An Identification Key for Selecting Methods for Sustainability Assessments

Michiel C. Zijp 1,2,3,*, Reinout Heijungs 2,4, Ester van der Voet 2, Dik van de Meent 1,3, Mark A. J. Huijbregts 3, Anne Hollander 1 and Leo Posthuma 1

1 Department of Sustainability, Environment and Health, National Institute for Public Health and the Environment (RIVM), P.O. Box 1, 3720 BA Bilthoven, The Netherlands; E-Mails: Dik.van.de.Meent@rivm.nl (D.M.); Anne.Hollander@rivm.nl (A.H.); Leo.Posthuma@rivm.nl (L.P.)
2 Institute of Environmental Sciences, Leiden University (CML), P.O. Box 9518, 2300 RA Leiden, The Netherlands; E-Mails: Heijungs@cml.leidenuniv.nl (R.H.); Voet@cml.leidenuniv.nl (E.V.)
3 Department of Environmental Science, Radboud University Nijmegen (RUN), P.O. Box 9010, 6500 GL Nijmegen, The Netherlands; E-Mail: M.Huijbregts@science.ru.nl
4 Department of Econometrics and Operations Research, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands; E-Mail: R.Heijungs@vu.nl

* Author to whom correspondence should be addressed; E-Mail: Michiel.Zijp@rivm.nl; Tel.: +31-30-2744087.

Academic Editor: Marc A. Rosen

Received: 2 December 2014 / Accepted: 22 February 2015 / Published: 2 March 2015

Abstract: Sustainability assessments can play an important role in decision making. This role starts with selecting appropriate methods for a given situation. We observed that scientists, consultants, and decision-makers often do not systematically perform a problem analyses that guides the choice of the method, partly related to a lack of systematic, though sufficiently versatile approaches to do so. Therefore, we developed and propose a new step towards method selection on the basis of question articulation: the Sustainability Assessment Identification Key. The identification key was designed to lead its user through all important choices needed for comprehensive question articulation. Subsequently, methods that fit the resulting specific questions are suggested by the key. The key consists of five domains, of which three determine method selection and two the design or use of the method. Each domain consists of four or more criteria that need specification. For example in the domain “system boundaries”, amongst others, the spatial and temporal scales are specified.
The key was tested (retrospectively) on a set of thirty case studies. Using the key appeared to contribute to improved: (i) transparency in the link between the question and method selection; (ii) consistency between questions asked and answers provided; and (iii) internal consistency in methodological design. There is latitude to develop the current initial key further, not only for selecting methods pertinent to a problem definition, but also as a principle for associated opportunities such as stakeholder identification.

**Keywords:** sustainability assessment; sustainability question; identification key; indicator selection; method selection; aggregation; sustainability assessment tools

1. Introduction

The recognition that mankind puts major pressures on the earth systems has resulted in publications reporting on the results of sustainability assessments, e.g. [1]. Not only environmental crises, but also social inequities at local to global scales trigger a societal drive to position sustainable development as a decision making strategy [2]. There is a call for sustainability assessments from the local scale, such as sustainable development of cities and neighborhoods [3], to the global scale, for example the United Nations sustainability goals [4]; and from the product level, e.g., eco labels [5], to the sector level [6]. This demand resulted in a plethora of methods that claim to provide answers to sustainability questions. In fact, we have entered an era in which there is an abundance of methods for sustainability analyses. Some of these methods are complimentary, but there are also many competing methods. In the meantime, sustainability science is a swiftly developing discipline. There is an ongoing debate on what sustainability or sustainable development is, and what sustainability assessment should encompass, whilst there is a need to bridge widely diverging disciplines, each with their own definitions and approaches. In this important, complex and swiftly developing field, the selection of appropriate methods for answering a particular sustainability question can be challenging.

Whilst scientific approaches ideally are “fit for use” and robust, the selection of sustainability assessment methods is thought to be frequently led by the expertise of the analyst and available capacity [7,8]. The choice of method(s), however, requires choices on scope, assumptions, values and precision. One question regarding sustainable development can therefore yield a manifold of answers, depending on which assessment method is selected to answer the question [9,10]. If the underlying choices in method selection are not considered explicitly, there is a chance of a mismatch between the results and the context in which the question was asked. Also, the answer becomes more dependent on the analyst to whom the question is directed, than on a transparent specification of the question and methods logically linked to such a question. Fundamental improvements in the analysis of sustainability questions seem warranted, given these observations.

Looking at optimal improvements, literature shows that there are attempts to organize available sustainability assessment methods. These primarily should support the selection of appropriate method(s) to address the problem. First, lists have been drawn of sustainability methods, including reviews per method on objectives, strengths, weaknesses, et cetera [11,12]. Although useful to gain insight in all current optional sustainability assessments, such lists and descriptions provide limited
guidance to selection of a method. They are not designed to compare methods, but to provide information per method. Because of the amount of methods and their details, the lists contain a substantial amount of information. The second type of organization is that methods are described or scored based on a selection of criteria that are found to be important, such as object and spatial focus (e.g., [13,14]). Different sets of criteria are used for this purpose. An overview of these criteria, based on 26 studies, can be found in Table 1. The list of potential criteria based on which one can distinguish between methods is long and the possible directions per criteria vary. This approach provides a structured way to organize all the information gathered per method. Still, it is a lot of information to process, since a comparison of all criteria for all methods is required in order to select the most suitable method for a question. As variant of the importance scoring, and to reduce the amount of information per method, some authors appoint a selection of their list of criteria as key-criteria. The basis for this selection, however, is often not found in the manuscripts or supporting information, e.g., [14–16]. An exception is found in Udo de Haes et al. [17] who, for the purpose of comparing Life Cycle Assessment (LCA) and Ecological Risk Assessment (ERA), give an argumentation per criterion why it is either a fundamental or secondary criterion. The third approach to organize methods is to categorize or frame them, based on a selection of features of the methods. Some much cited attempts of categorization can be found in literature (such as [17]). Although categorization might be valuable to gain a quick insight in the type of methods available, the approach provides limited support for method selection. Firstly, because the loss of information: the variability between methods is too large to be captured in three or four features. Secondly, because methods themselves are flexible and can often be applied in various modes for different purposes. This struggle with complexity and flexibility can for example be seen in the frequently cited categorization of Ness et al. [18]. Their framework provides a categorization based on, amongst others, the object of the sustainability analysis, but it has only room for the objects “spatial unit” and “products”. Further, various methods positioned in the framework would also fit on other places in the framework. For example, LCA could also fit under “integrated indicators” and all the methods under the integrated “prospective” categorization fit under “retrospective” as well. Other attempts for categorization frameworks [7,8,16,19] show the same type of struggles with existing variability and complexity of sustainability assessments. They are useful for a quick overview on the type of assessment methods that are available, but provide insufficient detail for method selection.

Scrutiny of the aforementioned approaches shows that they are often supply-driven. Methods used in sustainability assessments are organized based on the articulation of the available methods. Ideally, however, the selection of a method for answering a specific sustainability question would be performed based on a specific analysis of the assessment problem, and a subsequent specific articulation of the research question (demand-driven), which would result in explicit choices for assessment method(s). Some studies do experiment with an analysis of the assessment demand as organizing principle. For example, Wrisberg et al. [13] state that every demand has its specific object of analysis, spatial and temporal dimensions, required level of detail and required level of integration, which should be leading in choosing a method or a combination of methods. Furthermore, they distinguish five type of contexts (strategic, capital investment, design and development, communication and marketing and operational) and eight context specifications with each their specific profile of demands for method selection. Another example is from De Ridder et al. [19], who analyzed which phase in integrated assessment frameworks, like Strategic Environmental Assessment, asks for which type of method. When linking demand with
supply, both De Ridder et al. [19] and Wrisberg et al. [13] show that the context (e.g., the phase in the management cycle) does not seem to be leading in method selection, but rather in method design. The context specifies the role and design of a selected method, such as the thoroughness of the analysis or the way results are presented, and not the method selection itself [20].

Many examples exist in which method selection is based on explicit qualitative comparison (description of strengths and weaknesses) or characterization. These approaches are, however, case specific and do not cover the extent of relevant criteria and methods to choose from, e.g., [9,10,21–23].

Recent literature on novel ways to support method selection in the field of sustainability assessment is scarce. A recent innovation is described by Gasparatos and Scolobig (2012), who provide four different proposals to base method selection on [7], namely based on:

1. the perspective of the assessment (e.g., biophysical limits or human wellbeing);
2. desired features of the assessment (e.g., spatial or temporal focus);
3. the acceptability criterion of Pope et al. (2004) [24] (e.g., is the goal of the assessment to reduce impacts or to reach explicitly defined sustainability goals?);
4. values of the stakeholders (e.g., focus on general human well-being, personal well-being, or ecosystem well-being).

Given apparent limitations of current approaches in sustainability assessment method selection, we propose a next step. A next step in this field would be a framework for method selection that takes into account all four proposals of Gasparatos and Scolobig (2012) [7], and possibly more, but that is also capable of capturing the dynamics and complexity of both the question (demand) and method (supply) side of sustainability assessments. We propose to collate all aforementioned approaches and options into an identification key, inspired by e.g., Flora’s for plant determination. We expected that and evaluated whether this could be the step forward into systematic and transparent method selection. We expect that the provision of a Sustainability Assessment Identification Key (SA-IK) is of general support to:

(i) identify a method based on explicit choices and all methods available and not necessarily the well-known method by the analyst;
(ii) guide method selection from demand perspective (articulation of the question) rather than supply perspective;
(iii) report the results of the assessment referring to the explicit choices made with question articulation, making results easier to understand, interpret and compare with other assessments.
(iv) make method selection transparent and reproducible

Before these features can be substantiated, such an identification key had to be designed and tested. In this context, the aims of this study are:

(i) to confront the available assessment methods with the sustainability questions posed by society such as to propose a new organizing framework for selection of sustainability assessment methods: the sustainability assessment identification key;
(ii) to present the design of the sustainability assessment identification key;
(iii) to show how the sustainability assessment identification key (SA-IK) works.
It should be noted that the emphasis of this study, the SA-IK design and its application, mainly considers environmental assessment methods, focusing thus on the “planet” aspect of sustainable development. Further, we acknowledge that sustainability assessments may reflect subjective views of researchers on the definition of sustainability, and that sustainability assessments can differ due to that. Independent of such views, the SA-IK was designed to support clear and transparent support in choosing and applying methods. When the SA-IK is potentially used by various types of users, those users are expected to specify how sustainability is defined by them for each case specifically, transparently and explicitly.

2. Methods: Development of a Sustainability Assessment Identification Key

2.1. Terminology Used in This Article

The scientific literature in which sustainability assessment methods are being described and characterized shows a plethora of terms such as instrument, tool, method, methodology and procedure, which are more or less interchangeable, but having different meanings in different articles. In practice a certain degree of hierarchy can be discovered between them [15], see Figure 1. The definitions used in this paper are described below.

<table>
<thead>
<tr>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stepwise approach consisting of different phases of a process which provides the link with the societal and decision context (e.g. Strategic Environmental Assessment, Integrated Sustainability Assessment).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>The rationale and the structure for the integration of concepts, methodologies and methods (e.g. Sustainability Science conceptual framework, Life Cycle Sustainability Assessment, …).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>A collection of individual consecutive and/or complementary sub-methods. Methods can include 1) indicator selection, 2) quantification or qualification and 3) aggregation, but not necessarily all three aspects (e.g. Life Cycle Assessment, Risk Assessment, Dow Jones Index, …).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sub-methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>A method to agree on, estimate, calculate or aggregate results (e.g. delphi method, fate model, weighting).</td>
</tr>
</tbody>
</table>

- **Tool** is a brand that executes a method (software, application, …).
- **Indicator** is a parameter, or a value derived from parameters, which points to, provides information about, or describes the state of a phenomenon, with a significance extending beyond that directly associated with its value (OECD, 2003). The parameter could be quantitative or semi-quantitative or qualitative derived from a method or sub-method.

**Figure 1.** Terminology and their hierarchical relation adopted in this paper. Adjusted from Sala et al. [15].

A method is defined as a collection of consecutive and complementary sub-methods with which a specific question can be answered. Examples of methods are Life Cycle Analysis (LCA), Ecological Footprint (EF), EMergy analysis (EM), but also indexes like the Dow Jones Sustainability Index. Consequently, sub-methods are the consecutive and/or complementary analytical steps a method consists
of. For example methods with which indicators can be quantified or with which results can be aggregated. The distinction of methods and sub-methods is important, because in practice methods do not necessarily have to be applied as a whole. For example, LCA contains a sub method with which emissions and the use of resources can be translated into impacts: the Life Cycle Impact Assessment (LCIA). This sub method can (and is e.g., [25]) also be used to determine the impact of activities without taking into account the life cycle of that activity, thus without taking the LCA method into consideration as a whole.

Methods are the operationalized parts of a higher-level entity which we refer to as the framework in which the sustainability assessment is undertaken. A framework is the way sustainability is conceptualized. For example: Which pillars (that is: people, planet, prosperity) are thought to be important? How do they relate to each other? And: How are interactions between impacts at different spatial and temporal scales envisioned?

The framework itself is again part of wider context, which we refer to as procedures. Procedures consist of subsequent phases in a process of making decisions and in policy, given the existence of a societal problem. Methods can have varying roles in different phases of a procedure [19].

While this study focuses on the selection of methods, also the other levels (framework and procedure) play an important role. Namely, to select and especially to design a method tailored to the problem definition, the place in the procedure and the framework envisioned by the user must be made explicit and thus known. In other words, choices that are made in sustainability assessments regarding method(s) choice are co-influenced by the contexts of the framework and procedure within which problems can be addressed.

2.2. Review on the Derivation of an Identification Key in General

Identification keys exist in many fields, e.g., biology, medicines, social sciences and information architecture. These disciplines have in common that they intend to classifying large amounts of objects, at representing a complex system of characteristics, at easy adjustment when new insights or findings are added to the collation of objects, and that are—finally—of use to practitioners. In biology and other fields the result of an identification key is often referred to as a taxonomy of the objects under study. Although relevant for various disciplines, the available scientific literature on taxonomy development is limited [26]. Three types of processes towards a classification, or taxonomy, have been distinguished. First, there is the conceptual approach in which a classification system is drafted based on a conceptual model or idea of the field of interest in a deductive manner and subsequently improved or tested based on e.g., empirical data [27]. The second approach starts with empirical data and builds a classification system based on statistical methods (ibid). Thirdly, a classification system can be based on interaction with users of the information that needs to be classified, which is also named a “folksonomy” [28]. Expertise and experience on taxonomy in the fields of biology and knowledge engineering shows that there is not something like a perfect or ideal taxonomy for a field of interest. Complex information can be organized and unlatched in many ways. For example, in ecology, organisms are being characterized based on either physical characteristics (phenetics) or on evolutionary relationships (cladistics), resulting in significantly different groupings of organisms, but both function well as classification systems. Another lesson to be learned from taxonomic experiences in other fields is that designing a comprehensive,
widely supported identification key can take years [29,30]. Deriving a meaningful identification key is an iterative process (Figure 2). This will also hold for sustainability assessment, where basic principles of the field are still much debated [2,15,31]. Taking that in mind, a first design of an identification key, like presented in this paper, should be seen as an essential, but small step towards a complete and widely applicable and applied key.

![Figure 2: The identification key development method based on Nickerson et al. [26].](image)

Where taxonomy is the art of organizing and unlatching available information, the identification key we intended to design should also deal with information that is not yet available. Namely, for some (or perhaps many) sustainability questions, a method to derive the assessment outcomes does not (yet) exist. Because potentially a large group of sustainability assessment methods are not known yet, a conceptual approach is followed to build the sustainability assessment identification key and not a statistical approach. Nickerson et al. (2013) [26] provide a taxonomy development method, which we adjusted (Figure 2) and used for the development of the Sustainability Assessment Identification Key (SA-IK).

2.3. Step 1: Identify Criteria

As already mentioned, Table 1 presents an overview of the criteria used in 27 literature sources that focus on method selection, both from the supply perspective and the attempts to include the demand perspective. The column with references in Table 1 shows the difference in criteria sets that are thought to be important by the different sources. For the purpose of presentation criteria that consider similar concepts were grouped together under a single criterion and assigned a description, a step improving clarity, but removing nuance. For example, the criteria “values/view on sustainability” in Table 1 includes the choice of weak versus strong sustainability (see Table 1 and Table S1), the choice of world view (e.g., anthropocentric or ecocentric) and risk perception (e.g., Cultural Theory). In the identification key, these different aspects of values can be dealt with separately.
Table 1. List of criteria derived from literature sources and classified in different domains, with a short description and references per criteria.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Explanation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain: System boundaries/Inventory</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Object | What is the object of the assessment?  
Is it a physical object (product, chemical, process), or an organization, a region, a policy measure, an activity, etc... | [12–14,17,18,32] |
| Spatial focus | What is the spatial focus of the activity?  
Is the activity assessed on micro or macro scale, and if on macro on local, regional or global scale? | [8,12,13,17,18,32,33] |
| Temporal focus | What is the temporal focus of the assessment?  
Is the activity assessed retrospective, prospective or does a snapshot suffice? | [7,8,16,18,24,32] |
| Life cycle thinking | Which parts of the life cycle or supply chain are included in the assessment?  
Only one phase, the whole life cycle, or something in between? | [8,15,17,34] |
| **Domain: Impact Assessment/Theme selection** | | |
| What is to be sustained | What is to be sustained?  
Are these environmental, social, economic and/or institutional endpoints? | [8,16,18,32,33,35,36] |
| Theme and indicator selection | Which themes are selected?  
Is the method transparent in the selection and use of indicators? What place on the cause effect chain do the indicators have? etc. | [12,14,17,35,37] |
| Spatial focus of impact | What is the spatial scale of the impacts that should be taken into account?  
Does the assessment include intra-generational impacts? Or in other words: does the assessment aim at internal or external sustainability. Impacts at what scale are taken into account? Are they site-specific/dependent or independent? | [13,16,17,34,37] |
| Temporal focus of the impact | What is the temporal scale of the impacts that should be taken into account?  
Does the assessment include inter-generational impacts? What time-frame should be included for the impacts? | [7,16,33,36,38] |
Table 1. Cont.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Explanation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain: Aggregation/Interpretation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sustainability target</td>
<td>Is a sustainability target necessary? If the goal is to compare alternatives, to perform a hotspot analysis or to improve an object, a sustainability target is not essential. If the goal of the analysis is to determine the sustainability of an object, a target is required. This is also referred to as direction to target (no target needed) or distant from target (target needed); and assessment impact-led (least impact, no target needed), objective-led (best positive contribution, no target needed) or assessment for sustainability (like the other two, but in relation to a specific sustainability target)?</td>
<td>[7,8,16,24,34]</td>
</tr>
<tr>
<td>Values/View on sustainability</td>
<td>What view on sustainability should be leading in the assessment? Is sustainability understood as weak, strong or partly substitutional? In short: weak means that various capitals are interchangeable. Strong means that each capital should be preserved independently. Partly substitutional means weak until a critical level is reached, e.g., Critical Natural Capital (CNC) or planetary boundary. Also one’s world view (personal believes or risk perception) can influence the assessment.</td>
<td>[7–9,12,24,36,38–43]</td>
</tr>
<tr>
<td>View on integration of pillars</td>
<td>How should aggregation of information from different disciplines take place in the assessment? In a multi (separate), inter (connected) or trans (combined/holistic) disciplinary way?</td>
<td>[7,8,15,16,18,36,38]</td>
</tr>
<tr>
<td>Normalisation/weighting/aggregation method</td>
<td>Which aggregation level is preferred and which methods are used? Both normalisation (make data comparable), weighting (specify interrelationships) and aggregation (get functional relationships) need careful consideration.</td>
<td>[9,12,17,33–36]</td>
</tr>
<tr>
<td><strong>Domain: Method Design</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>View on stakeholder involvement</td>
<td>Who should be involved in the assessment in which way? Also referred to as legitimacy, in relation to indices or composite indicators.</td>
<td>[7,8,19,33,36,38,40,44]</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Explanation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context of the assessment</td>
<td>How and by whom are the results used?</td>
<td>[8,13,14,19,32,33]</td>
</tr>
<tr>
<td></td>
<td>In which (phase of a) procedure are the results of the assessment used?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Is the goal of the measure: decision making and management, advocacy,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>participation and consensus building or research and analysis? Or is it a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>strategic, capital investment, design and development, communication and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>marketing or operational question?</td>
<td></td>
</tr>
<tr>
<td>Uncertainties</td>
<td>How are uncertainties to be handled?</td>
<td>[7,8,13,16,33,36,45]</td>
</tr>
<tr>
<td></td>
<td>Salience, credibility and variability? Should an uncertainty, sensitivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and/or perturbation analysis be included?</td>
<td></td>
</tr>
<tr>
<td>Domain: Organisational</td>
<td>Should the method be formally recognized?</td>
<td>[13,33]</td>
</tr>
<tr>
<td>restrictions</td>
<td>ISO, EC, etc.</td>
<td></td>
</tr>
<tr>
<td>Formal requirements</td>
<td>Is there capacity for hiring expertise?</td>
<td>[12,13,36]</td>
</tr>
<tr>
<td>Expertise requirements and</td>
<td>Expertise requirements and availability</td>
<td></td>
</tr>
<tr>
<td>availability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software requirements and</td>
<td>Is there capacity for acquiring software?</td>
<td>[12,13]</td>
</tr>
<tr>
<td>availability</td>
<td>Software requirements and availability</td>
<td></td>
</tr>
<tr>
<td>Data requirements and</td>
<td>Is there capacity for gathering data?</td>
<td>[13,33,34,46]</td>
</tr>
<tr>
<td>availability</td>
<td>Data requirements and availability</td>
<td></td>
</tr>
</tbody>
</table>

2.4. Step 2: Assign the Criteria to Domains

For the design of the SA-IK we distinguished three domains that determine the method selected for a specific sustainability question, and two that determine the further design and use of the method. The first domain deals with the system boundaries of the activity under consideration or, in other words, the specification of the inventory (in LCA terms) or system quantification (in Material Flow Analysis terms) or Drivers and Pressures (in DPSIR terms; DPSIR is further explained in Section 2.6). It contains question articulation on, amongst others, the type of object, the spatial scale and other criteria that determine the system boundaries of the assessment. This first domain is further referred to as “System boundaries/Inventory”. The second domain: “Impact assessment/Theme selection”; articulates the type and scope of the impacts. For example, which themes or issues are thought to be important? Is the focus on environmental issues, or also on social, economic and institutional issues? And: is the focus on impacts on one specific location or impacts worldwide? The third domain articulates the need and specification for aggregation of the results and is named: “Aggregation/Interpretation”. The idea to distinguish these three domains for method selection is based on the observation that sustainability assessments methods in general consist of these three elements. For example, in LCA the activity and its resulting emissions and resource uses are quantified based on a set of boundary conditions in the inventory phase, which is followed by the impact assessment phase in which the emissions and resources used are expressed in themes (impact categories) that are thought to be important. Finally, choices are
made on if and how the results per theme need to be aggregated, e.g., by weighted summation. Another example of this triptych (1. system boundaries/inventory, 2. impact assessment/theme selection, 3. aggregation/interpretation) is (societal) cost-benefit analysis (CBA): CBA starts with defining the policy alternatives in a given situation (1. system boundaries), followed by translating these activities and its results in costs and benefits for different stakeholders (2. impact assessment). Then, choices are made in how the results are presented (aggregated or not) and weighted (3. aggregation and interpretation).

The domains are chosen such that the Criteria found in literature (step 1, paragraph 2.3) occur in just one of the domains. For example, the role of personal “values” for method selection (is ones view on sustainability strong, weak or partially compensatory) are influential for the choice of aggregation method, but less on the choice of system boundaries and theme selection.

Some of the Criteria derived from literature (Step 1, paragraph 2.3) do not determine the selection of methods, but rather steer the design or use of a method, e.g., the inclusion of uncertainty analyses or stakeholder involvement. For these Criteria we added the domain “method design”. Other Criteria are organizational restrictions provided by other influences then the question or the context of the use of its results. We refer to these Criteria as organisational restrictions. Table 1 groups the Criteria per Domain. Further details can be found in the Supporting Information.

2.5. Step 3: Build the Identification Key

Based on the Domains and Criteria specified above, a first design of the identification key for sustainability assessment was derived. Conceptually, the drafted SA-IK consists of a key that focuses first on the three domains System boundaries/Inventory, Impact assessment/Theme selection and Aggregation/Interpretation. When the Criteria in these Domains are addressed, the SA-IK delivers as main outcomes an articulated question and suggestions for method(s) selection per domain. That is, methods that fit the chosen boundary conditions, methods that fit the chosen themes and methods that fit the specifications regarding aggregation. The outcomes of these three keys are then collated and analyzed, yielding one selected method, a (new) combination of (sub) methods or the conclusion that a method for the specified question is not (yet) available in the SA-IK. The selected method(s) can be further designed with the help of the subsequent Method design key, for example to add a sensitivity analysis. Moreover, the Criteria labeled “organizational restrictions” (pragmatic constraints, like time and data availability), can influence the methods choice and design, but are left out of the scope of this first design.

As an example, Figure 3 visualizes the identification key for the part focusing on System boundaries/Inventory. The first question in the identification key is: what is the Object of the analysis? This can be products, geographic units, companies, et cetera. The follow-up question depends on the answer given. For example, the first question for the Object “products” could be: “Does the situation concern (a) single product(s) or (a) product group(s)?” whereas the first question for the Object “geographic unit” could be: “what is the spatial focus of the activity?” and so on. The identification key for the domain System boundaries/Inventory can be viewed as example in the Supporting Information.
2.6. Note on Theme Selection

The previous step results in a systematic evaluation of choices and evaluations relevant for selection of a method for sustainability assessment. Next to literature on method selection, on which our analysis and IK primarily focuses, an even larger amount of literature exists on theme and indicator selection. Themes are the issues that could be considered to take into account when looking at the sustainability of an object. Indicators are parameters that provide information about (or describes the state of) these themes [47]; see also Figure 1. For example, concentrations of nutrients in fresh water (the indicators) can be an indicator for ecological damage due to eutrophication (the theme). Indicators can be chosen on different places at the cause-effect chain. The DPSIR framework is often used to describe the relation between an indicator and the Driver (the activity), the Pressure (e.g., the resulting emissions), the State (e.g., the concentrations in fresh water), the Impact (e.g., on biodiversity) and the Response (e.g., policy measure or monitoring) [48], mostly in environmental assessments, but also broader [41,49]. Niemeijer et al. [50] provide an overview of criteria found in literature for indicator selection, which will not be repeated here, but which could be useful for a themes and indicators identification key, to be used following the SA-IK design. Although theme selection and method selection are closely related and always needed both, they are often treated as separate entities. This results in two approaches for designing a sustainability assessment. The first approach is that one or more methods are chosen, followed by theme selection. The second approach is that first themes are chosen, followed by a selection of methods with which these indicator representing the themes can be quantified. The drawback of the first approach is that existing methods are often limited in that they represent a current set of specific themes (and not: all potentially relevant themes), and thus the choice of method narrows down the options of considering potentially relevant themes (when uncritically applied). The second approach starts with theme selection and gathers methods able to quantify the selected themes. Though this
sustainability suggests a problem-driven choice of theme, and thus relevance of the final results, this working order also has potential drawbacks. Firstly, there is the probability that the results for the various themes are less comparable (e.g., life cycle based and not life cycle based; site dependent versus site independent impacts). Secondly, one runs the risk of applying methods in a wrong matter (e.g., linear extrapolation of LCA results from micro to macro level) to be able to compare the results of the different themes. Thirdly, there is the risk of double-counting when aggregating the results (e.g., both the “Pressure” and the “Impact” as indicator). Scientific literature reports many examples of both alternative approaches: method selection is leading versus theme selection is leading.

3. Results: Examples of How the Identification Key Works and What Type of Problems It Solves

3.1. Example of Sustainability Assessment Identification Key Application

Next to designing SA-IK, we aimed to test and improve it iteratively. The key we present has been subject to this iteration, and the final SA-IK (step 5 in Figure 2) was used to illustrate its use and usefulness, according to our third study aim.

First, we explored how systematic use of the SA-IK in the phase of sustainability assessment question articulation can result in method selection, based on the first design of the sustainability assessment identification key (SA-IK). An example of the results of this exploration is provided in Table 2. The table show three ways to articulate a realistic societal problem. Thus, three teams start with the same general question, but refined question articulation guided by SA-IK leads them to three different specific questions and thus three different methods selected. The focus of the example is not on why certain choices are made, but on what the consequences of the choices are for the method selected and thus the type of answer provided. The SA-IK provides the elements that require a choice and a general list of possible answers per choice. For example, the question in Table 2 is: “How sustainable is our food pattern?” The first choice provided by the SA-IK is: What is the object? This could be products consumed, but also lifestyles, the food sector, food policy, or a certain geographic unit. This choice, and its possible directions, are discussed and decided on by the actors responsible for the question articulation. The following choice depends on the answer to the previous choice. For example: the object “product” should be further specified in “single product(s) or product groups?”; a question that is not relevant for the object “geographic unit”. In that way, the SA-IK guides its user through explicit choices regarding all criteria in the System boundaries domain and then provides a list of methods that fit the choices made. The Impact assessment/Theme selection and Aggregation/Interpretation domains are specified similarly followed by comparison and selection of the method(s) that the SA-IK found suitable for answering the articulated question.
Table 2. Three examples on how the Sustainability Assessment Identification Key (SA-IK) applied to an apparently singular question leads to different method choices, given specifications identified by explicitly addressing SA-IK Domains and Criteria. The examples start with the same question, but they follow different contextual—and therefore assessment articulation—pathways, leading to different questions to be answered by different methods.

<table>
<thead>
<tr>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question</strong></td>
<td>How sustainable is our food pattern?</td>
<td>How sustainable is our food pattern?</td>
</tr>
<tr>
<td><strong>Sub Identification Key on System boundaries/Inventory</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>What is the object?</strong></td>
<td>Products</td>
<td>Products</td>
</tr>
<tr>
<td><strong>Single product(s) or product group(s)?</strong></td>
<td>Single products</td>
<td>Single product(s) or product group(s)?</td>
</tr>
<tr>
<td><strong>Should the product(s) life cycles be included?</strong></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Which part of the life cycle?</strong></td>
<td>Cradle to grave</td>
<td>Upstream</td>
</tr>
<tr>
<td><strong>What is the spatial focus of the activity?</strong></td>
<td>Local</td>
<td>Regional</td>
</tr>
<tr>
<td><strong>What is the temporal focus of the activity?</strong></td>
<td>Snapshot</td>
<td>Snapshot</td>
</tr>
<tr>
<td><strong>Results sub IK System boundaries/Inventory</strong></td>
<td>Life Cycle Inventory</td>
<td>Input Output Analysis, Material Flow Analysis, Substance Flow analysis, …</td>
</tr>
<tr>
<td><strong>Sub Identification Key on Impact assessment/Theme selection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>What is to be sustained?</strong></td>
<td>Environment</td>
<td>Resources</td>
</tr>
<tr>
<td><strong>Which location on the cause effect chain is required?</strong></td>
<td>Impact at endpoint</td>
<td>Pressure</td>
</tr>
<tr>
<td>Example 1</td>
<td>Example 2</td>
<td>Example 3</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Question</strong></td>
<td>How sustainable is our food pattern?</td>
<td>How sustainable is our food pattern?</td>
</tr>
<tr>
<td><strong>Sub Identification Key on Impact assessment/Theme selection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Select themes</strong></td>
<td>Climate change, acidification, eutrophication</td>
<td>Economy, energy and material use</td>
</tr>
<tr>
<td><strong>Results sub IK Impact assessment/Theme selection</strong></td>
<td>Endpoint Life Cycle Impact Assessment (LCIA) method</td>
<td>Input Output Analysis and Material Flow Analysis</td>
</tr>
<tr>
<td><strong>Sub Identification Key on Aggregation/Interpretation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>What type of analysis is required?</strong></td>
<td>Direction to target</td>
<td>Direction to target</td>
</tr>
<tr>
<td><strong>What type of analysis is required?</strong></td>
<td></td>
<td>Direction to target</td>
</tr>
<tr>
<td><strong>Which level of aggregation is required</strong></td>
<td>Capitals</td>
<td>Total</td>
</tr>
<tr>
<td><strong>Which level of aggregation is required</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>What is the view on sustainability</strong></td>
<td>Ecocentric</td>
<td>A Multi Criteria Analysis (MCA) like weighted summation or Multi Attribute Value Theory (MAVT)</td>
</tr>
<tr>
<td><strong>Result sub IK aggregation</strong></td>
<td>LCIA endpoint damage method</td>
<td>Result sub IK aggregation</td>
</tr>
<tr>
<td><strong>Match of sub IKS → method selection</strong></td>
<td>Life Cycle Assessment with Endpoint LCIA method (e.g., ReCiPe)</td>
<td>Match of sub IKS → method selection</td>
</tr>
<tr>
<td></td>
<td>Material Flow Analysis and Input Output Analysis aggregated with MCA, e.g., MAVT</td>
<td></td>
</tr>
</tbody>
</table>
3.2. Confronting Sustainability Assessments in Scientific Literature with the Identification Key

Next to illustration of hypothetical uses yielding vastly different methods for a single societal question (previous paragraph), we aimed to test the SA-IK for a suite of selected studies. We did that retrospectively. The goal of this analysis was to show whether and how the SA-IK may improve sustainability assessments on expected benefits mentioned above.

The case studies were selected by a two-step procedure: (1) a literature search from 2011 and beyond (Search engine: Scopus; search key: TITLE-ABS-KEY(“Sustainability assessment” OR “Sustainability evaluation” OR “Sustainability performance”) AND PUBYEAR > 2010); and (2) screening of titles and, in a second round, abstracts to sub-select the manuscripts that claim to describe a case study on sustainability assessment. Step 1 resulted in 1086 results and step 2 in a selection of 30 manuscripts with pertinent case studies (see Table S2). The introduction (goal/scope etc.), method description and results of the case studies were compared to the Domains and Criteria distinguished in the SA-IK. This analysis showed that the SA-IK potentially provides improvements in three directions: (a) more transparency in the link between the question and method selection, which is lacking or scanty available in most of the case studies; (b) more consistency between question and answer; and (c) more consistency in methodological design. These three conclusions are substantiated below.

3.2.1. Transparency on Method Selection

(a) Evaluation on Method selection. In 14 out of the 30 selected case studies, reasons for selecting a method are made explicit. Often one or two Criteria are mentioned as reason to select a method. As an example, considerations on “what is to be sustained” was mentioned as Criterion for method selection in 7 of the 30 case studies (Figure 4). Figure 4 shows that most case study descriptions revealed none or only little attention to method selection. This does not directly mean that methods were not carefully selected. It means that the relation between the question and the method selection was not explicitly described in the manuscripts (leaving room for variability in providing answers as illustrated in Table 2). When criteria are given attention in the manuscripts they might also be taken into account in method selection, also when this relation is not explicitly described as such. For example, all case studies described the Object of study, but none of them explicitly brought this in relation with the method to be selected. This is visualized in Figure 4. It shows the percentage of case studies in which the Criteria of the SA-IK are articulated (in blue) and the percentage of case studies in which the Criteria are explicitly brought into relation with method selection (in red).

The Criteria within the domain “System boundaries/Inventory” (object, spatial focus, temporal focus and life cycle or chain) are most frequently articulated in the case studies. For example, all case studies describe the Object under investigation. However, the temporal focus of the activity is discussed in less than 50% of the studies. In other words: it seems that in more than 50% of the case studies no explicit choices are made on the temporal focus of the activity under consideration. Of the criteria in the domain “Impact assessment and Theme selection”, most (93%) case studies discuss what is to be sustained. However, the spatial and (especially) the temporal focus of the impacts assessed are often not discussed. Aggregation seems to play a role in 50% of the case studies, but the details of the aggregation, e.g., “Is a sustainability target required?” and “What is the view on sustainability?”, are often not made explicit.
Figure 4. Percentages of case studies in which Criteria were explicitly described (blue), percentage of case studies in which Criteria were explicitly coupled to method selection (red) and percentage of case studies in which consideration concerning this Criteria were not mentioned (blank).

3.2.2. Prevent Inconsistencies between Introductions/Case Descriptions and Method Selection

(b) Evaluation on consistency between question raised and answer provided. Lack of specificity in question articulation and in making tailored and transparent choices was hypothesized to lead to a mismatch between the question asked and the answers given. We analyzed in the same thirty studies whether the mismatch occurred. Four type of questions were distinguished: (1) determine how sustainable an object is; (2) compare alternative objects; (3) perform a hotspot analysis (which part of an object has the highest positive or negative impact on sustainability?); and (4) improve an object. A noticeable percentage of the case studies (6 out of 30, 20%) showed a mismatch between the type of question described in the case description and the answer provided as a result of the method chosen. This exemplifies that missing transparency in the step between “a question to be answered” and “the methodological design to answer the question” might lead to “a question not answered”.

3.2.3. Prevent Inconsistencies in Methodological Design

(c) Analysis on inconsistencies in methodological design. We hypothesized that lack of question articulation can lead to different specifications for different indicators within one assessment. For example, are themes quantified based on comparable system boundaries? More specifically: when a life cycle approach is thought to be important, is it incorporated for all indicators or only a selection? Life Cycle Analysis conceptually focuses on environmental impacts of a product, but social (Social-LCA) and economic (Life Cycle Costing) aspects can be performed from a life cycle perspective as well. However, some case studies that do include environmental, social and economic themes only perform a Life Cycle Analysis for the environmental ones and not for the social and economic themes. Another finding considers inconsistencies for the spatial scale. Often, within one study, world-wide environmental impacts are taken into account, whereas economic and social impacts are taken into account on the organizational or regional level (e.g., [51–53]). Apparently, different capitals (People, Planet and Prosperity) tend to results in different spatial scopes. These inconsistencies were unexplained, such that SA-IK improvements can be gained here, though we also note that the observed inconsistencies are not necessarily wrong; we did not analyze the impacts of these inconsistencies on the results. However, in terms of question articulation the observed inconsistencies are remarkable. Logically, one view on the scope would be expected, e.g., one spatial focus, or otherwise an explanation for indicator-specific scope definition. Probably, some choices are not made explicit, leading to these inconsistencies in method design. Of the 30 case studies, 10 studies showed one or more of these types of inconsistency. With the SA-IK the preferred scope can be defined and compared with available indicators to choose from.

4. Discussion and Conclusions

Method selection is a crucial step between a sustainability assessment question raised and an answer given. By choosing a method, one selects, or at least narrows down, the choice in system boundaries, themes and personal values. Therefore, method selection should be based on careful articulation of the original (societal) question. This is especially true in the field of sustainability assessments, with its manifold of available methods and interpretations of what sustainability or sustainable development actually means. Expanding on existing approaches, we observed unnecessary unclearness in the general literature on this subject, suggesting that a novel approach could be valuable for transparent, reproducible and valid method selection. The available organizing approaches (describing, characterizing and categorization) were analyzed to be largely supply-driven instead of demand-driven and appeared not to be capturing the dynamics and complexity that is needed for guiding the selection of methods for sustainability questions. This manuscript encompasses a plea for the design and use of a sustainability assessment identification key (SA-IK) as that next step. A SA-IK provides a modular approach that helps to structure question articulation and that leads to demand-based method selection. Functioning foreseen similar to a flora key for the taxonomy and identification of plant species, this identification key was designed to support using and categorizing large amounts of information, and to present the information needs in a step by step manner, which makes it manageable. The SA-IK helps those responsible for the
Sustainability question and the sustainability assessment, ideally supported by the relevant stakeholders, to select a method based on explicit (and as appropriate: improved) articulation of the question.

Deriving an identification key that makes sense has proven to be a worthwhile, but extensive process; a process that can take years. This will especially be true in a relative young field like sustainability assessment, where there is still much debate on terminology and data interpretation. On the other hand, as a side effect, developing the identification key might contribute to clarifications in this field by clarifying the relation between choices in definitions/interpretations and the consequences for the assessment.

Given existing sustainability assessment studies, as well as some principles for designing a taxonomical key, we provide a first attempt to design a Sustainability Assessment Identification Key (SA-IK), and use it in various ways. In other words, a first iteration (Figure 2) of developing using and adjusting the SA-IK is provided. Based on examples, we have shown that although the design is incomplete and needs further development, the use of the SA-IK is supportive to:

(a) guide and make explicit choices in method selection and design, revealing assumptions that remain hidden in many studies;
(b) yield a better understanding of the question raised and how the question guides method selection;
(c) enable a more robust interpretation of the results, because the results can be placed in the context of methodological choices;
(d) producing eventually more transparent and reproducible assessments;

Furthermore, the SA-IK can provide insight in which type of question cannot yet be answered with the existing plethora of methods.

The proposed design is based on the observation that all sustainability assessments constitute three steps: (i) System boundaries/Inventory; (ii) Impact assessment/Theme selection; and (iii) Aggregation/Interpretation. Most Criteria for method selection found in literature have a role in only one of these three steps. The SA-IK itself consists of questions that can be answered for almost any problem definition, the answers guiding the assessors in various relevant directions. A single question can, depending on contextual aspects of the problem definition, result in different methodological choices, and a suite of questions can be analyzed systematically, such that the quality of sustainability assessments can be improved as compared to the current (reported) practices.

The SA-IK was not designed to provide answers to sustainability questions, but should serve in transparent and pertinent sustainability assessment method selection as such. Also, the SA-IK does not prescribe what sustainability assessment is and what sustainability assessment should encompass, but is designed to make these choices case specific, with all the relevant stakeholders, in a transparent reproducible and explicit way. The key reveals the consequences of choices for method selection, but does not prescribe these choices. Efforts to find consensus regarding the definition of sustainability and sustainability assessment exist, for example the development of the Bellagio STAMP principles [45], as do attempts to describe the ideal sustainability assessment method [15]. These were taken into consideration for the first design of the SA-IK.

Thirty case studies on sustainability assessment that were recently published in literature were evaluated based on the SA-IK. The analyses showed that using the SA-IK makes many hidden choices explicit, but also reveals inconsistencies, which would have been avoided had the SA-IK been used. In 6 of the 30 case studies, limited question articulation appears to have led to a mismatch between the
type of questions asked and the type of answer provided and in 10 of the 30 case studies to an inconsistent method design. Hence, the SA-IK use in its current format is potentially helpful in improving sustainability assessments.

We expect that the iterative process of using, discussing and further developing the identification key will at least lead to more transparency in method selection and potentially also to a better match between questions asked and answers provided.

Supplementary Materials


Acknowledgments

Funding by the Strategic Research Program of RIVM is gratefully acknowledged (project number S/607020).

Author Contributions

Michiel Zijp, Leo Posthuma, Reinout Heijungs and Dik van de Meent generated the idea of the SA-IK. Michiel Zijp led the development of the SA-IK and writing the manuscript, both to which the other authors contributed equally. All authors have read and approved the final published manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).