Field trials of a novel toolkit for evaluating 'intangible' values-related dimensions of projects☆

Gemma Burford [a], Ismael Velasco [a], Svatava Janoušková [b], Martin Zahradnik [b], Tomas Hak [b], Dimity Podger [a], Georgia Piggot [a], Marie K. Harder [a,c,*]

[a] Sustainable Development Coordination Unit, University of Brighton, Cockcroft Building, Lewes Road, Brighton BN2 4GJ, UK
[b] Charles University Environment Center, José Martího 2/407, 16200 Prague 6, Prague, Czech Republic
[c] Department of Environmental Sciences and Engineering, Fudan University, Shanghai 200433, PR China

ABSTRACT

A novel toolkit has been developed, using an original approach to develop its components, for the purpose of evaluating 'soft' outcomes and processes that have previously been generally considered 'intangible': those which are specifically values based. This represents a step-wise, significant, change in provision for the assessment of values-based achievements that are of absolutely key importance to most civil society organisations (CSOs) and values-based businesses, and fills a known gap in evaluation practice. In this paper, we demonstrate the significance and rigour of the toolkit by presenting an evaluation of it in three diverse scenarios where different CSOs use it to co-evaluate locally relevant outcomes and processes to obtain results which are both meaningful to them and potentially comparable across organisations. A key strength of the toolkit is its original use of a prior generated, peer-elicited 'menu' of values-based indicators which provides a framework for user CSOs to localise. Principles of participatory, process-based and utilisation-focused evaluation are embedded in this toolkit and shown to be critical to its success, achieving high face-validity and wide applicability. The emerging contribution of this next-generation evaluation tool to other fields, such as environmental values, development and environmental sustainable development, shared values, business, education and organisational change is outlined.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The notion of value and its assessment lies at the very heart of evaluation. Behind any frontline evaluation activity is an implicit set of values that have determined which indicators are worth assessing, and the standards for them. Efforts are usually made to ensure that any measurements or assessments carried out are faithful representations of the indicators, but this paper addresses the deeper and significant question of whether the indicators properly represent the reference values in the first place.

Although this is an important question in any evaluation and there are current debates in ecological economics, environmental education and environmental values (Heimlich, 2010; Spash,

2009; Turner et al., 2003), about the need for indicators to be more representative of the underlying values present in specific contexts, it has become a critical and pressing question in the domain of civil society organisations (CSOs),[1] who insist that key outcomes of their work are not only poorly represented but not being counted. This is because existing evaluation frameworks are not designed to capture dimensions relating to 'higher' ethical/spiritual values (such as equality or empowerment). Instead, they are based on frameworks centred on the views of funding donors, businesses or academia, and designed with some rigidity which is good for generalisability for the evaluators but not local face validity.

Ebrahim (2003) reports on the tendency of donor appraisals to focus on easily measurable and quantifiable outputs of CSO work, such as the number of schools built or hectares of land irrigated, whilst neglecting more ambiguous and less tangible changes in social and political processes. Bilateral donor agencies often rely on

[1] The term 'civil society organizations' is used for convenience to encompass both community-based organisations (CBOs) and non-governmental organisations (NGOs), varying in size and formality.

the Logical Framework Approach,[2] which has been criticised by Edwards and Hulme (1996) for overemphasising short-term, quantitative targets. There is widespread concern among NGOs that donors accordingly reward discrete, product-based approaches to development, while penalising innovative process-based approaches that prioritise sustainable behaviour change (Ebrahim, 2003; Riddel, 1999). Bamberger's (2000:96) observation that "donors' information priorities and evaluation methodologies continue to exert considerable influence on how evaluation is practiced and used", while initially made in the specific context of evaluating international development programmes, still remains highly relevant to diverse CSOs worldwide. As a result, CSOs may be pressured to develop onerous monitoring and evaluation systems that satisfy donor demands for information, yet detract from core project activities (Crishna, 2007) and have little relevance to internal decision-making and learning (Ebrahim, 2002, 2003).

There is evidence that donors themselves are similarly frustrated by difficulties in evaluating those things that really matter to them. The Ford Foundation's website states that "the philanthropic sector struggles with evaluating itself, grading itself: understanding whether its work is making a difference" (Ford Foundation, 2011). The CEO of the Vancouver Foundation (Wightman, 2010) criticises the trend towards using standardised metrics to 'quantify' the impact of charitable donations, and remarks that the benefits of certain types of philanthropy are abstract, difficult to measure, and manifested over several years, yet they are no less real or important.

This need for improved evaluation tools for the values-based work of civil society organisations (CSOs) was the main driver for a two-year European-funded research programme to develop values-based indicators, outlined elsewhere (Podger et al., 2010). It involved a consortium[3] of two university research groups and four CSOs broadly concerned with values and sustainable development. The CSO consortium members felt that intangible, values-related impacts of their sustainability work, which they could 'feel instinctively' as key to the success of their work but could not define concisely or illustrate with rigorous data, were omitted from evaluations and hence overlooked by donors because they were perceived to be impossible to measure. Some recognised that, in the current financial climate, their very survival might depend on their ability to convince donors of such wider achievements and potential. Others felt the under-appreciation of values-based work was generally a reason that 'Gross Domestic Product' (GDP) economics was failing, and saw critical importance in articulating their own core values in terms of measurable indicators and outcome data.

It is unsurprising that values-based aspects of CSO work have been considered 'intangible' and eluded evaluation for so long, because CSOs themselves may often be unable to define these aspects in terms familiar to evaluators. For example, many CSOs do not distinguish between outcomes, processes and stakeholder perspectives, or between formative and summative evaluations. They may also consider several overlapping outcomes or processes as 'dynamic clusters' rather than delineating them cleanly, which makes targeting for specific evaluation very difficult. Conversely, CSOs are usually clear about whether a project was 'good' or 'bad' according to their own ethos, even if they have difficulty articulating the reasons.

Contemporary evaluation literature describes several interrelated theoretical approaches and practical strategies that can be drawn on to develop an evaluation framework for such intangible aspects of CSO work. Firstly, it is essential to give due consideration

to process evaluation (e.g. Ellis & Hogard, 2006; Hogard, 2008) rather than focusing exclusively on outcome evaluation. Secondly, principles of participatory evaluation (e.g. Crishna, 2007; Springett, 2003) are indispensable, as external evaluators would be unlikely to understand the CSO perspective sufficiently to identify specific intangibles for evaluation. Thirdly, utilisation-focused evaluation, involving an iterative and flexible approach based on reflection and 'feedback loops' (Flowers, 2010), is necessary to ensure that an evaluation will ultimately be relevant and useful to the CSO stakeholders and bring about improvements their work. The growing literature on process use (e.g. Crohn & Birnbaum, 2010; Forss, Rebien, and Carlsson, 2002; Holte-McKenzie, Forde & Theobald, 2006; Patton, 1998) emphasizes that the utility of evaluation lies not only in the findings themselves, but also in processes that cause CSO staff to engage in systematic reflection on their projects and sometimes lead to changes in perspective. The studies in this paper indicate that such transformation seems to naturally occur when effective participatory, utilisation-focused process evaluation is executed.

A very important practical strategy for the measurement of intangibles is the combination of multiple qualitative and quantitative methods of data collection (Bouffard, Taxman, & Silverman, 2003; Odendaal, Marais, Munro, & van Niekerk, 2006). Specifically, participatory learning and appraisal (PLA) techniques such as voting, scoring and ranking were found to be useful and have been shown to reduce tensions between external evaluators and CSO staff (Pretty, Guijt, Thompson & Scoones, 1995).

In our own work (outlined by Podger et al., 2010) we have drawn on these evaluation approaches and strategies to co-design a toolkit that can be used by diverse CSOs in assessing previously intangible, values-related aspects of their projects and activities. This toolkit, which we have named WeValue, consists of a reference set of 166 values-based indicators, together with guidance on their application. The indicators are presented in a 'menu' format which greatly facilitates user CSOs to identify and crystallise what intangibles are important to them, and those they wish to evaluate (see Box 1 below). A second 'menu' is provided of assessment methods for potential use, accompanied by detailed guidance notes. In this paper, we report on field trials of the WeValue toolkit with three CSOs, and then discuss their significance within the broader discipline of monitoring and evaluation.

2. Evaluation setting

The first phase of the ESDinds research project entailed a rigorous consultative process aimed at identifying the ethical values that tend to promote 'success' in projects (according to the definitions

---

Box 1. Sample indicators from reference list (WeValue, 2012)

- Everyone has their place in the team
- Women feel that they are given equal opportunities to participate in decision-making processes
- People share their skills and abilities freely with one another, regardless of nationality, ethnic origin, skin colour, gender, sexual orientation, creed or religion
- Entity's activities or events connect participants emotionally to the community of life
- Mistakes are understood as opportunities to learn and improve
- People do not back-bite about others within the entity
- People have a sense of power that they can effect change
- Financial integrity is communicated internally or externally
- Action is consciously taken to contribute to a greater respect for nature
- The environment and community of life is celebrated

of the consortium's own CSOs) and iteratively co-developing and peer-eliciting generalised indicators for those values. From this large set, consortium members selected a subset focusing on clusters of values of high priority to them: Empowerment, Integrity, Justice, Trust, Unity in Diversity, and Care & Respect for the Community of Life. The prototype toolkit consisted of a list of 177 indicators regarded as enactments of these values in processes and outcomes (which, for convenience, we will term the 'WeValue Indicators'), together with guidelines on their application.

The second phase of the ESDinds research project focused on investigating whether the proposed toolkit could be applied by real CSOs, in a manner that was useful and meaningful to CSO stakeholders. A core requirement of this phase was that any assessment methods developed for the indicators had to be feasible within the respective CSO contexts; the indicators would be useless otherwise. Several field studies were carried out for this purpose, in a very diverse sample of organisations and environments. The earliest field studies showed that some indicators needed adapting, giving a final set of 166. Some also caused shifts in understanding, prompting changes in the processes used in applying the prototype toolkit. In subsequent field studies, however, no further changes of process were necessary, and it was possible to evaluate the feasibility and usefulness of the toolkit. That evaluation forms the content of this paper.

3. Evaluation design

3.1. Overview

The WeValue toolkit can be used by organisations to assess values-related processes and outcomes by having CSO consortium members collectively selecting relevant values-based indicators to use from the reference list of 166 items. The reference list contains a variety of indicators contributed by real CSOs, which stimulate the user CSOs to 'crystallize' what their intangible shared values look like when enacted, e.g. when embedded within organisational processes, systems, practices and behaviours. The toolkit then provides detailed guidance on the next step, namely the development of locally appropriate assessment tools for the CSOs' chosen indicators.

The WeValue toolkit thus had three separate elements to be evaluated – the indicators, the potential assessment tools and the process involved. The aim was to determine the overall usefulness of the entire toolkit to a wider sample of CSOs beyond the initial partners involved in Phase 1. (Note that we are thus evaluating an evaluation toolkit used for three local evaluations: in order to avoid confusion, we have used the terms 'toolkit evaluation' and 'local evaluation(s)' throughout the paper to distinguish these different levels.) While the WeValue toolkit can, in principle, be used through an online platform, it is anticipated that most CSOs will require the personal involvement of an external evaluator. Thus, in the toolkit evaluation described here, it is not anomalous that external researchers worked alongside CSO staff.

3.2. Aims and objectives

The aim of the toolkit evaluation was to determine whether the *WeValue* toolkit can assess values-related processes and outcomes in a range of CSOs. Specific objectives were:

1. To determine whether WeValue Indicators are perceived by participating CSOs as relevant to their work[4];

2. To determine whether participating CSOs find the Indicators useful in their own local evaluations
3. To explore whether the Indicators provided were capable of being assessed in real CSO settings, i.e. whether appropriate assessment tools could actually be developed and applied to generate useful data against those indicators.
4. To determine whether the resulting data had adequate face validity, in the sense of being relevant, meaningful and useful to the CSOs; and
5. To identify whether any major attitudinal changes occurred for the CSOs due to the application of *the WeValue tool.*

3.3. Method

Field trials of the WeValue toolkit were carried out in three organisations, chosen for diversity (see Section 3.4). For each one, visiting researchers from the universities worked alongside CSO staff to apply the toolkit in local values-based project or process evaluations of specific interest to them. The objectives above were investigated in each case, but from the field CSO's viewpoint the main activity was the local evaluation of their individually posed question, using the WeValue toolkit. The methodology of the toolkit contains the following five participatory processes, which were carried out in each case:

(a) Determination of which WeValue indicators in the reference list were generally relevant to the CSO;
(b) Localisation of indicator wording for the specific CSO context, and selection of a shortlist that the CSO would like to assess in its specific local evaluation;
(c) Selection of appropriate assessment methodologies from the WeValue 'menu' (Table 1) and development of related context-specific assessment tools congruent with existing CSO activities;
(d) Local data collection, usually led by the CSO staff; and
(e) Local data analysis, and CSO reflection, to examine whether meaningful conclusions could be drawn from data developed with the WeValue toolkit.

For each CSO, the processes listed as (a) and (e) above were led by the project director or coordinator. This 'key informant' usually involved further CSO participants for all of the other activities, although the WeValue toolkit purposely did not specify participation criteria.

In preparation for meta-analysis across the three studies, the researchers took detailed notes of which indicators were found relevant and which chosen for 'measurement'; how indicators were ultimately localised, which assessment tools were developed, and through which processes; what difficulties arose in applying the locally developed assessment tools; what were the main findings of the local assessment; and what was the perceived usefulness to the CSO of both the findings and the evaluation process. Recognising that some of the benefits from the field trials might not be immediately apparent, the researchers conducted follow-up semi-structured interviews with each key informant three to five months after the trials.

3.4. The choice of the three different study sites and environments

The consortium members, three of which were international 'umbrella bodies' with diverse affiliates, were asked to identify CSOs able and willing to test the toolkit within the time and resource constraints. Nominations could include projects in any country, of any size, involving any age or cultural group of people, of faith or not, locally or centrally managed, and independent or closely tied to a larger organisation. Three projects were ultimately chosen to maximise diversity. The characteristics of each project are summarised in Table 2 below.

---

[4] Implicit is the understanding that the indicators are, by origin, related to values-based outcomes and processes. However, the CSOs are not expected to make direct links between the indicators and their own espoused values from the start; the XYZ process allows them to proceed without that.

Table 1
Excerpts from the 'menu' of assessment methods presented to participants during field visits.

Method	Description
Survey	Gathering of information through a form containing a set of questions, especially one addressed to a statistically significant number of subjects considered to be representative of a whole.
Interview	A conversation with an individual in which the researcher asks questions in a systematic way. Can be structured (a specific set of questions asked in the same words and in the same order), semi-structured (a specific set of questions that are used as prompts for a broader discussion), and unstructured (a discussion with no pre-planned questions or order, naturally emerging from the conversation).
Focus group	A discussion by a small group of people (typically 6–12 individuals) of selected topics of interest in informal or formal settings. The focus group discussion is typically directed by a facilitator who guides the discussion in order to obtain the group's opinions about or reactions to specific themes or issues.
Observation	The systematic observation of an interaction, process, community or group. This can be done through observation alone or by both observing and participating (participant observation), to varying degrees, in the group's daily activities. Observers make careful notes about what they see, recording all accounts and observations as field notes in a field notebook. This can be structured (looking out for and/or rating a specific list of items), semi-structured (looking out for a broad set of themes) or unstructured (deriving any categories or questions from the observation itself), and can be carried out by individuals or by groups.
Document analysis	The systematic search for information, evidence or insight about a research question in documents directly or indirectly related to, and/or produced by, the research subjects.
Indirect measures	The systematic gathering of information in a way that does not interact directly with the participants.
Forum theatre	In this process the actors or audience members could stop a performance, often a short scene in which a character was being oppressed in some way. The audience would suggest different actions for the actors to carry out on-stage in an attempt to change the outcome of what they were seeing.

As Table 2 shows, the local evaluations conducted in Mexico and Sierra Leone were focused on values-related outcomes: empowerment, inclusiveness and non-discrimination. The local evaluation in Germany, conversely, was aimed at understanding the extent to which PT's espoused core values (consultative dialogue and decision-making, encouragement, unity in diversity, service orientation, and supporting individual initiatives) were actually enacted in project implementation processes. To achieve this, PT staff had first received assistance in clarifying what the visible manifestations of their core values would look like, as part of the first phase of the ESDinds research project.

4. Data analysis

Within each of the individual field studies, local data analysis was undertaken jointly by all co-evaluators. This entailed examining the collated local data in the light of the specific CSO evaluation question, which related to project outcomes in two of the evaluations and implementation processes in the third. Wherever a CSO wanted a rigorous test of a specific indicator, several lines of evidence deriving from different assessment tools were triangulated. However, this level of rigour was not always required, nor seen to be necessary, as the prime driver was utilisation for the CSO.

Across the field studies, the overall evaluation of the WeValue toolkit was undertaken via multiple and iterative discussions

between the university researchers and also the entire project consortium.

5. Results

5.1. PIMAUG, Guanajuato University

A researcher from the University of Guanajuato, Mexico [a project consortium member] worked alongside the PIMAUG coordinator and two peer educators as co-evaluators, but the peer educators took the lead in choosing ten indicators from the reference list and localising them to fit the specific context of the Earth Charter workshops. Doing this helped the PIMAUG representatives to clarify exactly what they felt would be useful for their local evaluation question – the extent to which their Earth Charter workshops contributed towards an increased sense of empowerment among participants. For example, indicator #113 from the reference list was chosen: "People feel a sense of power to effect change", and modified to become 113a, "Participants feel a sense of power to effect a positive change", and 113b, "Participants feel that their sense of power to effect a positive change has increased after doing the workshop". The final set of customised indicators represented the consensus within the PIMAUG team about what needed to be evaluated to answer their CSO question, and can be seen in full in Table 3.

The co-evaluators then reflected carefully on the menu of localisable assessment methods (Table 1) to develop assessment tools that were congruent with the pre-defined workshop programme. This resulted in the following being planned: (i) a spatial/corporal survey with follow-up focus group discussion; (ii) semi-structured non-participant observation of workshop exercises; (iii) semi-structured key informant interviews; and (iv) word elicitation using a 'What/Why' grid.

The CSO workshop was then run as normal by two peer educators. A semi-structured non-participant observation was carried out concurrently by two independent observers, one of them the visiting university researcher and the other the PIMAUG coordinator. At the end of the normal workshop programme, but before the participants dispersed, the co-evaluators then led four short spatial/corporal surveys, a 15-min focus group discussion, and the word elicitation task, as detailed in Table 3. Afterwards, semi-structured interviews were conducted with the peer educators and project coordinator as key informants.

Together, the team of four co-evaluators analysed the data set. The overall conclusion was that the workshop was very effective in increasing participants' sense of empowerment, particularly among those whose baseline was lower (i.e. those who started the workshop feeling that they had 'some' power to effect positive change, rather than 'a lot'). Overall, approximately two thirds felt that their sense of power to effect positive change had increased, one third felt it had stayed the same, and none felt it had decreased. However, about a quarter of respondents said they felt that participants had 'sometimes' just agreed with the ideas of others rather than identifying their own responses to issues, which was thought to suggest that the group dynamic did not achieve full trust, and the co-evaluators felt that longer workshops might obtain better results.

An even more valuable insight, which emerged from both the focus group discussion and the word elicitation task, was that the key aspect of the workshop that increased participants' sense of empowerment was being in touch with like-minded individuals (in contrast to their feeling of relative isolation in such goals and aspirations prior to the workshop). Furthermore, those who had reported feeling only 'some' or 'a little' power to effect positive change had explained that they lacked clear ideas to translate their motivation into action, and were conscious that much well-meant

Table 2
Overview of the characteristics of organisations chosen for field studies.

Name of organisation	Relationship to project consortium	Geographical and cultural setting	Type of organisation	Local evaluation needs
People's Theater	Full member of the consortium	Offenbach, Germany: youth volunteers aged 18–25	Small, independent local CSO in which youth volunteers are trained to run interactive drama workshops in schools, helping the pupils to explore social responsibility and non-violent conflict resolution. The volunteers work together as teams for a full year to rehearse performances and develop workshops.	Evaluating the extent to which PT's publicly espoused core values were enacted in the interpersonal behaviour of youth volunteers on a day-to-day basis, and in organisational processes such as decision-making and communication flow.
PIMAUG (Environmental Institutional Programme of the University of Guanajuato)	Affiliate of the Earth Charter Initiative, a consortium member	Guanajuato, Mexico: university students	Cross-faculty environmental initiative within a university, structured around six strategic areas: (i) assisting students to develop a holistic vision of the environment; (ii) promoting sustainable resource use; (iii) environmental awareness; (iv) interdisciplinary research; (v) training in environmental issues through degree courses; and (vi) social participation.	Evaluating the effectiveness of a peer education programme in which university students train to deliver workshops promoting the Earth Charter to fellow students. The WeValue toolkit was applied during a workshop to evaluate the extent to which it contributed to an increased sense of empowerment.
Sierra Leone Red Cross Society (SLRCS) 'Youth Empowerment Project: Youth as Agents of Behavioural Change'	No relationship: initial connection made at conference hosted by a consortium member	Kabala, Sierra Leone: vulnerable youth (e.g. former child soldiers; unemployed; homeless; abusing drugs; victims of violence)	National humanitarian society piloting an initiative designed by the International Federation of Red Cross and Red Crescent Societies (IFRC), a large international NGO. Vulnerable youth participate in YEP-YABC peer education workshops (which promote a culture of peace, gender equality, social inclusion and intercultural dialogue) while working together as ethnically diverse teams on agricultural sites.	Evaluating the extent to which the project had been successful in creating an environment of inclusiveness and non-discrimination, and in empowering the youth. The WeValue toolkit was applied in the context of a workshop for 60 youth at an agricultural site.

action is ultimately ineffective. This highlighted to the evaluators the importance of providing participants with clear ongoing paths of service, and of either integrating members into service-focused networks or providing them with volunteer support to help them turn their aspirations into effective action.

In summary, the CSO felt that the data from their assessment provided useful, meaningful and relevant information to them and thus had face validity. They not only achieved an assessment of the effectiveness of the workshop in increasing empowerment, as hoped, but also gained some rich insights into the reasons for its success and how it might be further improved. Table 3 provides rich details.

5.2. 'Youth Empowerment Project: Youth as Agents of Behavioural Change', Sierra Leone

This field trial involved two CSOs: the International Federation of Red Cross and Red Crescent Societies (IFRC) which is developing the 'Youth as Agents of Behavioural Change' initiative at a global level, and the Sierra Leone Red Cross Society (SLRCS), a member society in which the initiative is being piloted under the name YEP-YABC (Youth Empowerment Project: Youth as Agents of Behavioural Change). These separate stakeholders were essentially interested in the same local evaluation questions, namely the extent to which the project had been successful in creating an environment of inclusiveness and non-discrimination and in contributing to the empowerment of marginalised youth; but they were asking them for different reasons. For SLRCS, the goal was to understand how well their own project was 'working' at a local level. IFRC, however, wanted to explore the effectiveness of the YABC pedagogy in a real field setting, and the extent to which it helped participants to enact the espoused IFRC fundamental

principles (humanity, impartiality, neutrality, independence, voluntary service, unity and universality) and their underlying values such as empathy, cooperation, respect and inclusiveness (K. Beeckman, Head, Principles and Values Unit, IFRC: personal communication, 2010). Accordingly, the core team of co-evaluators included an IFRC programme officer with responsibility for Fundamental Principles and the YABC initiative, the National Youth Coordinator of SLRCS, and two researchers. One of the local youth volunteers from the Kabala agricultural project was also involved in the planning for some evaluation activities.

An initial selection of indicators from the reference list was made by IFRC, focusing on what was most useful for evaluating the success of YABC as a global initiative. The Head of the Principles and Values Department of IFRC asked ten youth leaders from the global YABC network to comment on the extent to which they felt the indicators could be useful *for them* in the field. The feedback was used for selecting ten indicators for evaluating the SLRCS YEP-YABC project.

Immediately before the field visit, the two university researchers together previewed the menu of assessment tools for those potentially suitable for use with 60 predominantly non-literate youth in the upcoming project. On arrival the conversations with the CSO staff were iterative, beginning with a broad discussion of WeValue and YABC, and gradually achieving clarity about indicators, methodologies and specific assessment tools that were suitable. Although the YABC youth leaders had previously selected relevant indicators, this 'expert group' of four people reviewed the full reference list again before confirming the original selection. They then localised the wording (see Table 4) and planned a complex evaluation combining five methods, namely (i) spatial/corporal surveys; (ii) a secret ballot survey; (iii) focus groups, with and without role-play; (iv) structured and semi-structured

Table 3
Results of *WeValue* tool testing at PIMAUG. Indicator numbers from the reference list (WeValue, 2012) are shown in square brackets.

PIMAUG localised indicator	Assessment tool(s)	Results
Workshop participants are given autonomy and trust to fulfil responsibilities [6], at the same time receiving encouragement and support [7]	**Semi-structured non-participant observation**: Two independent external observers (the university researcher and a staff member from the Earth Charter Initiative) watched the workshop and separately took notes of activities that appeared relevant to this indicator, later comparing their notes.	Observers noted that in several workshop activities, participants were given tasks to complete without supervision, but the facilitators made themselves available to support the group in these tasks.
	Semi-structured key informant interviews: The external researcher interviewed the PIMAUG team, using this indicator as a theme to guide the discussion.	PIMAUG team members felt that much more could be done to support participants in translating the enthusiasm generated by the workshop into sustained and effective action.
Participants feel that their worth and value has been acknowledged in the workshop [35]	**Spatial survey**: Workshop participants were asked to step to their left to answer NO, stay in their line to answer MORE OR LESS, or step to their right to answer YES.	All 20 participants answered 'yes', although there was concern that the results may have been influenced by social desirability bias and/or group conformity bias. Anonymous surveys might be more appropriate for this indicator.
Participants are encouraged to express their opinions [56]	**Semi-structured non-participant observation**: As above	Observers noted that the facilitators repeatedly prompted for participants' opinions, and provided a number of activities based on individual self-expression.
Participants are given the opportunity to explore and reflect upon their own ideas and traditions, and to develop their own visions and goals [66]	**Semi-structured non-participant observation**: As above	Observers noted that participants were repeatedly invited to reflect on their circumstances, their personal histories, their values and beliefs, their society, etc. in a multitude of formats and ways, and that a personal goal setting exercise was incorporated at the end of the programme to help participants translate the message of the workshops into action.
Participants have identified their own responses to an issue, rather than just agreeing with the ideas of others [73]	**Spatial survey**: Workshop participants were asked to step to their left to answer VERY LITTLE, stand still to answer MORE OR LESS, or step to their right to answer MOSTLY.	11 participants answered 'mostly', nine 'more or less', and none 'very little'.
The workshop has an emotional effect on participants [75]	**Word elicitation with What/Why grid**: Participants were asked to write on one side of the whiteboard three emotions which the workshop stirred in them, and on the other, a sentence expressing the reason for the emotions (with repetition allowed).	The most common emotions were *joyful* and its synonyms, followed by *motivated, committed, persevering*, etc., and then by words such as *connectedness, love, solidarity and identity*. Most cited "sharing experience with like-minded people" as the reason for those emotions, while a few cited fun, learning, and "sense of possibility".
Participants are provided with opportunities for personal growth [78]	**Semi-structured non-participant observation**: As above.	Observers noted that the workshop explicitly encouraged participants at different points to think of their potential contribution and personal development in their lives, to cultivate a sense of possibility, and to develop their own vision and goals (many of which related to personal growth).
Participants feel a sense of power to effect a positive change [113]	**Corporal survey**: Workshop participants were asked to cross their arms to answer VERY LITTLE, put their fists against their hips to answer SOMETIMES, or flex their biceps to answer MOSTLY.	19 out of 20 individuals answered 'mostly', one answered 'sometimes', and none said 'very little'.
Participants feel that their sense of power to effect a positive change has increased after doing the workshop [113]	**Corporal survey**: Workshop participants were asked to sit on the floor to answer IT HAS DECREASED, stand still to answer IT HAS STAYED THE SAME, or raise their arms to answer IT HAS GROWN.	17 out of 20 felt their sense of power to effect positive change had increased, three felt it had stayed the same, and none said it had decreased.
	Focus group discussion: The peer educators led a 15-min discussion, which centred on the reason why people felt different levels of power to effect positive change, and why for some the workshop had increased this sense of power to effect a change, while for others it had not affected it.	The key factor that made participants feel an increased sense of power was being in touch with like-minded individuals, compared to a feeling of relative isolation prior to the workshop. Those who felt that their sense of power had not increased reported that they had already participated in a previous PIMAUG Earth Charter workshop, or were already aware of the possibilities and obstacles discussed. Another factor was the lack of clear ideas for translating their motivation into effective action.

non-participant observation; and (v) informal interviews. The CSO asked the external researcher to lead on data collection, which took place throughout a two-day workshop, interspersed with standard YABC Toolkit activities (e.g. role-plays and games) facilitated by the CSO representatives. Afterwards, the four co-evaluators analysed the data together. The most exciting finding was a clear, statistically significant difference ($p < 0.05$, using a t-test) in the

extent to which youth felt discriminated within their YEP-YABC project teams (65% responding 'never', 30% 'sometimes', and only 3% 'very much or all the time') versus their local communities (31% never, 43% sometimes, 26% very much or all the time).

With some of the spatial surveys, there were concerns about social desirability bias (Arnold & Feldman, 1981) and group conformity bias (Jetten, Spears, & Manstead, 1996), as well as a

Table 4

Results of *WeValue* tool testing at SLRCS 'Youth as Agents of Behavioural Change' project. Indicator numbers from reference list (WeValue, 2012) are shown in square brackets.

SLRCS localised indicator	Assessment tool(s)	Results
Decision-making processes within teams are ethical [14] and democratic [15], and provide for equal representation [16]	**Structured non-participant observation (with informal follow-up interviews)**: Three independent observers watched group discussions and decision-making processes during YABC workshop exercises and used an assessment sheet to keep a tally of how many individuals (a) talk spontaneously, (b) talk after being prompted, (c) are passive, or (d) dominate the conversation. In one group, the researcher informally asked the passive members why they were not participating, and whether they felt this situation was typical.	In the first group, 3 of the 8 men dominated the conversation, while 5 of the 7 women were passive and did not speak at all. Passive members reported that the situation *was* typical: "the big boss always makes the decisions". The second group also had many passive members, but the leader made more effort to prompt people to voice their opinions. In the third and fourth groups, decision-making processes were much more participatory, in that only two members were passive in each case. These groups differed, however, in that Group 3 the conversation was largely dominated by five men whereas in Group 4 two men and two women were dominant. Co-evaluators attributed the differences to individual personalities and facilitation styles.
Team members share their skills and abilities freely with one another, regardless of ethnic origin [47]	**Semi-structured non-participant observation**: Two independent external observers (university researchers) watched the workshop and separately took notes of activities that appeared relevant to this indicator, later comparing their notes.	Observers noted that when the youth were asked to perform traditional dances, they could not spontaneously distinguish members of different ethnic groups, and had to ask among themselves. To support under-represented ethnic groups, others who knew the songs readily joined in. No ethnic discrimination was observed in any workshop exercises.
Team members feel they are encouraged to reach their potential [76] and provided with opportunities for personal growth [78]	**Spatial survey**: Team members were asked to step to their left to answer A LITTLE, stand still to answer MORE OR LESS, or step to their right to answer A LOT, in response to questions about whether, in the project, they felt encouraged to (a) make a good living, (b) have good relationships with others, or (c) be a better person.	No valid data could be collected, as researchers felt that the questions were too abstract and the youth did not fully understand them. In addition, there were concerns about group conformity bias, social desirability bias, and participant fatigue due to the hot sun.
Team members feel they are treated equitably and with fairness [120]	**Spatial survey**: Team members were asked to step to their left to answer NEVER, stand still to answer SOMETIMES, or step to their right to answer VERY MUCH/ALL THE TIME, to the question "In your [home] village, do you feel discriminated because of your past, tribe, gender, or anything else?"	18 members responded that they never felt discriminated in their home villages, 25 answered 'sometimes', and 15 answered 'very much/all the time'.
Entity acts in a manner that is impartial and non-discriminatory (not discriminating on the basis of nationality, ethnic origin, colour, gender, sexual orientation, creed or religion)	**Spatial survey (divided by gender)**: Members were asked to step to their left to answer NO, or to their right to answer YES, to the question "Do you feel that women are treated as they should be within your team?"	In the men's group, 29 answered 'yes' and three 'no'. In the women's group, 27 people unanimously answered 'yes'.
Women feel that they are valued and given equal opportunities to participate in decision-making processes	**Secret ballot**: Team members were given three ballot papers with spots of different colours to represent NEVER, SOMETIMES and VERY MUCH/ALL THE TIME, respectively. In response to the question "In your project team, do you feel discriminated because of your past, tribe, gender, or anything else?", they placed the paper representing their response into a ballot box and discarded the others.	39 members answered that they never felt discriminated in their project teams, 18 answered 'sometimes', and 3 answered 'very much/all the time'.
	Focus group with role-play (men only): Team members organised in groups of 3–4 were asked to enact through role-play, and then to discuss, examples of discrimination and good treatment (non-discrimination) respectively in the wider communities and in their teams. They were also asked to enact ways in which the situations of discrimination might be changed, and to discuss opportunities and barriers to behaviour change.	Team members were able to act out examples of feeling discriminated within the project teams, e.g. because they were poor or had less attractive clothing than others, as well as examples of non-discrimination such as sharing food and working together in the fields. Expectation of violent reaction was identified as a possible barrier to intervention.
	Focus group discussion (women only): Female team members discussed the question of whether women are treated as they should be within the project teams, without any role-play. Co-evaluators chose to avoid introducing men into the women's focus group as facilitators, interpreters or camera operators, as it was felt that the women might be reluctant to discuss this sensitive issue in front of men.	Data could not be collected because of the poor standard of English of the female interpreter, and the absence of a female camera operator to record the session for later analysis. An observer noted, however, that the women talked much more freely in the single-sex group than in mixed-sex groups, and that there were lively discussions.

recognition that one of the questions was too abstract for the youth to understand. These problems could have been overcome by pre-testing the assessment tools. Nonetheless, the spatial survey method in combination with other methodologies provided valuable insights into the nature of discrimination and non-discrimination within the project, and the dynamics of the project teams. Focus groups provided a safe space for youth to act out examples of feeling discriminated within the teams, e.g. because they were poor or had less attractive clothing than others, and to discuss possible solutions. Gender segregation was found to be critically important for the focus groups, however, the women-only focus group was adversely affected by the lack of skilled female translators. The structured observation of individual youth and group discussions revealed that one of the four teams had dominating vocal individuals and passive (especially female) members, but found the other teams relatively well balanced.

Localisation of the assessment methods was critical to the success of this evaluation. As mentioned above, there were

Table 5
Results of *WeValue* tool testing at People's Theater. Indicator numbers from reference list (WeValue, 2012) shown in square brackets.

PT localised indicator	Assessment tool(s)	Results
Ethical values and principles are used by youth volunteers in guiding decision-making and activities [104]	**Structured non-participant observation, with self-assessment and a follow-up dialogue**: Youth volunteers: • [do not] interrupt others; • [do not] put others down and provoke • [do not] focus on their own personal goals while ignoring the preferences of others; • [do not] use aggressive non-verbal communication when dealing with others (e.g. frowning, squinting eyes critically, using a critical, loud or yelling tone of voice, using fast, clipped speech, moving into people's space); • [do not] waste other people's time; • [do not] always agree with others regardless of what is said, trying to 'sit on both sides of the fence' to avoid conflict; • [do not] stand aside, failing to express their own wants and feelings, and allowing others to make choices on their behalf; • [do not] use passive non-verbal communication when dealing with others (nodding the head very often, downcast eyes, fidgeting, low volume of voice, slow speech or very fast, anxious and hesitant speech); • [do not] have difficulty implementing plans when they are made; • know what is needed and develop a plan to get it; • are realistic in their expectations; • express their own feelings and wants honestly, while also considering the feelings and wants of others; • use assertive nonverbal communication when dealing with others (natural gestures, direct eye contact, appropriate volume of voice, varied rate of speech)	The main purpose of the structured observation was not to generate aggregated results showing (for example) how many youth interrupted others or how many used assertive nonverbal communication. Rather, it was to build up a clear picture of the behaviour, or 'values in action', of each individual participant during the rehearsal. For most items, there was agreement between at least 2 out of 3 observers, and PT staff who spent time with the youth on a daily basis throughout the year confirmed that the observations described the personality of each youth participant well. Comparison of the independent observation and self-assessment enabled fruitful discussions with each youth about his/her behaviour, based on concrete findings. The youth greatly appreciated the opportunity for self-reflection.
Everyone has their place in the team [1]	**Structured non-participant observation, with self-assessment and a follow-up dialogue**: Youth volunteers (i) participate in group work; (ii) find their (strongest) role in the team; (iii) ask relevant questions and promote thoughtful discussion; and (iv) help in cooperation with others to achieve a set goal	See comment for indicator #104 above
The organisation is transparent about the processes [23] and outcomes [24] of decision-making, and shares information openly with the youth volunteers [26] Youth volunteers feel there is the right information flow [25]#	**Questionnaire**: Youth volunteers (i) feel that they get information at the right time; (ii) feel that they get complete information; (iii) feel that they get all the information needed for their work in PT (from staff and from co-workers/other team members)	1 out of 6 youth volunteers answered 'always' 5 out of 6 answered 'mostly' 1 out of 6 youth volunteers answered 'always' 5 out of 6 answered 'mostly' 4 out of 6 participants answered 'always' 2 out of 6 answered 'mostly'
Individuals in a team feel they have an equal opportunity to voice their opinions [54] and their opinions are respected and listened to [58] Youth volunteers feel they are treated equitably and with fairness [120]	**Questionnaire**: Youth volunteers feel that both in their team and in PT as a whole, (i) they have an equal opportunity to voice their opinions in decision making processes; (ii) their opinions are listened to; (iii) their opinions are respected	5 out of 6 youth volunteers answered 'always' 1 out of 6 answered 'mostly' All participants responded that their opinions are always listened to in the youth team. In PT as a whole, 4 out of 6 responded 'always' and 2 out of 6 'mostly'. 3 out of 6 participants answered 'always' 3 out of 6 answered 'mostly'
Youth volunteers encourage others to express their opinions [57]	**Structured non-participant observation, with self-assessment and a follow-up dialogue**: Youth volunteers (i) listen to others' opinions; (ii) help quieter members to express their opinions; (iii) use encouraging words to support others; and (iv) use nonverbal communication to support others (gestures, eye contact, voice, touch, etc.)	See comment for indicator #104 above

still several issues that were felt to reduce the validity of some of the evaluation data, but none of these would have been insurmountable if more time had been available. Overall, the assessment tools achieved their local evaluation objective – a notable achievement, in the light of the still-fragile sense of group cohesion among individuals who had once fought on opposite sides in a civil war.

The evaluation was very useful to IFRC because it generated preliminary evidence for the efficacy of the YABC approach as a whole, and to SLRCS because it confirmed that the YEP-YABC project was working well towards its intangible values-related objectives. It also demonstrated to both stakeholders that much work remains to be done, and highlighted specific areas for improvement. See Table 4.

5.3. People's Theater, Germany

At PT, the co-evaluators were the CSO director, a social worker and a visiting university researcher: the youth volunteers themselves did not participate actively in the evaluation planning. One aim of the evaluation was to assess in the youth the development of manifestations of higher values, each of which could have several indicators. The CSO's local evaluation question decided on was: to evaluate the extent to which the youth (and staff) were enacting the five espoused core values of PT in their day-to-day activities and organisational processes (consultative dialogue and decision-making, encouragement, unity in diversity, service orientation, and supporting individual initiatives). From the reference list of WeValue indicators the co-evaluators identified 11 that they felt were appropriate (see Table 5). All of these indicators were considered desirable for the evaluation, when localised to the specific PT context.

The challenges of developing suitable assessment tools at PT differed from the two preceding case studies, in that rather than evaluating the perceived impact of a discrete event such as a workshop, the evaluation concerned the behaviour of the youth during ongoing activities. The co-evaluators decided to focus on *rehearsals* for the theatre performances, one of the central project activities of the youth volunteers, which provided an ideal opportunity for observation of interpersonal behaviour in a natural but goal-oriented environment. The plan was to do this by combining three separate methods: (i) structured non-participant observation; (ii) youth self-assessment of their own behaviour; and (iii) follow-up one-on-one dialogues between the youth and their observers, comparing findings from the observation and the self-assessment. These methods were supplemented by a questionnaire to collect information on organisational processes involving both staff and youth. See Table 5 for details.

For the structured observation, co-evaluators converted the indicators into a standard assessment pro-forma on which observers could record the occurrence of specific behaviours on a three-point scale ('very much', 'sometimes' or 'not at all'). In the case of very general indicators such as "Ethical values and principles are used by youth volunteers in guiding decision-making and activities", the co-evaluators first had to define specific positive behaviours whose presence would indicate enactment of the espoused core values, and negative behaviours whose absence would indicate enactment, according to PT's own codes of conduct. Three separate observers (two CSO staff members, and one independent external evaluator included for comparison purposes) watched a normal rehearsal, simultaneously recording it on video for later cross-checking. Participants were aware of the evaluation, but to minimise social desirability bias, they were not told which youth were being assessed or which specific elements of their behaviour would be observed. Immediately after the rehearsal, each participant was given a blank copy of the assessment form and asked to reflect on their own behaviour during that specific rehearsal and complete the form themselves. Seven youth were observed altogether, in two separate rehearsals. The differences between the results of the observation and those of the self-assessment were then discussed, item by item, in face-to-face dialogues between individual youth and their observers. During the course of these dialogues, as in unstructured interviews, the researchers acquired further qualitative data about the feelings and motivations of the youth. Finally, all of the youth were asked to complete an anonymous closed-ended questionnaire about their views on internal processes within PT. Descriptions of these methods, with results, are shown in Table 5.

The co-evaluators later analysed the data together (see Table 5). It confirmed the facilitators' instinct that interpersonal interactions among the youth were generally very positive and that, in a broad sense, the core values were being enacted. The observation findings showed that the majority of participants were able to find their strongest role in the team, participate in group work, cooperate with others to set goals, and encourage others to express their opinion. While some of the youth did not always promote thoughtful discussion or ask relevant questions, the dual nature of the assessment alerted them to these shortcomings, and the follow-up dialogue provided project staff with a clear forum for encouraging them. Co-evaluators also noticed that youth self-assessments tended to underestimate the manifestation of positive behaviours during the rehearsal, in comparison to the consensus of the three observers, probably because they were asked to reflect retrospectively. This meant that the follow-up dialogues were uplifting, and served to boost the self-confidence of the youth.

The results of the questionnaire were also very positive and revealed that there were no areas of major concern with respect to communication flow, decision-making or the fulfilment of responsibilities, also indicating enactment of the core values. Co-evaluators acknowledged, though, that the information gleaned from closed-ended questions did not reveal much about the real situation. Incorporating some open-ended questions would have made the questionnaire more useful.

In summary, the CSO felt that the local process evaluation with the WeValue toolkit was useful because it evidenced that there was an overall positive nature to interactions within the organisation, and that the youth were generally 'walking their talk' within their day-to-day activities. It also proved useful as a tool that could alert the youth themselves to the existence of value-action gaps, and help staff find new ways to support them to close these gaps.

6. Discussion

In this section, we first discuss issues relevant to the five objectives listed in Section 3.2, followed by wider topics relating to evaluation practices.

6.1. Relevance of the WeValue indicators to participating CSOs

The main driver for developing a new evaluation tool was the stated CSO need for one which could measure values-based dimensions. Meaningful CSO participation in such evaluations is crucial if the specific 'intangibles' requiring assessment are to be correctly identified. Specifically, Daigneault and Jacob (2009) observe that an important dimension of participatory evaluation is the extent to which relevant stakeholders (e.g. policy makers, project implementers and beneficiaries) are involved in selecting the evaluation questions and issues. Springett (2001) has emphasised that true participatory evaluation requires these stakeholders to develop their own indicators of success from scratch. Our novel approach has been to use a peer-generated 'menu' of indicators which can be used to expedite indicator development in similar organisations. These were developed by the ESDinds project consortium, and their relevance to their own, affiliated or non-affiliated CSOs needed evaluating. People's Theater was a consortium member and thus had extensive involvement in the peer-generation of the indicators list. It was thus not surprising that PT staff perceived some 40 draft indicators as highly relevant to their work. PIMAUG is loosely affiliated to a consortium member (Earth Charter). Where 74 indicators were discussed, 56 were initially perceived as 'relevant' for the work of PIMAUG and 24 were later prioritised as 'highly relevant'. The IFRC movement's YABC initiative was not affiliated to the consortium. In a separate study of indicator relevance 61 youth leaders of the YABC initiative, meeting in a global summit, rated the WeValue

indicators. Every indicator was returned as either 'relevant' or 'highly relevant' by at least 15 individuals.

The process of scoring and then discussing the indicators for relevance can be regarded as an example of participatory learning and appraisal. In common with Pretty, Guijt, Thompson, and Scoones (1995), we have observed catalysis of vigorous debate, often generating exciting and unexpected results. At both PIMAUG and PT, the simple act of reading the indicator list triggered new realisations about what was usually missing from their projects and how those might be improved. We believe this aspect of heightened relevance is due to the fact that the indicator reference list is peer-generated though a 'values lens' (see Section 6.6). These are important design features to carry forward into future work.

6.2. Usefulness of the indicators for meeting specific local evaluation goals of the CSOs

The CSO field staff were asked to determine whether the WeValue indicators were not just generally relevant but also potentially useful for meeting their local specific evaluation goals as discussed in Section 5. All three of the CSOs identified several indicators as useful: the ESDinds consortium member (PT) shortlisted 11 indicators; the loosely affiliated PIMAUG, 12; and the unaffiliated SLRCS project, 17.

What is clearly evident, however, is that the indicators could not have been considered very useful if the reference list were treated as rigid and inflexible. In all three CSOs it was used only as an initial starting point, and the indicators then localised to the specific CSO contexts. Iterative processes of short-listing, editing, discussion of possible assessment methodologies and final selection were used to refine the indicators to local contexts, with varying degrees of participation by the different stakeholders (decision-makers, implementers and beneficiaries) at each stage. The extent of localisation varied from simple changes in wording, such as replacing the generic term 'people' with 'youth volunteers', to profound changes involving the definition of specific local meanings for indicators initially worded in general terms. At PIMAUG, for example, the generic indicator "People feel a sense of power to effect change" was modified to two sub-indicators, "Participants feel a sense of power to effect *a positive* change" and "Participants feel that their sense of power to effect a positive change has increased after doing the workshop". At PT, the indicator "Ethical values and principles are used in guiding decision-making and activities" was linked to the CSO's specific code of conduct for its youth volunteers. This promoted assertive communication styles over aggressive or passive ones, and explicitly discouraged certain behaviours, such as interrupting others (see Table 5). In both of these cases, the general indicators could have been rigidly interpreted in a myriad different ways, but the WeValue approach required the interpretation to be confirmed and agreed by the co-evaluators for each specific evaluation. This prescribed level of localisation and agreement ensured the great usefulness and clear definition of the indicators locally.

6.3. Assessment tools and methods

For every localised indicator chosen for use in these studies, it was possible to develop at least one successful assessment tool that could be used to collect relevant data. All participating CSOs chose to utilise a mixture of qualitative and quantitative methods in their local evaluation, and in some cases, both types of methods were used to measure a single indicator. This approach was encouraged (not imposed), as convergence of the results from different measurements is a strong indication of validity (Campbell

and Fiske, 1959; see also Lather, 1986; Mathison, 1989 regarding triangulation). Even when individual assessments are 'methodologically imperfect', they can still be aggregated to generate a valid conclusion, provided that they are congruent in their findings but differ in their inadequacies (Cox, Karanika, Griffiths, & Houdmont, 2007).

The outcomes of the WeValue field trials confirm several previously reported advantages of mixed-methods evaluation. Onwuegbuzie et al. (2007) refer, for example, to "weakness minimisation legitimation", a scenario in which a weakness of one approach is compensated by the strengths of another; while Bouffard, Taxman, and Silverman (2003) have observed that combining methods generates a richer understanding of processes and increases the benefits of process evaluation to stakeholders. In all three of our case studies, quantitative results, which showed the 'big picture' but provided only superficial answers to the CSO questions, were complemented by rich qualitative data that filled in the detail of *why* respondents felt or behaved in the ways they did. This enabled strengths, weaknesses, opportunities and threats to be clearly identified and addressed:

(a) In the PIMAUG workshop, the spatial and corporal surveys (quantitative) revealed that the majority of participants feel empowered, while the focus group (qualitative) highlighted many also lack awareness of how to translate their enthusiasm and motivation into effective practical action. In subsequent workshops PIMAUG has offered a 'menu' of specific volunteer opportunities.

(b) In the SLRCS project, the structured observation (quantitative) allowed facilitators to count how many individuals monopolised discussions or were passive, while follow-up interviews (qualitative) provided insights into the reasons for reluctance to participate. This allowed identification of effective future support actions.

(c) At PT, the structured observation and self-assessment of interpersonal behaviour (quantitative), gave the number of youth who felt that they exhibited specific 'positive' and 'negative' behaviours versus those who actually exhibited them. The follow-up dialogues (qualitative) created new opportunities for meaningful conversations between staff and youth – 'how you see yourselves' versus 'how we see you'. These results, taken together, helped the staff to identify ways of supporting individuals more effectively to fulfil their potential.

6.4. Face validity: meaningfulness of the results

The main driver for the initial development of the WeValue toolkit was the need for utilisation by partner CSOs, so their ability to relate to the results was a crucial required outcome. As illustrated in Section 5, the participating CSOs found the data highly relevant, meaningful and useful. In the follow-up interviews, key informants from PIMAUG and People's Theater confirmed that their evaluation questions had been answered to their satisfaction and had given them further insights, although PT felt that they needed more time to draw out the learning. Representatives from IFRC and SLRCS reported that the result showing that youth experience less discrimination in the YABC project teams than in their home villages was very useful and this has been incorporated into various YABC reports. They also, however, perceived a need for further evaluation to answer their questions about empowerment more fully. This was not because the toolkit is inherently unsuitable for this question, but because time constraints prevented co-evaluators from overcoming challenges associated with the context in Sierra Leone and the assessment tools initially selected for it.

Table 6
Process use benefits of the *WeValue* evaluation tool observed in the three field studies.

Process use benefits	Where?	How?	Specific outcomes
Learning how to learn about intangible values-related processes and outcomes	All field sites	CSO staff greatly developed their critical thinking skills and sense of "evaluative perception". CSO staff learned to develop and apply specific customised assessment tools that complement their existing activities.	A culture of learning, monitoring and self-evaluation has been strengthened within all three CSOs. Assessment tools used in the field visits have been incorporated into CSOs' ongoing M&E strategies, both in association with the values-based indicators and in relation to other types of assessment.
New shared understandings of the role of values in CSO's work, and their relationship to behaviour change	PIMAUG	Participants' consciousness of the presence and importance of values in their work was greatly heightened. One peer educator reported that it helped her to view the Earth Charter in terms of holistic personal development (starting with herself).	The content of peer education workshops has moved away from a primary focus on concrete behaviours, such as recycling waste, to a holistic values-centred approach that is anticipated to generate the desired behaviours in a more deep-seated and sustained way.
	SLRCS	The youth gained a deeper understanding of how IFRC core values can be translated into action.	Motivation and commitment increased within the youth teams.
	IFRC	The IFRC representative learned that to address values at a deep level, individuals need to start by reflecting on their own behaviour, and from this derive the power and confidence to address others.	The IFRC representative was able to communicate the expectations and goals of the YABC initiative more effectively to a Government department, which has accordingly modified the way in which it understands and communicates the concept of behaviour change.
	People's Theater	The *WeValue* tool brought values consciousness to the forefront of PT's activities, and strengthened its identity as a values-based organisation.	Both the orientation programme for new volunteers, and the way in which the goals of PT's work are communicated to new schools, have been restructured to centre around values.
		Reflecting on the indicators prompted staff to review their list of espoused core values.	There is now an increased focus on 'Justice', which was not in PT's original list of espoused core values.
Strengthening the projects	PIMAUG	Merely reading the list of indicators helped the PIMAUG team to identify and promote specific processes that are important for project success, which they had not previously thought about. The PIMAUG coordinator has recognised the potential for transformation that is inherent in values-based evaluation, and is now seeking to scale it up.	The content of the PIMAUG Earth Charter workshops has been modified to address concerns highlighted by some of the indicators, such as the need for appropriate channels for expressing grievances or reporting ethical violations. WeValue indicators are already being mainstreamed within PIMAUG. There is a plan to pilot them at departmental level, and if successful, to then embed them in the entire university system.
	People's Theater	Staff gained awareness that participatory values-based M&E can help the youth to identify and close their own values-behaviour gaps. Staff reviewed the espoused core values of the project.	The motivation questionnaire previously used for internal M&E with youth has been replaced by a *WeValue* approach aiming to catalyse deeper conversations and promote sustainable behaviour change. Consideration of 'Justice', which was not previously viewed as a core espoused value of PT, was incorporated into a revised orientation programme for new volunteers.
Boosting morale	All field sites	Staff/facilitators gained confidence in their ability to evaluate their own projects, as well as a deeper appreciation of the significance of their work.	Visible improvements in morale among both staff/facilitators and youth/beneficiaries.
Developing professional networks	All field sites	CSO staff built strong working relationships with the university partners, and each key informant became aware of the learning from the other field visits.	Co-development of an online *WeValue* platform and CSO handbook, and potential for joint publications.

6.5. Attitudinal changes

The WeValue evaluations exceeded all expectations in the sense that it was not only the data that proved useful, but also the *process* of conducting participatory evaluations with a values-based tool (c.f. Crohn & Birnbaum, 2010; Patton, 1998). Forss, Rebien, and Carlsson (2002) have identified five categories of process use benefits, namely 'learning how to learn', creating shared understandings, strengthening the project, boosting morale, and developing professional networks. Table 6 illustrates how these were manifested in the case studies.

All three CSOs effectively *learned how to learn* about intangible values-related processes and outcomes. This was evident both in their enhanced evaluative perception and critical thinking skills, and in their newly acquired ability to create and apply appropriate assessment tools that complement their ongoing activities. In each, internal monitoring and evaluation strategies have since been changed, with integration of the WeValue tool on an ongoing basis, showing CSO empowerment in accordance with work by Patton (1998).

Conducting evaluation through a values lens created shared understandings within all four field CSOs. Using values-based indicators for local evaluations deepened their understanding of both the significance of values in their work, and the relationship between their values and behaviour, which has led to substantial changes in their communication strategies.

In PIMAUG and PT, ongoing projects have been strengthened through concrete changes to policy and practice. For both CSOs there has been a concurrent shift in the way in which monitoring and evaluation are perceived: no longer merely informative (identifying values-behaviour gaps) but potentially transformative (catalysing sustainable behaviour change to close the gaps). At PIMAUG, this attitudinal shift was so profound that the team is now designing and implementing a staged and iterative process to

scale up the application of the WeValue tool – first within the university-wide Environmental Management System and PIMAUG's own network of allies, followed by a separate intensive pilot at departmental level, to be applied across the whole university if successful.

Finally, as shown in Table 6, the morale of staff and youth has been substantially boosted in all CSOs. Furthermore, the emergence of professional networks involving the user CSOs and ESDinds research team has resulted in an online WeValue platform, handbooks for CSOs, and ongoing collaborations to produce this paper and other, forthcoming academic publications.

6.6. Other contributions to evaluation practice

This non-traditional approach to evaluation (in particular, values-based and CSO-utilisation focused) has generated results relevant to several current issues in the evaluation literature. These include, among others, fitness for purpose; mixed-methods evaluation; and the validity of the dichotomy between 'process' and 'impact' evaluation.

6.6.1. Fitness for purpose

The tension between pursuing scientific rigour and remaining flexible to participants' diverse needs and expectations, reported in action research literature (Peterson, 2010), is highly relevant to evaluation. An over-emphasis on multiple assessment methodologies and systematic cross-checks, especially when imposed by external evaluators or researchers, could convince CSO staff that approaches such as WeValue demand impracticable levels of time and resource investment. An useful contribution to this debate has been made, however, by Cox et al. (2007) with the concept of 'fitness for purpose'. Research that is fit for purpose, they assert, addresses a problem that is important to society; has a clearly defined purpose and research question; and utilises theoretically and ethically legitimate methods that can answer the research question and meet the purpose of the study. Within a CSO, there may be multiple purposes for undertaking monitoring and evaluation, leading potentially to the development of parallel systems. On the one hand, small-scale data collection and analysis could be incorporated into ongoing project activities (as in the PIMAUG example) to improve their effectiveness without a large resource commitment; while, on the other, rigorous external evaluations with the WeValue indicators could be conducted at key stages in project life cycles to provide accountability to donors and the public.

6.6.2. Mixed methods evaluation

Odendaal et al. (2008) advocate combining qualitative and quantitative methods as a means of accommodating the methodological preferences of diverse stakeholders: some prefer anonymity, for example, while others appreciate personal interactions. The WeValue approach goes even further, by providing an extensive 'menu' of possible assessment methodologies and encouraging customisation of assessment tools from these methods to ensure congruence with ongoing project activities. Thus, co-evaluators can respond to the needs and preferences of diverse stakeholders (subject to time and human resource availability), collecting data in even the most challenging of contexts. This was evidenced at the SLRCS-YABC evaluation – integrated within a time-limited workshop involving large groups of non-literate youth, affected by numerous cultural, linguistic, conflict and gender-related issues that posed a threat to the validity of the local evaluation. Some difficulties still arose as a result of social desirability response bias and group conformity bias, but the lesson learned was that appropriate assessment tools could indeed be developed to evaluate values-based dimensions;

the difficulties could have been overcome by taking more time to use iterative 'feedback loops' (Flowers, 2010).

In each local evaluation discussed here, the CSOs initially found it difficult to develop appropriate assessment tools, and success would have been unlikely without external assistance – this was expected. However, after undertaking the co-evaluation, all of the key informants reported that they had gained capacity to (a) subsequently identify appropriate assessment methods for different WeValue indicators and (b) develop further creative tools, congruent with their existing activities – for WeValue evaluations or for other uses. Reflecting on indicators and constructing assessment tools helped them to develop a deeper understanding of the inherent values associated with their activities, and to build a strong consensus within their groups about the nature of their core mission, in several cases causing earlier statements of 'espoused values' to be replaced by others regarded as considerably more meaningful. We anticipate that unless their own staff possess relevant research experience, CSOs will continue to require professional support at the tool design and data analysis stages in order to maintain overall validity.

6.6.3. Validity of the process/impact dichotomy

Evaluation literature generally delineates between 'process evaluation' and 'impact evaluation', but we would suggest that the WeValue indicators, by their very nature, blur the boundaries to the extent that this distinction becomes almost meaningless. On the one hand, these indicators broaden the domain of 'impacts' by highlighting those that were previously regarded as immeasurable, and on the other hand, they broaden the domain of 'processes' by focusing evaluators' attention on implementation. This corroborates a point made in the recent policy evaluation guidelines issued by the UK Government (HM Treasury, 2011), namely that 'process' and 'impact' evaluations often consider similar questions and issues, and what is an impact indicator in one context may be a process indicator in another. This is exemplified by the indicator "[Youth] feel they are treated equitably and with fairness" – effectively an impact indicator at SLRCS, where creating an environment of non-discrimination was an explicit objective of the project, whereas it was used as a process indicator at PT, where there was no such objective.

As Cox et al. (2007) explain, a process evaluation that is limited to examining intervention processes – i.e. listing the actual steps taken – may come to the erroneous conclusion that a programme is inherently flawed, when its failure in a particular context was in reality attributable to the way in which it was delivered or managed. We have discovered, through the research project that created the WeValue indicators (outlined by Podger et al., 2010), that implementation processes such as communication flow, interpersonal relationships, decision-making strategies and the engagement of beneficiaries can exemplify some of the 'intangibles' that can make or break a civil society project; yet they are often overlooked in evaluations, and almost never addressed in a systematic way. A detailed exploration of the way in which the WeValue indicators relate to these implementation processes is beyond the scope of this paper, but will be addressed in future work.

7. Lessons learned

The WeValue toolkit appears to successfully fill a known gap in evaluation methods. CSOs have used it to generate data on previously 'intangible' outcomes which, by their own judgement, are considered to be values-based. The data it produced have intrinsic face-validity because the indicators used are chosen by the CSO as representative of the values-based work they wish to assess, and then localised to ensure relevance and agreed

definition in the applied context. In effect, the CSOs are facilitated to develop local evaluations which are based on their own values frameworks, rather than have an external values framework imposed on them. That this is possible at all, is a significant lesson learned.

Furthermore, it was learned that CSOs are able to develop appropriate assessment methods, usually with initial assistance from external evaluators. They were later able to transfer their learning to independently develop new assessment methods for other purposes.

The evaluations carried out produced information of use to wider stakeholders, even though devised in terms of local CSO values (e.g. see Table 6). In their words, once crystallised, evaluated and communicated, values-based outcomes were deemed to have use and importance externally. The lesson learned is that those previously 'intangible' quantities are important, can be made tangible, and can now be used to build up a new vocabulary between values-based organisations and stakeholders such as donors and wider society.

Wanting and achieving participation and face validity (i.e. utilisation-focus) in evaluation are two different things. The approach taken here to facilitate CSOs to crystallise their own values, stimulated by the peer-generated lists, seems to elicit deep participation and thus face-validity, leading to good ownership and utilisation of the results. This suggests two very important lessons for evaluation: a localisable framework is powerful, as is use of peer knowledge and language in its construction.

The framework of WeValue appears to work across a range of organisations – not only the three presented here, but also over 40 others involved in the ESDinds project. This appears to be related to the point made above – that WeValue has built-in peer experience and flexibility of localisation.

Finally, a very important lesson learned is that the core approach of using a 'values lens' to examine the work of essentially values-based organisations, such as CSOs, works well. Readers not familiar with the derivation of the WeValue list of indicators (WeValue, 2012) may not be aware that the list results from empirical information of "what is important (valued)" to CSOs, and processing the responses into clusters of higher values, from which the indicators were derived. Thus, even though a small number are reminiscent of those found in corporate evaluations, as a whole they are values-based. The key difference is that the CSOs strongly relate to this set, affiliating themselves with it. A lesson seems to be that using a values lens is key to working with CSOs (and other values-based organisations).

Evaluation principles of participation, inclusion and utilisation-focus are seen to be critically important for evaluation of values-based dimensions. However, these principles worked when combined in an approach centred on values crystallisation and assessment rather than as an end in themselves. Such an approach often resonates strongly with CSOs and can lead to a deeper voluntary engagement in evaluation processes, and deeper ownership and utilisation of the results, than might be anticipated in a conventional evaluation. It can also lead to sustained change. Nonetheless, the relative importance of actual participation and utilisation-focus in a given evaluation can vary, according to circumstances.

8. Conclusion

The WeValue toolkit with its original components has been shown to fill a known gap in evaluation methods by making tangible previously 'intangible' values-based outcomes. It has been successfully used for local evaluations by a range of CSOs and found useful in disparate evaluation contexts – from assessing values-related outcomes of workshops only a few hours in

duration, to monitoring interpersonal relationships and implementation processes within longer-term projects. We suggest that the power of the WeValue approach, and the key to its remarkable applicability, lies in the fact that it is both structured and localisable. By providing an indicator framework which allows localisation to specific contexts, it elicits good engagement and participation, and thus ownership and utilisation of the evaluation results. However, results from a given indicator on the list can be compared across organisations (albeit with some normalisation). WeValue thus provides a bridge between 'participatory' and 'conventional' evaluation.

In all three of the field trials presented here, the CSOs reported that the evaluation retuned information of use, relevance and meaning to them. They did not feel that the framework imposed external values, but rather, that it facilitated evaluation of their own. This is a remarkable achievement of the elusive CSO-based evaluation. The information obtained in each case was of significant use and interest to wider stakeholders, providing new vocabulary to allow future joint planning.

The WeValue toolkit with its novel components thus represents a 'new-generation' evaluation tool that makes use of key recent ideas in participatory, utilisation-focused and process-based evaluation, which are themselves framed through a values lens compatible with the work of CSOs and other values-based organisations. The toolkit thus shows promise for use with a very wide range of organisations including educational institutions, health care providers, universities and businesses, whether or not they have an explicit commitment to higher values. However, new sets of indicators may need to be peer-generated for each sector.

This original approach has contributions to make in current debates in environmental values, environmental education and indicators for sustainable development where difficulties in crystallising and measuring values-based concepts in different contexts and groups are of key interest.

The toolkit described in this paper is accessible at the online platform www.wevalue.org, which provides case studies, guidelines and contact details for experienced WeValue evaluators. A 'library' of assessment tools, and a variety of training packages for evaluators and CSO staff, are currently being developed.

References

Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. *Academy of Management Journal, 24*(2), 377–385.

AusAID. (2005). *AusGuideline 3.3: The logical framework approach*. Canberra: AusAID.

Bamberger, M. (2000). The evaluation of international development programs: A view from the front. *American Journal of Evaluation, 21*(1), 95–102.

Bouffard, J. A., Taxman, F. S., & Silverman, R. (2003). Improving process evaluations of correctional programs by using a comprehensive evaluation methodology. *Evaluation and Program Planning, 26*(2), 149–161.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56*, 81–105.

Cox, T., Karanika, M., Griffiths, A., & Houdmont, J. (2007). Evaluating organizational-level work stress interventions: Beyond traditional methods. *Work & Stress, 21*(4), 348–362.

Crishna, B. (2007). Participatory evaluation (I)—Sharing lessons from fieldwork in Asia. *Child Care Health and Development, 33*(3), 217–223.

Crohn, K., & Birnbaum, M. (2010). Environmental education evaluation: Time to reflect, time for change. *Evaluation and Program Planning, 33*(2), 155–158.

Daigneault, P. M., & Jacob, S. (2009). Toward accurate measurement of participation rethinking the conceptualization and operationalization of participatory evaluation. *American Journal of Evaluation, 30*(3), 330–348.

Ebrahim, A. (2002). Information struggles: The role of information in the reproduction of NGO–funder relationships. *Nonprofit and Voluntary Sector Quarterly, 31*(1), 84–114.

Ebrahim, A. (2003). Accountability in practice: Mechanisms for NGOs. *World Development, 31*(5), 813–829.

Edwards, M., & Hulme, D. (1996). Too close for comfort? The impact of official aid on nongovernmental organizations. *World Development, 24*(6), 961–973.

Ellis, R., & Hogard, E. (2006). The trident: A three-pronged method for evaluating clinical, social and educational innovations. *Evaluation: The International Journal of Theory, Research and Practice, 12*(3), 372–383.

Flowers, A. B. (2010). Blazing an evaluation pathway: Lessons learned from applying utilization-focused evaluation to a conservation education program. *Evaluation and Program Planning, 33*(2), 165–171.

Ford Foundation. (2011). *Our approach to impact assessment and learning* Retrieved from http://www.fordfoundation.org/impact/approach.

Forss, K., Rebien, C. C., & Carlsson, J. (2002). Process use of evaluations: Types of use that precede lessons learned and feedback. *Evaluation, 8*(1), 29–45.

Heimlich, J. E. (2010). Environmental education evaluation: Reinterpreting education as a strategy for meeting mission. *Evaluation and Program Planning, 33*(2), 180–185.

HM Treasury. (2011). *The magenta book: Guidance for evaluation.*

Hogard, E. (2008). Purpose and method for the evaluation of interpersonal process in health and social care. *Evaluation and Program Planning, 31*(1), 34–40.

Holte-McKenzie, M., Forde, S., & Theobald, S. (2006). Development of a participatory monitoring and evaluation strategy. *Evaluation and Program Planning, 29*(4), 365–376.

Jetten, J., Spears, R., & Manstead, A. S. R. (1996). Intergroup norms and intergroup discrimination: Distinctive self-categorization and social identity effects. *Journal of Personality and Social Psychology, 71*(6), 1222–1233.

Lather, P. (1986). Issues of validity in openly ideological research. *Interchange, 17*(4), 63–84.

Mathison, S. (1989). Why triangulate? *Educational Researcher, 17*(2), 13–17.

Odendaal, W. A., Marais, S., Munro, S., & van Niekerk, A. (2008). When the trivial becomes meaningful: Reflections on a process evaluation of a home visitation programme in South Africa. *Evaluation & Program Planning, 31*(2), 209–216.

Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal, 44*(1), 113–160.

Patton, M. Q. (1998). Discovering process use. *Evaluation, 4*(2), 225–233.

Peterson, J. C. (2010). CBPR in Indian Country: Tensions and implications for health communication. *Health Communication, 25*(1), 50–60.

Podger, D., Piggot, G., Zahradnik, M., Janouskova, S., Velasco, I., Hak, T., et al. (2010). The Earth Charter and the ESDinds initiative: Developing indicators and assessment tools for civil society organisations to examine the values dimensions of sustainability projects. *Journal of Education for Sustainable Development, 4*(2), 297–305.

Pretty, J. N., Guijt, I., Thompson, J., & Scoones, I. (1995). *Participatory learning and action: A trainer's guide.* London: IIED.

Riddel, R. C. (1999). Evaluating NGO development interventions. In D. Lewis (Ed.), *International perspectives on voluntary action: Reshaping the third sector* (pp. 222–241). London: Earthscan.

Spash, C. L. (2009). The new environmental pragmatists, pluralism and sustainability. *Environmental Values, 18*(3), 253–256.

Springett, J. (2001). Participatory approaches to evaluation in health promotion. In M. G. I. Rootman, B. Hyndman, D. V. McQueen, L. Potvin, J. Springett, & E. Ziglio (Eds.), *Evaluation in health promotion: Principles and perspectives* (pp. 83–105). Copenhagen: WHO Regional Publications.

Springett, J. (2003). Issues in participatory evaluation. In M. N. Wallerstein (Ed.), *Community-based participatory research for health.* San Francisco, CA: Jossey-Bass.

Turner, R. K., Paavolaa, J., Coopera, P., Farberb, S., Jessamya, V., & Georgiou, S. (2003). Valuing nature, lessons learned and future research directions. *Ecological Economics, 46*, 493–510.

WeValue (2012). *WeValue Indicator List accessible via 'Participate' page.* Retrieved from: http://www.wevalue.org/participate/index.php.

Wightman, F. (Fall 2010). *Measures of success: Not everything that counts can be counted.* Vancouver Foundation Magazine. Retrieved from http://vancouverfoundation.ca/documents/VanFdn-Magazine-Nov2010.pdf.

Gemma Burford is a Research Officer in Sustainable Development at the University of Brighton, where her research focuses on values-based indicators and empowerment. She previously worked in Tanzania as Founder and Co-Director of the NGO Aang Serian ('House of Peace'), which promotes values-based intercultural education, collaborative research and indigenous knowledge. She is a Trustee of the Global Initiative for Traditional Systems of Health; a co-editor of the World Health Organisation Global Atlas of Traditional, Complementary and Alternative Medicine; and a co-author of diverse publications on health, environment, gender and development issues.

Ismael Velasco was a Research Fellow in Sustainable Development at the University of Brighton during the work presented in this paper, conducting a number of the field visits and designing a number of the research tools. He is now Director of the non-governmental organization Community Resolve in Bristol, UK. With 20 years experience in community development and social policy, his work has specialised in stakeholder engagement, cross-cultural dynamics, and social exclusion.

Svatava Janoušková is a Research Fellow at Charles University Environment Centre in Prague. Her main field of interest is the development and use of educational and well-being indicators. She cooperates with a number of civil society organisations, local authorities and schools, helping them to develop and use indicators for programme or curriculum evaluation.

Martin Zahradnik is a research fellow at Charles University Environment Center in Prague. His field of interest includes evaluation of education for sustainable development and policy decision making processes together with his doctoral thesis, exploring conditions of freedom in contemporary society. He cooperates with several universities and CSOs in projects oriented towards interdisciplinary networking in the field of policy development for sustainable development.

Tomáš Hak is deputy director of Charles University Environment Center in Prague. He is interested in indicators of sustainable development, quality of life and environmental sustainability, in particular. He participates in several international projects, collaborates with a number of local and international organisations (UNEP, SCOPE, EEA, etc.) and teaches these subjects at Charles University faculties.

Dimity Podger joined ESDinds as a Research Fellow following her doctoral work, which examined conceptual and practical contributions of a faith organisation to Education for Sustainability (EfS). Her research and consultancy work has included: the use of participatory action research and transformational pedagogies to assist academics transform curricula in the higher education sector and corporate staff to evaluate environmental education programmes; and the development of EfS learning experiences for youth and communities. She is Associate Lecturer at Macquarie University teaching courses in Action Research and sustainability to graduate students.

Georgia Piggot is currently completing her PhD in Environmental Sociology at the University of British Columbia. Prior to joining UBC, Georgia worked as a research officer for the ESDinds project described in this manuscript. Her contributions to this research included assistance with research design, and synthesis of theory and qualitative data, to produce a pool of indicators for field testing by civil society organisations.

Marie K. Harder is a Professor of Sustainable Waste Management at the University of Brighton, and coordinated the EU FP7 Research project entitled "The Development of Indicators & Assessment Tools for CSO Projects Promoting Values-based Education for Sustainable Development." She has considerable experience in applied research in waste and recycling across the physical and social sciences, working with businesses, government agencies, civil society and NGOs, leading the Waste & Energy Research Group (WERG). Through this broad sustainable development approach she has addressed issues in co-design, communities of practice, indicator design and community engagement in waste issues, as well as physical sciences. Marie started working at Fudan University, Shanghai from September 2011.