, G., Hoover, E., Velasco, I., Janouskova, S., Jimenez, A., Piggot, G., Podger, D., Harder, M.K., 2013a. Bringing the 'missing pillar' into Sustainable Development Goals: towards intersubjective values-based indicators. Sustainability 5, 3035-3059.

Burford, G., Velasco, I., Janouskova, S., Zahradnik, M., Hak, T., Podger, D., Piggot, G., Harder, M.K., 2013b. Field trials of a novel toolkit for evaluating 'intangible' values-related dimensions of projects. Evaluation and Program Planning 36, 1-14.

Clugston, R., 2011. Ethical Framework for a Sustainable World: Earth Charter Plus 10 conference and follow up. Journal of Education for Sustainable Development 5, 173-176.

Corcoran, P.B., Vilela, M., Roerink, A., 2005. The Earth Charter in Action: Toward a Sustainable Development. Royal Tropical Institute (KIT), Amsterdam, The Netherlands.

Cousins, J.B., Chouinard, J.A., 2012. Participatory evaluation up close: an integration of research-based knowledge. Information Age Publishing, Inc., Charlotte, NC, USA.

Cousins, J.B., Whitmore, E., 1998. Framing participatory evaluation. New Directions for Evaluation 80, 3-23.

Creswell, J.W., 2013. Qualitative Inquiry and Research Design: Choosing Among Five Approaches, 3rd ed. Sage, Thousand Oaks, CA.

Crishna, B., 2007. Participatory evaluation (I) - sharing lessons from fieldwork in Asia. Child Care Health and Development 33, 217-223.

Crompton, T., Brewer, J., Chilton, P., Kasser, T., 2010. Common Cause: The Case for Working with our Cultural Values. WWF-UK.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16, 239-334.

Dahl, A.L., 2012. Achievements and gaps in indicators for sustainability. Ecological Indicators 17, 14-19.

de Groot, J.I.M., Steg, L., 2007. Value orientations and environmental beliefs in five countries - Validity of an instrument to measure egoistic, altruistic and biospheric value orientations. Journal of Cross-Cultural Psychology 38, 318-332.

De Leo, J., 2012. Quality education for sustainable development: an educator handbook for integrating values, knowledge, skills and quality features of Education for Sustainable Development in schooling. UNESCO APNIEVE Australia Publishing, Adelaide, Australia.

ECI Secretariat, 2010. Earth Charter Initiative Handbook. Earth Charter Initiative, Costa Rica.

ESI, 2005. Environmental Sustainability Index: Benchmarking National Environmental Stewardship. Yale Center for Environmental Law and Policy (YCELP) and Center for International Earth Science Information Network (CIESIN), Columbia University. .

Estrella, M., Gaventa, J., 1997. Who counts reality? Participatory monitoring and evaluation: a literature review. Participation Group, IDS, University of Sussex, Brighton.

European Commission: EUROSTAT, Sustainable development indicators and national accounts.

European Union, 2015. Beyond GDP: measuring progress, true wealth, and the well-being of nations.

Fien, J., Scott, W., Tilbury, D., 2001. Education and conservation: lessons from an evaluation. Environmental Education Research 7, 379-395.

Fraser, E.D.G., Dougill, A.J., Mabee, W.E., Reed, M., McAlpine, P., 2006. Bottom up and top down: Analysis of participatory processes for sustainability indicator identification as a pathway to community empowerment and sustainable environmental management. Journal of Environmental Management 78, 114-127.

Fuad-Luke, A., 2009. Design activism: beautiful strangeness for a sustainable world. Earthscan, London.

Gallopin, G.C., 1997. Indicators and their use: information for decision-making, in: Moldan, B., Bilharzia, S. (Eds.), Sustainability Indicators. Report on the Project on Indicators of Sustainable Development. John Wiley & Sons, Chichester, pp. 13-27.

Global Leaders of Tomorrow Environment Task Force, 2005. Environmental Sustainability Index. Yale Center for Environmental Law & Policy, New Haven, CT.

Grainger, A., 2012. Forest sustainability indicator systems as procedural policy tools in global environmental governance. Global Environmental Change 22, 147-160.

Grundy, S., 1987. Curriculum: product or praxis? Falmer Press, Lewes, UK.

Gudmundsson, H., 2003. The policy use of environmental indicators - learning from evaluation research. Journal of Transdisciplinary Environmental Studies 2, 1-12.

Hak, T., Moldan, B., Dahl, A., 2007. Sustainability indicators: a scientific assessment. SCOPE (Scientific Committee on the Problems of Environment based in Paris), Vol. 67. Island Press, Washington, D.C.

Handy, R., 1970. The measurement of values. W. H. Green, St Louis.

Harder, M.K., Burford, G., Hoover, E., 2014a. From 'sustainable production' to 'production as sustainability': the emergence of values-focused evaluation, Global Research Forum on Sustainable Production and Consumption. Fudan University, Shanghai, Shanghai, China.

Harder, M.K., Velasco, I., Burford, G., Podger, D., Janouskova, S., Piggot, G., Hoover, E., 2014b. Reconceptualizing 'efffectiveness' in environmental projects: can we measure values-related achievements? Journal of Environmental Management 139, 120-134.

Hardi, P., Zdan, T., 1997. Assessing Sustainable Development: Principles in Practice. International Institute of Sustainable Development, Winnipeg: Manitoba.
Hawkes, J., 2001. The fourth pillar of sustainability: culture's essential role in public planning. Common Ground Publishing Pty Ltd in association with the Cultural Development Network (Vic), Victoria.
Hedlund-de Witt, A., 2011. The rising culture and worldview of contemporary spirituality: A sociological study of potentials and pitfalls for sustainable development. Ecological Economics 70, 1057-1065.
Hinkel, J., 2011. ''Indicators of vulnerability and adaptive capacity'': Towards a clarification of the science–policy interface. . Global Environmental Change 21, 198-208.
Hitlin, S., Piliavin, J.A., 2004. Values: Reviving a dormant concept. Annual Review of Sociology 30, 359-393.
Hoover, E., Ribeiro, M.M., Buchebner, J., Burford, G., Lindenthal, T., 2015. Values: a bridge between sustainability and institutional assessment - a case study from BOKU University. International Journal of Sustainability in Higher Education in press.
Hulme, M., 2010. Problems with making and governing global kinds of knowledge. Global Environmental Change 20, 558-564.
Interreligious Statement Towards Rio+20, 2012. Towards Rio+20 and Beyond - A Turning Point in Earth's History. Southern African Faith Communities' Environment Institute, Kalk Bay, South Africa.
Kagan, S., 2011. Art and sustainability: connecting patterns for a culture of complexity. Transcript Verlag, Bielefeld, Germany.
Lather, P., 1986a. Issues of validity in openly ideological research: between a rock and a soft place. Interchange 17, 63-84.
Lather, P., 1986b. Research as praxis. Harvard Educational Review 56, 257-277.
Lele, S., Norgaard, R.B., 1996. Sustainability and the scientist's burden. Conservation Biology 10, 354-365.
Littig, B., Griessler, E., 2005. Social sustainability: a catchword between political pragmatism and social theory. International Journal of Sustainable Development 8, 65-79.
McCool, S., Stankey, G., 2004. Indicators of sustainability: challenges and opportunities at the interface of science and policy. Environmental Management 33, 294-305.
Meadows, D., 1998. Indicators and Information Systems for Sustainable Development: Report to the Balaton Group. The Sustainability Institute, Hartland Four Corners.
Messick, S., 1989. Validity, in: Linn, R.L. (Ed.), Educational measurement, 3rd ed. Macmillan, New York, pp. 13-103.
Othman, M.R., Repke, J.-U., Wozny, G., Huang, Y., 2010. A modular approach to sustainability assessment and decision support in chemical process design. Industrial & Engineering Chemistry Research 49, 7870-7881.
Oxford English Dictionary, 2013. Definition of 'values'.
Palmer, K., Conlin, R., 1997. Sustainable Seattle: The Indicators of Sustainable Community, in: Hardi, Zdan, T. (Eds.), Assessing Sustainable Development: Principles in Practice. International Institute of Sustainable Development, Winnipeg: Manitoba, pp. 117-128.
Peterson, C., Seligman, M.E.P., 2004. Character strengths and virtues: A handbook and classification. Oxford University Press, Oxford.
Podger, D., Velasco, I., Amezcua Luna, C., Burford, G., Harder, M.K., 2013. Can values be measured? Significant contributions from a small civil society organisation through action

research evaluation. Action Research e-publication ahead of print DOI:10.1177/1476750312467833.

Pope, J., Annandale, D., Morrison-Saunders, A., 2004. Conceptualising sustainability assessment. Environmental Impact Assessment Review 24, 595-616.

Rametsteiner, E., Puelzl, H., Alkan-Olsson, J., Frederiksen, P., 2011. Sustainability indicator development: Science or political negotiation? Ecological Indicators 11, 61-70.

Reed, M., Fraser, E.D.G., Morse, S., Dougill, A.J., 2005. Integrating Methods for Developing Sustainability Indicators to Facilitate Learning and Action. Ecology & Society 10, 1-6.

Reed, M.S., Dougill, A.J., 2002. Participatory selection process for indicators of rangeland condition in the Kalahari. Geographical Journal 168, 224-234.

Reed, M.S., Dougill, A.J., Baker, T.R., 2008. Participatory indicator development: What can ecologists and local communities learn from each other? Ecological Applications 18, 1253-1269.

Reed, M.S., Fraser, E.D.G., Dougill, A.J., 2006. An adaptive learning process for developing and applying sustainability indicators with local communities. Ecological Economics 59, 406-418.

Reid, A., Nikel, J., Scott, W., 2006. Indicators for Education for Sustainable Development: a report on perspectives, challenges and progress. Anglo-German Foundation for the Study of Industrial Society, London, UK.

RFSC, 2013. Reference framework for European sustainable cities: a toolkit for the integrated approach.

Robottom, I., Hart, P., 1993. Research in Environmental Education: Engaging the Debate. Deakin University, Geelong.

Rokeach, M., 1973. The Nature of Human Values. The Free Press, New York.

Rokeach, M., 1979. Understanding human values. Free Press, New York.

Rosenström, U., 2006. Exploring the policy use of sustainable development indicators: interviews with Finnish politicians. The Journal of Transdisciplinary Environmental Studies 5, 1-13.

Schlater, J.D., Sontag, M., 1994. Toward the measurement of human values. Family and Consumer Science Research Journal 23, 4-25.

Schwartz, S.H., 1992. Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. Advances in Experimental Social Psychology 25, 1-65.

Schwartz, S.H., 1994. Are there universal aspects in the structure and content of human values? Journal of Social Issues 50, 19-45.

Schwartz, S.H., 2007. Basic Human Values: Theory, Methods, and Applications. The Hebrew University of Jerusalem.

Schwartz, S.H., Bilsky, W., 1987. Toward a universal psychological structure of human values. Journal of Personality and Social Psychology 53, 550-562.

Schwartz, S.H., Bilsky, W., 1990. Toward a theory of the universal content and structure of values - extensions and cross-cultural replications. Journal of Personality and Social Psychology 58, 878-891.

Schwartz, S.H., Lehmann, A., Roccas, S., 1999. Multimethod probes of basic human values, in: Adamopoulos, J., Kashima, Y. (Eds.), Social psychology and culture context: Essays in honor of Harry C. Triandis. Sage, Newbury Park, CA, pp. 107–123.

Schwartz, S.H., Melech, G., Lehmann, A., Burgess, S., Harris, M., Owens, V., 2001. Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. Journal of Cross-Cultural Psychology 32.

Smolko, R., Strange, C., Venetoulis, J., 2006. The Community Indicators Handbook. Redefining Progress, Oakland, CA.

Sollart, K., 2005. Framework on Indicators for Education for Sustainable Development: Some conceptual thoughts. . Netherlands Environmental Assessment Agency (MNP). .

Spangenberg, J.H., 2002. Institutional sustainability indicators: an analysis of the institutions in Agenda 21 and a draft set of indicators for monitoring their effectivity. Sustainable Development 10, 103-115.

Spangenberg, J.H., Pfahl, S., Deller, K., 2002. Towards indicators for institutional sustainability: lessons from an analysis of Agenda 21. Ecological Indicators 2, 61-77.

Springett, J., 2003. Issues in participatory evaluation, in: Wallerstein, M.M.N. (Ed.), Community-Based Participatory Research for Health. Jossey-Bass, San Francisco, CA.

Stake, R.E., 2000. Case studies, in: Denzin, N.K., Lincoln, Y.S. (Eds.), Handbook of qualitative research, 2nd ed. Sage Publications, Thousand Oaks, CA, pp. 134-164.

Sterling, S., 2001. Sustainable education: revisioning learning and change. Green Books, Cambridge, UK.

Sterling, S., 2007. Riding the storm: towards a connective cultural consciousness, in: Wals, A.E.J. (Ed.), Social learning toward a more sustainable world: principles, perspectives, and praxis. Waginengen Academic Publishers, Waginengen, NL.

Stern, P.C., Dietz, T., 1994. The value basis of environmental concern. Journal of Social Issues 50, 65-84.

Stiglitz, J.E., Sen, A., Fitoussi, J.-P., 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress.

Talamo, A., Pozzi, S., 2011. The tension between dialogicality and interobjectivity in cooperative activities. Culture & Psychology 17, 302-318.

Taylor, S.J., Bogdan, R., 1998. Introduction to qualitative research methods: A guidebook and resource (3rd ed.). John Wiley & Sons Inc, Hoboken, NJ, US.

Tilbury, D., Cooke, K., 2005a. A National Review of Environmental Education and its Contribution to Sustainability in Australia: Frameworks for Sustainability, Volume 1. Australian Government Department of the Environment and Heritage and the Australian Research Institute in Education for Sustainability, Canberra.

Tilbury, D., Janousek, S., 2006. Development of a National Approach to Monitoring, Assessment and Reporting on the Decade of Education for Sustainable Development – Stage 1: Identification of National Indicators. Report to the Australian Government Department of the Environment and Heritage. Australian Research Institute in Education for Sustainability.

Tilbury, D., Wortman, D., 2004. Engaging people in sustainability. IUCN Commission on Education and Communication, Gland, Switzerland, & Cambridge, UK.

UN Division for Sustainable Development. Department of Policy Co-ordination and Sustainable Development, 1995. Work Programme on Indicators of Sustainable Development, Document UN/E/CN.17/1995/18. United Nations, New York.

UN Division for Sustainable Development. Department of Policy Co-ordination and Sustainable Development, 1996. Indicators of sustainable development, framework and methodologies. United Nations, New York.

UN Division for Sustainable Development. UN Department of Economic and Social Affairs, 2000. Indicators of sustainable Development, framework and core set. United Nations, New York.

UNESCO, 1997. Educating for a sustainable future: a transdisciplinary vision for concerted action. UNESCO, Paris.

UNESCO, 2007. Asia-Pacific Guidelines for the Development of National ESD Indicators. UNESCO Asia and Pacific Regional Bureau for Education, Bangkok.

UNESCO, 2013. Culture in the post-2015 sustainable development agenda: why culture is key to sustainable development. Background paper to the 2013 Hangzhou Congress on Culture and Development, 14-17 May 2013, Hangzhou, People's Republic of China.

United Nations General Assembly, 2000. United Nations Millennium Declaration. United Nations General Assembly, New York.

Walker, S., 2006. Sustainable by Design: Explorations in Theory and Practice. Earthscan / James and James, London, UK.

Wang, S., 2002. Wicked problems and metaforestry: is the era of management over? . Forestry Chronicle 78, 505-510.

Woodley, E., Crowley, E., Dookie, C., Carmen, E., 2006. Cultural indicators of Indigenous Peoples' food and agro-ecological systems. Food and Agriculture Organization of the United Nations, Rome.

Jan – Nov 2009	**Background research into existing values-based indicators - within academic literature and on-ground actions of CSOs**
	Develop a pool of values and possible indicators (first set)
Nov 2009 – May 2010	**Develop and trial first set of indicators with ESDinds CSO partner projects and develop assessment tools**
	Refine and add to first set of indicators (second set)
Jun – Nov 2010	**Develop and trial second set of indicators with ESDinds CSO partner projects and with 50-80 additional CSOs**
	Refine second set of indicators and process of values identification and indicator development
Dec 2010 – Jan 2011	**Disseminate results to a broad audience through conference, workshops, website and publications**

Fig 1. Outline of the key stages in the ESDinds project.

Core Group Meeting
Establish Research Focus and Expectations for First Iteration

Background Research
Review of literature to identify existing values-based indicators and conversations with CSOs to learn about the organisations and identify espoused values frameworks

Research Meetings
Define Variables, Design Case Study Research (tailored to each CSO), Develop Interview Guide

Value Identification and Indicator Development
First Iteration

Data Collection with CSO Case Studies Research Participants
Key informant interviews, group interviews, documents (vision statements, website material, project reports, workshops), participant and non-participant observation.

Data Analysis
Develop initial code book; design Atlas.Ti database; carry out content analysis for value identification for each CSO; describe initial indicator ideas; and continuously update code book and value meanings

Value Set Selection
Review comprehensive set of values for each CSO and identify set of five primary values

Indicator Development
Convert indicator ideas for these values into pilot indicators; Categorise indicators into themes and write up value descriptions.

Suggested Measurement Methods and Assessment Tools
Review academic literature for each of the values to refine value meanings, and identify current measurement methods and assessment tools. Include those identified for possible testing in the second iteration.

First CSO Indicator Evaluation
Review value set, value themes and potential indicators with CSO research participants and partners for meaningful and useful indicators

Second CSO Indicator Evaluation
Technical review by CSO partners to finalise value set and prioritise meaningful and useful indicators for the Second Iteration

Fig 2. Outline of the key steps to identifying values and developing indicators in the First Iteration.

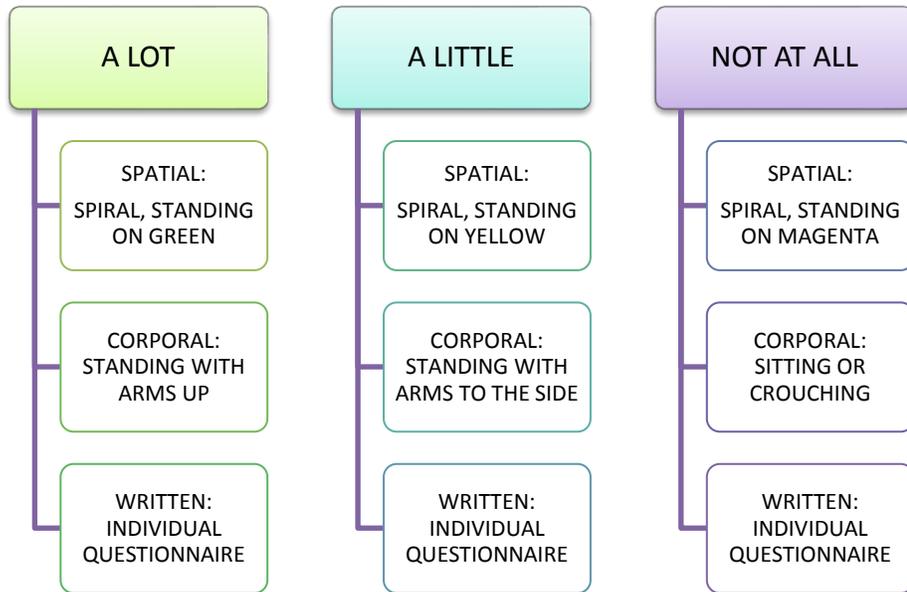| A LOT	A LITTLE	NOT AT ALL
SPATIAL: SPIRAL, STANDING ON GREEN	SPATIAL: SPIRAL, STANDING ON YELLOW	SPATIAL: SPIRAL, STANDING ON MAGENTA
CORPORAL: STANDING WITH ARMS UP	CORPORAL: STANDING WITH ARMS TO THE SIDE	CORPORAL: SITTING OR CROUCHING
WRITTEN: INDIVIDUAL QUESTIONNAIRE	WRITTEN: INDIVIDUAL QUESTIONNAIRE	WRITTEN: INDIVIDUAL QUESTIONNAIRE

Fig. 3. Overview of the data collection process, showing the different representations (spatial, corporal, or written) for each of the three possible responses to the survey questions

Fig. 4. Collective development of the "values tree", visual representation of the YLSF's values system. The tree was called Huhub / Amor / Love (photography of the tree in development).

Table 1

CSO values and themes, the number of indicators, and a sample of pre-test indicators provided for illustrative purposes (examples are those indicators identified as meaningful and useful by a majority of the CSOs at the Core Group Meeting in November 2009).

Values	# of Indicators	Themes	Illustrative Examples
Trust/ Trustworthiness	57	Perception and presence of trust; Atmosphere of trust; Building and maintaining trust; Living ethical principles	The organisation is transparent about the process and outcomes of decision-making, openly sharing information with employees; Individuals, colleagues, organisations, partners are perceived to be trustworthy, truthful, honest, transparent, respectful and practice integrity in their interactions with others
Integrity	56	Application of ethical values; Moral conduct; Authenticity; Consistency; Reliability; Ecological integrity [a]	Ethical values and principles are used by individuals/team/organisation in guiding decision-making and activities; Individuals/team/ organisation/partners follow through on their commitments
Justice	55	Social justice; Economic justice; Participative democratic decision-making; Environmental justice	People feel they are treated equitably and with fairness; People/organisations participate actively in making decisions about issues that affect their lives
Empowerment	65	Developing and awakening capacity; Providing encouragement and support; Providing purpose or reason; Responsibility, ownership and	People/partners are encouraged to express their opinion; People /teams/ organisations are given autonomy and trust to fulfil responsibilities, at the same time receiving

		autonomy; Manifestation of empowerment	encouragement and support (*)
Unity in Diversity	55	Presence of diversity; Appreciation of Diversity; Use of diversity; Spirit of collaboration	People feel they create something better/greater as a group than on their own; Teams include members with different characteristics (e.g. gender, culture, age and other aspects of individual difference such as personality)
Respect and Care for the Community of Life	76	Recognition of a caring role; Social and environmental sustainability; Respectful treatment of people and the community of life; Ecological literacy; Celebration of the environment and community of life; Interrelatedness between environment and other elements of society; Raising awareness of environmental sustainability and the community of life; Acknowledgement of worth and contribution; Appreciation of identity and individuality; Self-respect; Openness and inclusion	People understand the complexity of natural systems; The organisation's activities or events connect participants emotionally to the community of life
Total	364		

[a] IND-SD did not focus on specific indicators associated with environmental impact as there are already well-established indicators that CSOs can refer to.

Table 2

RM localised interpretation of different parts of the 'tree of values' metaphor

Part of `Tree'	Interpretation
Roots	Pre-existing values. Motivation, compassion, kindness, creativity and identity
Trunk	Values acquired by decision. Companionship, honesty, tolerance, empathy, friendship, coherence, responsibility, loyalty, truthfulness, empowerment, integrity, equity
Branches	Values learned through learning and training. Knowledge, patience, transparency
Fruits	Values acquired through strength and will. Transcendence, fulfilment, autonomy, empowerment
Leaves	Values visible from outside. Harmony, admiration, good example
Sap	Essential values. Love, communication, trust, solidarity

Table 3

Examples of indicators used in the RM evaluation, with their associated 'mother values'

#	INDICATOR	SPHERE	ASSOCIATED VALUE
14	Within the project, decision-making processes are ethical	Projects	Autonomy
20	People actively participate in developing the YLSF programme's code of ethics	YLSF	Responsibility
25	Members of RM involved in YLSF feel there is an appropriate flow of information	RM	Communication
28	Conscious actions are taken to improve the way in which YLSF members are treated	YLSF	Responsibility
53	In the project, people can express their opinions	Projects	Communication
122	Remuneration/payment policies are perceived to be fair by all involved in the YLSF programme	YLSF	Responsibility
125	Confidential, truth-seeking and non-judgemental channels exist for teams/individuals seeking guidance on the applications of ethics, denouncing and examining violations of ethics	YLSF	Communication
127	Performance goals are measured	YLSF	Responsibility
130	The financial integrity of RM is communicated internally and externally	RM	Communication
146	Your project acts to protect the environment, without waiting for governments or others to act	Projects	Responsibility
149	YLSF has a policy to purchase all or part of its energy from renewable sources	YLSF	Autonomy
152	People know how many projects and activities take place towards the goal of environmental sustainability	YLSF	Responsibility
155	Members of the project share their knowledge on how to protect and restore the natural environment with others	Projects	Solidarity
156	Education processes in YLSF are delivered to develop capacities and knowledge to act according to principles of sustainable development	YLSF	Knowledge
158	In your projects, you create long term pledges to protect the natural environment.	Projects	Responsibility

Table 4: Summary of scores for all eight `mother values' across the four spheres of evaluation. Results are shown as a percentage of strongly positive responses across the full dataset for these indicators, in relation to a benchmark figure of 75%,. Note that not all of the indicators were deemed to be applicable at the organisational level, so there are some mother values with no data.

	RM as an organization	YLSF project	Grassroots projects	Youth leaders' personal lives
Love	N/A	58	71	90
Trust	N/A	45	98	70
Identity	N/A	70	71	95
Solidarity	N/A	80	87	88
Communication	13	56	77	80
Autonomy	54	56	67	57
Knowledge	N/A	72	75	68
Responsibility	29	52	75	42