Agroecology TPP

TRANSITIONS

Developing holistic assessments of food and agricultural systems

A meta-framework for metrics users

Christine Lamanna
Richard Coe
Mary Crossland
Lisa E. Fuchs
Carlos Barahona
Brian Chiputwa
Levi Orero
Beatrice Adoyo
Matthias Geck

WORKING PAPER 4

FEBRUARY 2024

WORKING PAPER 4

Developing holistic assessments of food and agricultural systems

A meta-framework for metrics users

Christine Lamanna,[a] Richard Coe,[a,b] Mary Crossland,[a] Lisa E. Fuchs,[a] Carlos Barahona,[b] Brian Chiputwa,[a] Levi Orero,[a] Beatrice Adoyo[a] and Matthias Geck[a]

a CIFOR-ICRAF, United Nations Avenue, PO Box 30677-00100, Nairobi, Kenya
b Statistics for Sustainable Development, Reading, UK

The Transformative Partnership Platform on Agroecology (Agroecology TPP)

Working Paper 4

DOI: 10.17528/cifor-icraf/009081

Lamanna C, Coe R, Crossland M, Fuchs LE, Barahona C, Chiputwa B, Orero L, Adoyo B, Geck M. 2024. *Developing holistic assessments of food and agricultural systems: A meta-framework for metrics users.* Working Paper 4. Bogor, Indonesia and Nairobi, Kenya: CIFOR-ICRAF: The Transformative Partnership Platform on Agroecology.

The Transformative Partnership Platform on Agroecology

CIFOR Headquarters
Jalan CIFOR
Situ Gede, Sindang Barang
Bogor Barat 16115
Indonesia

ICRAF Headquarters
C/O World Agroforestry Centre
United Nations Avenue, Gigiri
Mailing address: P.O. Box 30677 – 00100, Nairobi,
Kenya

T +62-251-8622-622
E agroecology-tpp@cifor-icraf.org

bit.ly/AgEc_TPP

This work was conducted as part of the EU-funded and IFAD administered “Development of Holistic Metrics for Agriculture and Food Systems” Project (Grant number 2000003774) under the Agroecological TRANSITIONS Programme

We would like to thank all donors and partners of the Agroecology TPP. For a full list of the Agroecology TPP partners and donors, please visit: bit.ly/AgEc_TPP

Contents

List of figures, tables and boxes

Tables

Boxes

Figures

1 The need for holistic assessments of food and agricultural systems

Food and agricultural (agrifood) systems feed the world and form the basis of our economic, political and cultural systems. At the same time, they are also drivers of environmental change at local and global levels. Given how central agrifood systems are to lives and livelihoods and the environment, it is not surprising that we try to measure, monitor and assess how they are performing. However, agrifood systems are complex with many connected parts, things happening at different scales in space and time, and great heterogeneity, making measurement challenging. For much of history, our measurement of agrifood systems has focused on economic aspects – production, prices and trade. The Babylonians kept price data for agricultural commodities from as early as 500 BCE (Temin 2002), and such metrics still make up the bulk of international databases such as FAOStat.

A key aim when data on system performance are collected is that those data will be interpreted and used in future decisions about the system. An implication is that those system properties that are measured will improve over time with management decisions, while properties that are not measured will change positively or negatively depending on what drives them and their correlation with measured properties. Hence for a long time, researchers and activists have argued that we need to measure much more than just production and economic aspects or dimensions of agrifood systems. For example, the Sustainable Livelihoods Framework was proposed in 1991 (Chambers and Conway 1992) and has prompted measurement of human, social and natural dimensions of food and agricultural systems in addition to production. The IAASTD (McIntyre et al. 2009) and UN Sustainable Development Goals (2015) led to a proliferation of schemes for system measurement, particularly in relation to sustainability (Kalachevska et al. 2022).

One particular concern with assessment of agrifood systems is that of the 'level playing field'. As an example, agroecologically intensified systems have been critiqued for being less productive or profitable than industrially intensified systems. While the validity of this criticism has been challenged, it would not be surprising if it were true in the short term. Agroecological systems are *developed with the aim of providing environmental and social benefits,* not only economic benefits, and to be sustainable in the long term. To assess them *only* in short-term production or economic terms misses the whole point of the approach. Likewise, assessing conventionally intensified systems without including their longer-term social and environmental effects leads to degradation of those aspects. The same principle – that only assessing part of the system leads to worsening outcomes in other parts – applies to any agrifood system. Agrifood systems are a major contributor to our collective problem of exceeding planetary boundaries (Conijn et al. 2018; Gerten et al. 2020; Rockström et al. 2020). At the same time, they contribute to dietary and health problems (Willett et al. 2019), rural and urban livelihood challenges, and local environmental degradation. We need to be able to measure alternative systems holistically and using the same terms, and then compare them within a common framework – the level playing field.

2 Metrics, metrics everywhere...

The emerging consensus on the need for holistic assessment of food and agricultural systems has led to a recent proliferation of available frameworks based on differing needs, objectives and conceptual frames. A systematic review of 216 assessments reveals an increasing number of holistic assessments published annually over the past 30 years, with a dramatic rise since 2019 (Figure 1) (Crossland et al. in preparation).

These assessments have been designed to meet different objectives. For example, a major axis of variation in assessments for agroecology is whether it is for system characterization (assessing to what degree the system is agroecological or adheres to agroecological principles), or for measuring system performance (how well the system meets various goals or outputs) (Figure 2, redrawn from

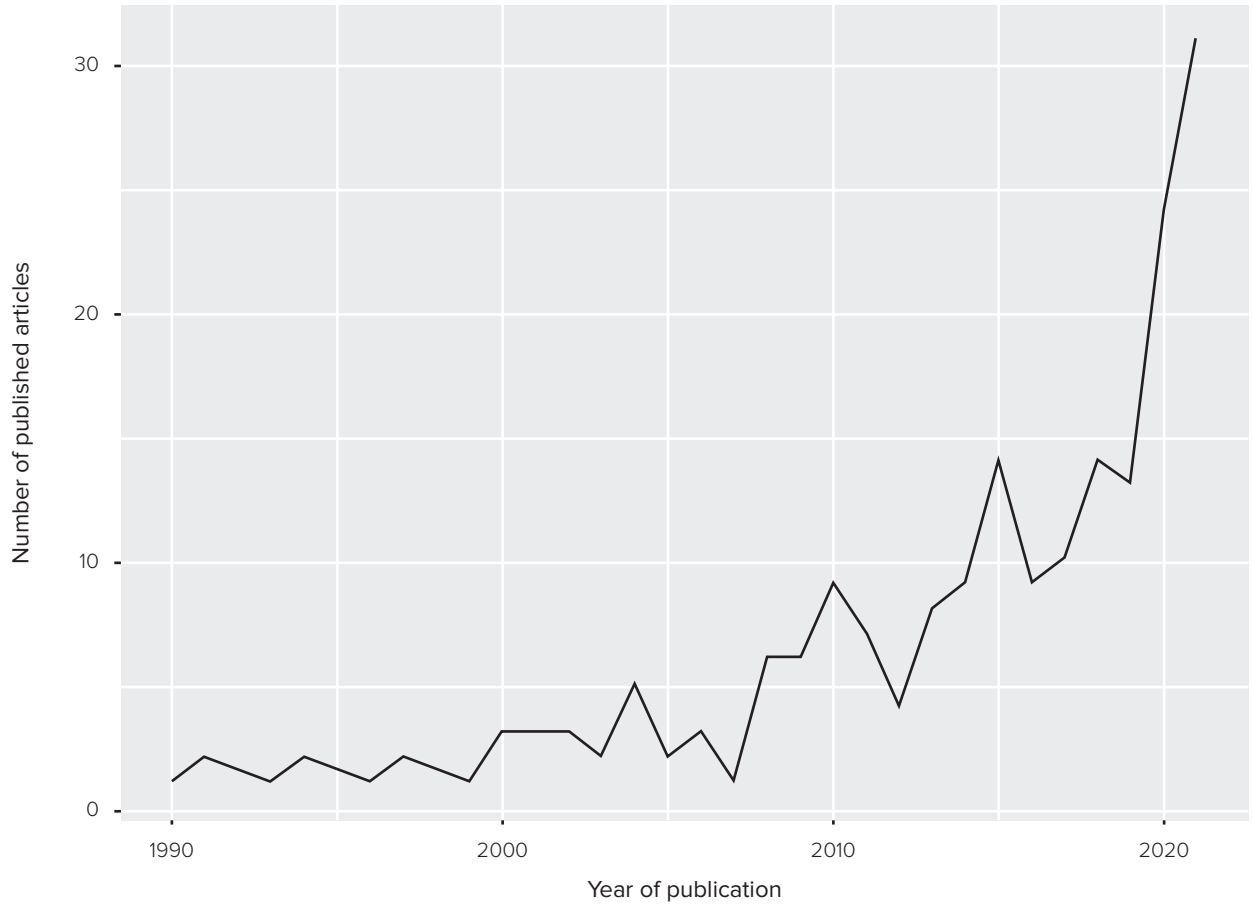

Figure 1. Holistic assessments published over time

Note: Number of holistic assessments of food and agricultural systems published annually since 1990 based on a review conducted by Crossland et al. (in preparation)

Geck et al. 2023). While many approaches are strictly for either characterization or performance assessment, several do both, often with the goal of linking characterization to performance (e.g., do more agroecological farms perform better than fewer agroecological farms?).

Another key axis of variation amongst tools is the scale at which they are intended to operate. While many tools are available for assessment of fields, farms and farming households, relatively few target the landscape or food system scale. Other tools may focus on different components of a value chain, from on-farm production to marketing to consumption. Still others focus on analysis of projects or portfolios of projects rather than on-the-ground analysis. Tools can also differ in the degree to which they utilize quantitative versus qualitative data and analysis approaches, or in the level of participation or co-creation with stakeholders (Geck et al. 2023).

Many frameworks are developed in response to identified gaps existing at the time (Rasmussen et al. 2017; Talukder et al. 2020). Some frameworks are proposed as candidates for globally applicable and standardized assessments, such as FAO's Tool for Agroecology Performance Evaluation (TAPE) (FAO 2019) which has now been applied in many countries. While there are advantages to using a standardized framework, the context specificity in agrifood systems and the ever-changing objectives of projects and policies means that there will always be a need for new frameworks.

For anyone seeking data about agrifood systems and deciding to initiate data generation, there are three broad choices:

1. Take a common framework and apply it: For example, FAO's TAPE is being promoted as universally applicable and has other characteristics, such as ready-made tools, that may make it attractive. Using such a ready tool is clearly efficient if it meets the user's objectives.

2. Review existing frameworks before selecting: Recent reviews of frameworks, such as for understanding socioecological systems (C. Binder et al. 2013) or sustainability (Alaoui et al. 2022) are excellent starting points. Such a systematic look at multiple frameworks gives insights for choosing, merging and developing further holistic assessment frameworks. However, this approach requires the time and expertise available to conduct a systematic review and analysis of the ever-growing number of assessment frameworks available. There is also no certainty that the set of frames considered will include anything suitable.

3. Develop or adapt a new framework: This route has the potential to meet objectives but is certainly not easy. It is the subject of the remainder of this publication.

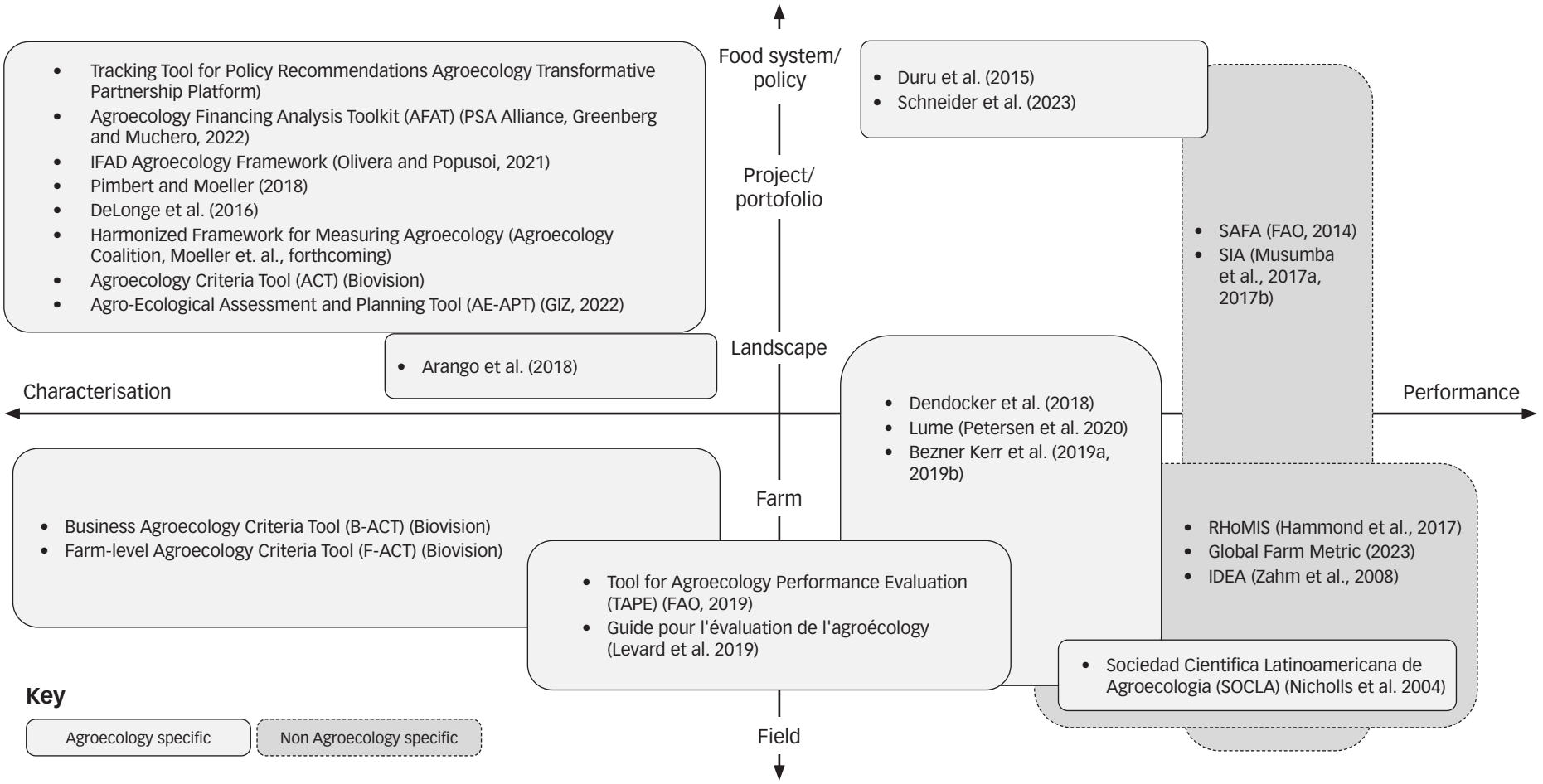

Figure 2. Typology of tools and frameworks for agrifood system assessments

Note: Assessment goal (characterization versus performance) on the x-axis and scale of assessment (field to food system) on the y-axis

Source: Redrawn from Geck et al. 2023

3 A meta-framework as a guide to designing an assessment

Given the diversity of agrifood systems globally, as well as the diversity in objectives of those who wish to collect data on the performance of those systems, there can be no single assessment framework which can meet every objective in all possible contexts. **There is no one-size-fits-all solution when it comes to measuring food and agricultural systems**. Any group that finds a need for holistic system assessment data has the challenge then, of selecting from available frameworks, adapting them to their needs, or innovating when there is nothing that meets those needs. Hence, we need a guide through the maze of frameworks to help choose an approach, metrics and process that meet specific objectives. We therefore propose the development of a meta-framework, a scheme to guide or support those groups planning or promoting holistic agrifood systems measurement.

In this document, we put forward general steps for developing a holistic agrifood system assessment, as well as principles to guide decision making at each step of the design process. Holistic measurement of agrifood systems may be planned as part of research studies; in connection with development projects and interventions; by activists collecting evidence; by groups and communities engaged in action research cycles; or as routine data collection by governments at local, national or international scales. Our guiding framework should be relevant to any of these.

To organize the ideas, we use the concepts of:

1. The **Destination**: The goal of providing a holistic assessment of an agrifood system
2. The **Compass**: Principles for designing an assessment
3. The **Landscape**: The features and elements of the assessment system
4. The **Path**: The steps in developing an assessment.

4 Definitions and concepts

Agrifood system: An agrifood system includes the agricultural production of food and non-food products, as well as all of the steps to its consumption and disposal, including harvesting, storage, processing, value addition, transportation, marketing, etc. Additionally, an agrifood system includes environmental and ecological processes that support food production as well as the human capital, social and governance structures it is embedded in (Eigenraam et al. 2020).

Principles: Here, 'principles' are defined as fundamental, overarching guidelines that can be used to steer and shape an approach, methods and decision-making process. They are meant as guides to practice that can be interpreted in contextually appropriate ways, rather than ridged rules (Patton 2018).

Framework: Here, a 'framework' refers to a coherent set of principles, concepts and criteria that are linked by theory. They provide organized structures to guide the design, implementation and evaluation of agrifood systems.

Dimensions: While often interconnected, a 'dimension' of agrifood systems refers to one of the specific areas of focus or consideration within the concept of agrifood system performance. Generally, there are three key dimensions that are often considered when assessing agrifood systems: environmental; social; and economic, and each of these can be broken down into further dimensions in many ways.

Holistic assessment: an assessment is 'holistic' when it describes the system as a whole, not just one part or aspect of the system. In the following section, we explain the core characteristics of a holistic assessment.

Metrics and indicators: Although a distinction between 'metrics' and 'indicators' can be made, we use these terms interchangeably to mean something measured (quantitative or qualitative) that can be used to describe system performance. They may be directly observable, or composite metrics based on multiple components spanning several dimensions (e.g., the Women's Empowerment in Agriculture Index, or the Multidimensional Poverty Index).

Assessment system: The landscape of an assessment comprises the measurements made, along with the processes and people that design the assessment, generate and use the data.

System performance and characterization: System assessments can focus on performance, meaning what the system produces or achieves, or on characterization, meaning the description of what the system comprises. This is a distinction between ends and means. We are focusing on system performance or ends.

5 The Destination – Holistic assessment of agrifood systems

The use of 'holistic' in relation to assessment of an agrifood system implies the system as a whole is assessed, rather than only some properties or components of it. Various authors have described what that means in practice. For example (Sandhu et al. 2019) say "Our proposal is to consider all social and environmental externalities – both negative and positive, in global agriculture and food systems and reflect them in an economic system by evaluating comprehensive costs and benefits". However, they take a commodity value chain approach, which only meets some of the breadth of possible objectives. More importantly, it is never possible to measure or assess 'everything' as there is no limit to what that means, so a 'holistic assessment' is something to strive for rather than an absolute standard. We can say that one approach is more holistic than another, but not that one is while the other is not. However, there are characteristics that we expect to be included in any assessment that claims to be holistic.

That it:

1. is **multi-dimensional**, and includes assessment of at least some economic, social and environmental aspects;

2. is conducted from **multiple perspectives** – As different actors in the system and users of the data are likely to assess system performance differently, these multiple views will be included in a holistic assessment;

3. generates insights into **synergies and trade-offs** in the system – Synergy and trade-offs are not usually measured directly, but inferred during data analysis and interpretation. However, the intention to interpret data in this way has implications for the way data are collected.

4. includes assessment of **emergent properties**[1] that only appear at, or are defined at, the level of the system as a whole. Examples are resilience, sustainability, circularity and justice. Evidence on these aspects may be assembled from evidence on each of the separate dimensions, but will also need some additional assessments that would not be included if we were only interested in some of those individual dimensions.

1 Emergent properties of a complex system are properties of the system as a whole that are not properties of the constituent parts.

6 The Compass – Principles to guide holistic assessment development

We start by proposing principles for design of holistic measurement systems. Here, 'principles' are meant as guides to practice that can be interpreted in contextually appropriate ways (Patton 2018). The principles are normative statements that can guide the user in making decisions at each step in the assessment design process to achieve the goal of holistic assessment and co-creation of knowledge.

The choice of principles for decision making reflects the overall goals of the assessment design or guidance. For example, in the UN's recently published principles for measuring progress beyond GDP, key principles include "Universally applicable" and "Comparable across time and countries" (UN 2023). This reflects the UN's need for an assessment system that can be applied in all of the member states. Principles have been proposed for the application of FAO's TAPE tool (FAO 2019) for conducting agroecological assessments (Namirembe et al. 2022). These principles stress the co-design and localization of TAPE in order to increase its context-specificity and applicability in diverse contexts. Both sets of principles include an emphasis on iteration, trying, reflecting and improving the assessment through the design process.

An initial set of proposed principles for designing holistic assessments of agrifood systems are stated below (Figure 3), each followed by an explanation and some of the consequences. These principles are normative statements, meant to guide a user at each stage of the assessment design process.

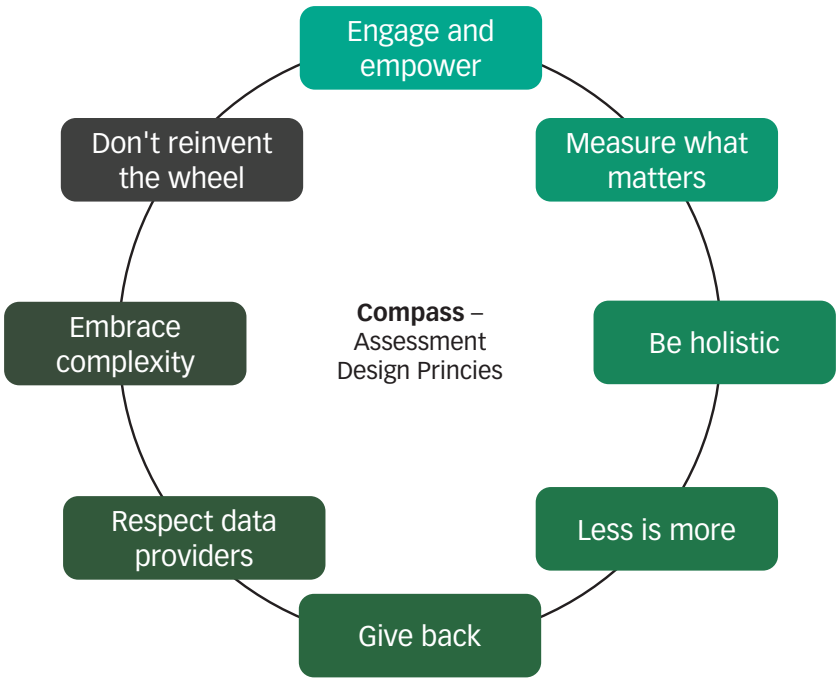

Figure 3. Eight principles for designing holistic assessments

Though we write down discrete and numbered principles, they overlap and interact. The list of principles is not exhaustive, but draws from a network of principles related to agroecology, ethics of collecting and using data, professionalism and so on. This set of principles are those we feel it is most important to consider when designing holistic assessments for food and agricultural systems.

6.1 Engage and empower

The process of collecting data for an assessment is typically extractive, particularly for those on the front lines of data collection, such as farmers, respondents or interviewees. However, those collecting the data do so with the intention that the data and information it provides will benefit those who are participating in its collection by improving farm management or supporting policy development. Changing data collection from an extractive process done by outsiders to knowledge co-production between beneficiaries and researchers or development practitioners can increase the rate of innovation and uptake. This approach is increasingly valued by many actors involved in development processes and has been elaborated in connection with food systems assessment (Alrøe et al. 2016).

6.2 Measure what matters

This principle was outlined in the introduction. It says that the choice of what to measure must be determined by what 'you care about' or 'what matters' (Stiglitz et al. 2018; Stiglitz 2020). While the aim is to be holistic, there is an unlimited number of aspects or properties of any system that could be assessed and measured, and choices must be made. The first requirement is to measure those dimensions on which you want systems to improve – that is, dimensions that you care about.

Exactly who is meant by 'you' is addressed in the first Principle: Engage and Empower. It likely includes those with a vested interest in data collection (the researchers, development practitioners, donors, project managers, etc.) as well as the intended beneficiaries or participants (farmers, value chain actors, etc.) and interested stakeholders (project partners, civil society, government, etc.).

Identifying what matters most – what types of system changes are most desired – can increase the buy-in of stakeholders to the project and data collection activities. Empowering stakeholders to contribute to the assessment design process does not mean relinquishing control over what data is collected. It instead means a dialogue between all partners on what matters most to measure. The likely outcome is that the things that matter most to measure include both the external targets (e.g., key results or impact areas of a donor, policy targets of a government) as well as the aspirations and motivations of the beneficiaries and participants.

6.3 Be holistic

In a previous section we have given our definition of what it means for an assessment to be holistic. Assessment of social, economic and environmental dimensions have to be included. Despite common use, the three categories are somewhat arbitrary. There are dimensions that fall into more than one category (for example soil health is an ecosystem service in 'environmental' and a production input in 'economic') and others that are unclear (human health, animal welfare). The value in using these categories, and the many subcategories that could be added, is to prompt thinking beyond the few dimensions that are typically assessed.

Engaging the people who an assessment is meant to serve and identifying the things that matter to them often leads to a more holistic assessment system. However, it can also present challenges for designing the assessment. For example, important parts of the social dimension are often the

aspirations, motivations, values and attitudes of people who are part of the system (Alrøe et al. 2016; Soini Coe and Coe 2023). Assessing these dimensions often requires subjective and qualitative approaches, along with the objective and quantitative data.

6.4 Less is more

The desire to be holistic and participatory in designing an assessment system can lead to an overwhelming number of indicators that would be great to collect. In reality, anyone designing an assessment system will have limited resources for data collection, from time to equipment to funds for enumerators, etc. Collecting every possible indicator will be prohibitively lengthy and complex data collection can result in low quality data. Choose indicators that enable holistic assessment, but are not redundant or burdensome. Try not to collect more information than you need to adequately assess your system.

6.5 Give back

Generate data and information that can be used by those it is meant to serve. This principle is a consequence of *engage and empower and measure what matters* principles, which focus on who the assessment is for and what their objectives are. It is often said that data needs to be 'actionable', meaning it informs decisions, and those decisions and actions will be different from what would happen without the data. Being actionable or useful is not purely a function of data, but also of the capacity of users. This includes the capacity to access, process and interpret data. As an example, if the data is supposed to meet farmers' needs but can only be interpreted through sophisticated analytical models or in combination with other data that they do not have access to, then it is not appropriate.

6.6 Respect data providers

There are many principles and guidelines in usage on the ethics of data collection, storage and usage, particularly when that data contains personal information about people, their households and their livelihoods. The importance of high ethical standards in data collection warrants repeating here. Respecting data providers means following basic ethical guidelines on survey design, informed consent, data storage, data usage and privacy. The specific ethical guidelines you must follow as a data collector will vary based on your location and the type of institution you represent, but will involve a minimum of an ethical review board clearance at institutional or national level (or a combination of the two). Researchers and data providers based in Europe will need to follow GDPR regulations on the protection of personal data.

Beyond basic data ethics, the *respect data providers* principle asks us to consider the relationship between the data collectors and data providers. Engaging and empowering those data providers in the design of the assessment increases respect, as does giving back information to those who have provided the data. Reducing the burden of survey time, complexity or number of visits also respects data providers by acknowledging the time it takes to provide data. Engaging in a co-creation or participatory research process can be time-consuming for stakeholders, so understanding how and what stages they would like to be engaged will contribute to the success of your data collection efforts.

6.7 Embrace complexity

When designing an assessment or collecting data, we often want that data to give us a simple answer to questions such as 'how sustainable is this practice?', 'how agroecological is this farm?' etc. Although reducing information to a simple score can be appealing, it often hides the nuances behind the numbers. Particularly for holistic assessments, being able to understand the potential trade-offs in performance across different dimensions (e.g., environmental versus economic performance) is critical to managing the whole system. There is also context specificity. A practice or intervention may perform differently in different farming systems or for different types of users (households with different assets and capacities, men and women, etc.). Embrace the complexity inherent in a holistic system assessment by developing ways of analysing trade-offs and synergies or context specificities while still being accessible and actionable by the target users (the *give back* principle).

6.8 Don't reinvent the wheel

As we have shown, there are a plethora of available assessment frameworks, tools and metrics for conducting holistic assessments for food and agricultural systems. Consider whether existing tools are fit for your purpose before designing a new one. You might just save a lot of time as well as contribute to larger comparable datasets (Rosenstock et al. 2017).

7 The Landscape – Elements of an assessment system

The preceding sections imply that the design and successful implementation of a scheme for holistic measurement of a food and agricultural system is very much more than the selection of metrics and indicators. There are many interacting components to understand and ensure they are in place and functioning – the landscape or ecosystem of the assessment (Figure 4). As with any ecosystem, this one can operate at different scales, from a landscape (such as done in some project monitoring) to a global system (such as supported by the UN or World Bank). There are many ways in which it can be described, with functions and processes lumped or split. In Figure 4 we provide a summary of the assessment ecosystem and emphasize three main components: framing, metrics and data, and their various sub-components, each of which needs navigating/consideration when developing and carrying out a holistic assessment.

The **framing** brings together the users or stakeholders, their objectives, the definition and boundaries of the system being measured, and the theory that supports it all. Interactions between these are likely to be ongoing and iterative. For example, the objectives of one group of users or stakeholders are likely to be updated during discussion with others.

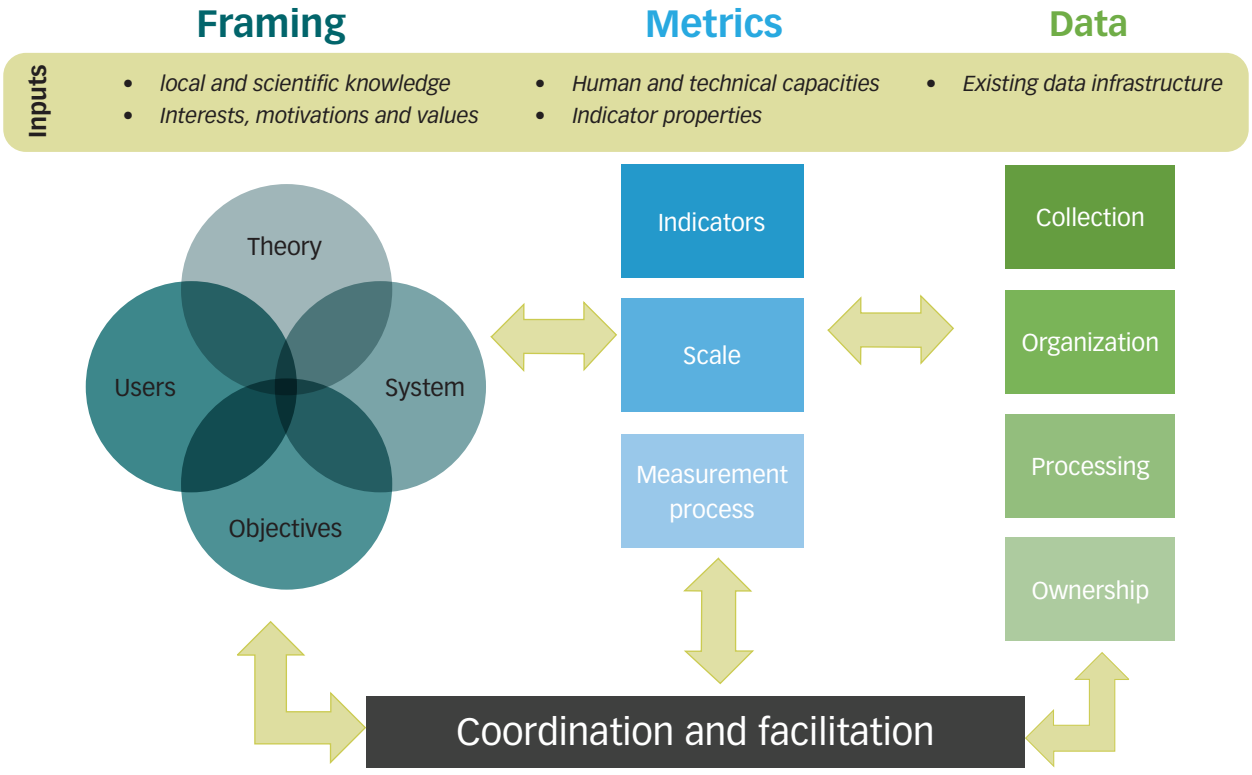

Figure 4. The landscape of a holistic performance assessment

The **metrics** or indicators are the part of the ecosystem to which researchers pay most attention. When designing a practical measurement scheme, the measurement process (cost, skills needed, etc.) is likely to be as important as the conceptual basis for choosing an indicator.

The **data** section includes tools and processes for generating, managing and using the data from an assessment. No objectives will be met unless this is functioning.

Finally, this landscape is unlikely to be self-organizing. Someone will need to start it up, **coordinate and facilitate** it – perhaps provide the energy that drives it. Whoever that is will probably also be among the users or stakeholders (framing), but the coordination and facilitation function is separate, so it is represented as a piece in the ecosystem diagram.

When imagining and designing the landscape of an assessment, we need to be aware of the power relations involved. State actors (governments) supported by international bodies such as the UN and World Bank are in a very different position from, for example, an NGO, activist group or community. States can *require* data to be collected and provided, as happens in agricultural censuses, and provide the resources needed for all parts of the assessment ecosystem. An NGO or activist group can raise resources to conduct data collection and use but cannot require anyone to provide it. A community or farmers' group can often do neither, so will depend on simple data collection and handling methods and voluntary compliance.

8 Charting your Path – Steps for developing a holistic assessment

Here, we briefly outline a stepwise process for designing a holistic assessment system for food and agricultural systems (Figure 5). Although we present the steps in a linear order, designing an assessment system should be an iterative process, particularly if it is done in a participatory way. The first few steps could in practice be done in any order, as each decision informs the scope of the steps after it. Each potential assessment designer may have a different mandatory starting point based on their reasons for measuring. They may have a set geographic boundary to work in (system boundary), they may have a specific stakeholder platform to work with (who will be involved), or they may have an a priori goal for the assessment.

There are many ways that the steps can be completed, depending on the needs and goals of the data collectors. The principles we previously discussed in the "Compass" section inform choices made at each step and are meant to guide users to designing a more holistic and inclusive assessment system. Not all principles inform each step; rather certain principles are more important at various stages in the assessment design process (Table 1).

Below we discuss what each step entails, and some guidance on how to complete it. In the future, this section will be further elaborated for specific use cases of holistic assessment design. This first iteration is meant as a general description of each step.

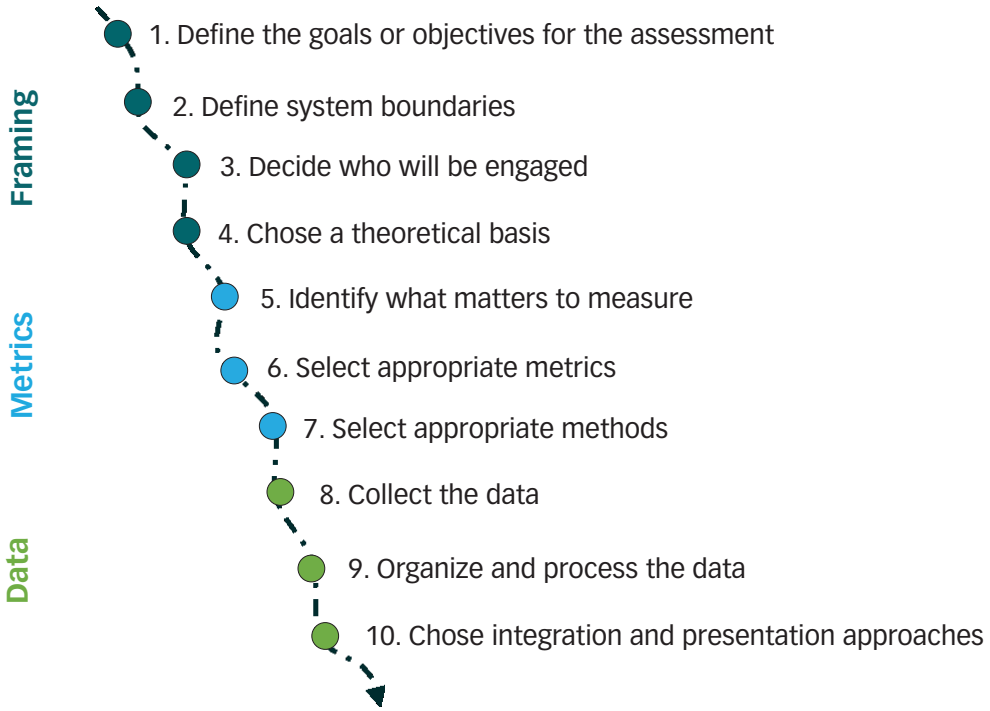

Figure 5. The ten steps for designing a holistic assessment system

Table 1. Steps in designing a holistic assessment, and the principles invoked

	Step	Principles
Framing	1. Define system boundaries	Be holistic
	2. Decide who to engage	Empower and engage, respect data providers
	3. Set goals and objectives	Empower and engage, be holistic
	4. Chose theoretical basis	Be holistic
Metrics	5. Identify what matters to measure	Measure what matters, empower and engage, be holistic
	6. Select appropriate metrics	Don't reinvent the wheel, measure what matters, less is more
	7. Select appropriate methods	Don't reinvent the wheel, measure what matters, respect data providers
Data	8. Collect data	Empower and engage, less is more, respect data providers
	9. Organize and process data	Respect data providers
	10. Integrate and present	Generate information, embrace complexity, empower and engage

8.1 Framing (steps 1–4)

Step 1: Set the goals and objectives for the assessment

What is the purpose of the assessment, and what are you trying to achieve by assessing your system and collecting data? You may already have a clear idea of what the goal of the assessment is, determined a priori in the project design or policy objectives. Alternatively, you may have some ambiguity in the goals, or multiple perspectives (different stakeholders or interests) to consider. Having a clearly defined and agreed upon goal for the assessment will greatly help in designing the assessment and making choices about what to include and what not to.

Some key questions to ask when defining the purpose of your assessment, and considering who will be involved and impacted, include:

- Purpose: Why do we need this assessment? What is the issue we are attempting to address? What are we trying to find out, test or prove? What are our hypotheses?
- Stakeholders: Does everyone have the same objective or goal for the assessment? Do we need to consider multiple perspectives?

Every aspect of the assessment depends on exactly what it is for – the reason it is being done and its specific objectives. "We need data on x" is not a sufficient statement of objective, or sufficient to design the rest of the assessment. Referring to Principle 1 (*engage and empower*), setting of objectives should be done in consultation with those who have an interest in the output. For example, the objectives will determine whether the assessment needs to track changes over time, compare alternatives at a single time point, or do something else. These have profound implications for the design of the assessment.

If different systems are to be compared on equal terms, then the focus of the assessment should be on system performance, meaning the extent to which the system meets defined goals or purposes. For example, a goal might include producing nutritious food without degrading the environment. Indicators of performance would include indicators of food quality and quantity, and of environmental impacts. It is possible that one system is more diverse than another, and diversity helps meet these goals. Then diversity per se is a means (a characteristic of the system) rather than an end or goal. Hence diversity would not be an indicator to include in the assessment. However, only a slight change in objectives and conception of the purpose of a system might shift aspects of diversity into

being considered goals, and require indicators of diversity to assess performance. This example illustrates one reason why a careful statement of objectives is needed.

Step 2: Define the boundaries of the system to be assessed

Holistic assessment of agrifood systems could include everything from farm management practices to value chains; from consumer behaviour to the policy environment, and more. Assessing everything is unfortunately not possible, so it is key to define the boundaries of what you will assess. Limits or boundaries should include the parts of the agrifood system considered, the geographical extent and the time period.

Determining the boundaries of the assessment requires careful consideration. For example, if the focus of the assessment is on agricultural production, it could mean a specific commodity (e.g., maize) or it could mean all the production from the farms that produce the maize (e.g., other crops, livestock, tree products). If focus is only on one commodity such as maize, exactly which/whose maize production will be included? What is the geographical scope of the maize production under scrutiny (a farm? a district? a 'foodshed'?). If the focus is instead on household consumption, which of the dozens of value chains implicated in anyone's diet do we look at? The conception of the system must be bounded before it can be assessed (Ostrom 2009).

The multi-level perspective can be useful when visualizing the extent of the system being studied (Geels 2019). It views the system, and system change, as having three levels:
- Niche (place where things happen, such as a farm or physical landscape. Rapid change and innovation are possible)
- Sociotechnical System (rules and institutions that control changes in Niche – can change fast)
- Sociotechnical Landscape (long-term trends and drivers – usually slow and hard to change).

Decisions on the scope and boundaries of the system to assess must be based on objectives, and if objectives are stated first, it is likely that they will need to be updated once the scope of the system has been mapped out.

Step 3: Decide who will be engaged

While you may already have a clear goal and know who will be driving the assessment, it is important to reflect on what you are trying to achieve, who is it for, and who are key players – for example, there may be actors you have overlooked and who will be important in helping you reach your goal. TEEB considers three key groups of stakeholders: 1) the assessment audience; 2) partners contributing to the doing the assessment; and 3) people who may impact or be impacted by the findings (Eigenraam et al. 2020). These may not include the *subjects* of the assessment (farmers, consumers, people along the value chain between them).

Given your goals for the assessment, and boundaries of the system to be assessed, a key question is: **Who are the stakeholders, or people whose interests are served by the assessment?**

Co-designing the assessment with the key actors you have identified helps ensure assessment is from multiple perspectives, and increases the relevance and utility of the assessment and its outcomes. This can be achieved through structured multistakeholder engagements and co-design processes.

Step 4: Choosing a theoretical basis

The process of designing an assessment is not only one of selecting indicators. If there is no coherent basis for those selected, we may end up with a collection of measurements that cannot be usefully integrated and hence does not meet the requirements of holistic assessment. The

Feedback and Iteration

We have presented the first three steps (set goals and objectives; define system boundaries; decide who will be engaged) in a linear order, but these three steps are interdependent. Choices at one step will interact with choices in the others. For example, a goal of the assessment might be that it is participatory, or that it reflects local stakeholder priorities. In such cases, identifying the stakeholders to engage (the ones with a stake in the system you are interested in) and convening a process to set goals and objectives in a participatory way might be in order. Thus, there is often some amount of iteration that happens when designing an assessment system – especially in the framing phase.

underlying conceptual basis of an assessment will determine much of the rest of the design of a set of indicators and the way they are used. To ensure that we select metrics that meet our objective and can be usefully integrated, it is important to clearly identify the theory on which the assessment will be based.

Often, elements of the conceptual or theoretical basis will be explicit in the reasons for designing an assessment. Many projects are based on testing or upscaling an approach, which carries conceptual elements that should be included in an assessment (e.g., climate-smart agriculture, agroecology, regenerative agriculture, ecosystem-based adaptation, etc.). Other times, the goals of the engagement will point to conceptual bases to include (e.g., food security, sustainability, resilience, adoption of innovations, etc.). More broadly for holistic system assessment, Binder (2013) provides the basics for selection of a theoretical basis of social-ecological system assessment. Other commonly used frameworks include the Sustainable Development Goals, Doughnut Economics (Raworth 2018), versions of footprint analysis (Halpern et al. 2022) and energetics (Wacha et al. 2022). Any of these could be modified to guide choice of theoretical basis of our assessments.

An explicit theoretical basis will help to structure the subsequent steps where we identify and choose key metrics to collect data on – ensuring we gather all the data necessary to achieve the assessment goals without being overly burdensome for the data collectors and providers.

8.2 Metrics (steps 5–7)

Step 5: Identifying what matters to measure

Once you have defined the assessment goals, the system boundary, and engaged the key actors in that system, the next step is to identify what is most important to measure or capture data on in the assessment.

Here, tools from systems thinking and analysis can be helpful. This can include the use of causal and systems diagramming tools and approaches, which allow you to map out the components of a system, their interactions and the potential outcomes and impacts of system changes (e.g., new farming practices, approaches, policies). This can help identify what aspects or dimensions will be of interest and need to be measured in the assessment.

Again, this process can be done as a participatory process, and will require negotiation, trade-off, and a balance between being holistic and what is feasible. Overambitious assessment designs end up not being used (de Olde et al. 2017). There may also be a need to nudge people into considering dimensions they had not yet considered as important to measure/in their interest to measure (e.g., dimensions that are important to those impacted, but not driving the assessment). Useful approaches and sources of information on such approaches include adaptive management, community-based management, multiple knowledge systems, farmer indicators, etc.)

The Metrics Library

To assist in identifying appropriate metrics for the prioritized concepts or aspects you have decided to measure, we are developing an interactive **Metrics Library**. It is the result of a systematic review of frameworks and metrics for holistic assessments of food and agricultural systems (Crossland et al. in preparation). This interactive database will offer a one-stop shop where users can view, explore and select the right metrics for their objectives, context and practical constraints.

Step 6 and 7: Selecting appropriate metrics and methods for collecting them

Once you have identified what should be measured, the next step is choosing appropriate metrics that will give information on the desired factors. Often in participatory processes for indicator selection, relatively broad concepts can emerge. For example, a key priority for measurement might be soil health, but there are dozens, if not hundreds of metrics available for assessing soil health. These metrics will differ in what is measured (soil carbon, soil moisture, erosion, cation exchange capacity, colour and texture), and how it is measured (collecting a sample, qualitative assessments, consulting geospatial data, etc.). The same is also true for more abstract or qualitative concepts, such as agency, equity, attitudes, etc. Navigating these choices is the subject of steps 6 and 7 in the assessment design process.

The metrics used in the assessment will have to be based on the purpose and theory, and the principles described above. In addition, there will be consideration of:

- Scope in terms of geography (geo boundaries) and time (one-off, repeated, etc.)
- Scale or granularity (space and time) and the observational unit
- The measurement or data collection process and the skills and resources needed (time, money, equipment, laboratory analyses, enumerators, etc.)
- The precision and sample sizes required
- The inclusion of local knowledge
- The inclusion of qualitative measurements.

Table 2. Comparisons for interpretation of indicators

Possible comparison	Implications for measurement
1. Same location at a different time: This is the comparison used when tracking the evolution of a system over time, and applies at any scale from a single farm to global.	This necessitates maintaining the same definitions, methodology and choices of metrics over time. This may be a challenge given Principle 5.
2. Different cases in the same context: For example, this is the comparison used to compare two farm types in the same location, or farms that have and have not taken part in an intervention.	This requires agreement on 'what matters' and relevant indicators in all cases being compared. We also need to understand which contextual factors matter so that we are sure they are constant across the cases compared. For example, comparisons of an agroecologically intensified farm in a humid area with an industrially intensified farm in a dry area are generally meaningless.
3. Different cases in different contexts (e.g., farms in different locations)	This requires agreement on what matters in both contexts and in what sense things are comparable across contexts. Sensible measurements are possible in both. Numbers can be interpreted, perhaps after standardization
4. Actual levels compared with theoretical, expected or desired levels	This requires clear and uncontested agreement on the basis of the comparison. For example, crop yields are sometimes compared with 'potential yield', but that requires both agreement on what 'potential yield' means and agreement on the estimates of that value.

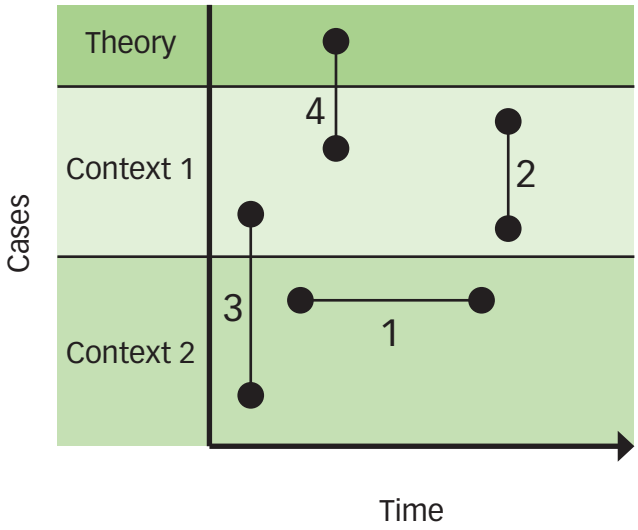

Figure 6. Possible comparisons of indicators

Note: Numbers refer to Table 2

Indicators only make sense when compared to another instance of the same indicator. This sets limits to adaption, to context, and the extent to which an assessment can be standardized and widely adopted. Possible comparisons are summarized in Table 2 and illustrated in Figure 6. Table 1 also summarizes implications of the type of comparison for the indicators and way data are collected. Combinations of these comparisons in the form of 'differences of differences' are commonly used. For example, comparison of change over time (type 1) between two different contexts (type 3).

8.3 Data (steps 8–10)

Step 8: Collecting the data

Once the team has decided what metrics to collect – it is time to actually collect the data! This process involves deciding: Who will be carrying out the assessment? When? With what resources? And how will data be captured? To ensure a robust and ethical approach to data collection, important considerations will include:

- Ensuring those collecting data are adequately trained and equipped for the task;
- Piloting your methodology and refining your approach before wider role out;
- Ensuring free, prior and informed consent of those contributing data;
- Periodic review of data collection and quality to address any emerging challenges.

Step 9: Organizing and processing data

Once data collection is complete, the next critical step is organizing and processing the data. This involves several key considerations from how data is stored and secured, and how is it checked and cleaned, to who will have access to it, and under what terms. Key issues here relate to data ethics and where following FAIR (Findable, Accessible, Interoperable and Reusable) principles can enhance the usability and longevity of the collected data.

Step 10: Choose integration and presentation approaches

Now that the data have been analysed, the next and final step is to use that data to generate information that will address the goals of the assessment and the interests of the various stakeholders who were engaged in its design.

Research has shown that individuals have strong preferences for different types of data visualizations, and that no one visualization is likely to suit all stakeholders (Bourne et al. 2023). In this step we will describe choices of different presentations of the data, and their strengths and weaknesses. This will include approaches to methods that reveal rather than hide complexity, heterogeneity and the interdependencies in data sets.

Key to this step is also “giving back” to the data providers, i.e., sharing back results and information to respondents or farmers in a way that responds to their interests in participating in the project. There are many methods for sharing research results with beneficiaries, from producing individualized reports, to apps, to radio programmes or SMS messages. Sharing information back with the data providers can show the value of the project to the beneficiary as well as increasing uptake of innovations or positive behaviour changes that were the goal of the engagement.

9 Use cases

Once the previously described principles and steps for designing a holistic assessment have been elaborated and validated by review, who might be able to use them? Our aim for the meta-framework is to describe a generic process that is useful to many different audiences. The results of following this process for each of these different audiences will be different and result in assessments adapted to their needs. Possible examples of different use cases are outlined in Table 3.

Table 3. Potential uses of the holistic assessment meta-framework

Coordinated by	Aim	Others involved	Likely characteristics
Project managers, M&E specialists	Monitor impacts of project activities on key performance areas	Implementing partners, donors, beneficiaries	Tailored to project goals, Short term (project lifespan), Monitoring and evaluation Impact assessment
International alliances or organizations	Track changes in food and agricultural systems at global scale Compare progress across countries or regions	Researchers	Common approaches adapted to share learning and expense of development
National governments	Track changes in food and agricultural systems within their countries for national statistics, to track implementation of policies or progress towards policy objectives	Local governments	Nationally representative Long term Idiosyncratic? Difficult to change?
Development funders	Track changes in food and agricultural systems within project areas they are supporting Inform evaluations	Project implementers	Common approaches with additional location-specific foci relevant to individual objectives
Research organizations	Test hypotheses about alternative system trajectories and the processes involved	Local governments and other research partners within selected locations	Use cross-sectional not longitudinal data because of short-term research funding Linked to other research objectives Use some novel and complex indicators
Local activist organizations or CBOs	Provide evidence to support their demands for action	NGOs, community groups	Use simple and cheap indicators

10 Conclusions

The goals of developing a holistic assessment framework and the proliferation of available tools and metrics often leave the would-be measurer feeling like they need to measure everything, everywhere, all at once but with limited time and resources. The principles and framing provided here are intended to support metrics-users in navigating the jungle of available approaches and selecting something that meets their needs. By developing a guiding framework for the design of holistic assessment and metrics, we aim to support their wider use, levelling the playing field for sustainable agrifood systems.

References

Alaoui A, Barão L, Ferreira CSS, Hessel R. 2022. An overview of sustainability assessment frameworks in agriculture. *Land* 11(4): 1–26. https://doi.org/10.3390/land11040537

Alrøe HF, Moller H, Læssøe J, Noe E. 2016. Opportunities and challenges for multicriteria assessment of food system sustainability. *Ecology and Society* 21(1): https://doi.org/10.5751/ES-08394-210138

Binder CR, Hinkel J, Bots P, Claudia P-W. 2013. Comparison of frameworks for analyzing social-ecological systems. *Ecology and Society* 18(4): 26. https://doi.org/10.5751/ES-05551-180426

Bizikova L, Larkin P, Mitchell S, Waldick R. 2019. An indicator set to track resilience to climate change in agriculture: A policy-maker's perspective. *Land Use Policy* 82(December 2018): 444–456. https://doi.org/10.1016/j.landusepol.2018.11.057

Bourne M, Magaju C, Neely C, Peterson N, Lamanna C, Chesterman S, Vagen T-G. 2023. Enhancing co-production of knowledge: Visualisation and engagement approaches for evidence-based decision making within the Kenya Agroforestry Strategy. *Environmental Science & Policy* 140: 152–162.

Chambers R and Conway GR. 1992. Sustainable rural livelihoods: Practical concepts for the 21st century. *In* IDS Discussion Paper (Vol. 296).

Conijn JG, Bindraban PS, Schröder JJ, Jongschaap REE. 2018. Can our global food system meet food demand within planetary boundaries? *Agriculture, Ecosystems and Environment* 251(December 2016): 244–256. https://doi.org/10.1016/j.agee.2017.06.001

de Olde EM, Moller H, Marchand F, McDowell RW, MacLeod CJ, Sautier M, Halloy S, Barber A, Benge J, Bockstaller C, Bokkers EAM, de Boer IJM, et al. 2017. When experts disagree: The need to rethink indicator selection for assessing sustainability of agriculture. *Environment, Development and Sustainability* 19(4): 1327–1342. https://doi.org/10.1007/s10668-016-9803-x

Eigenraam M, Jekums A, Mcleod R, Obst C, Sharma K. 2020. *Applying the Teebagrifood Evaluation Framework* (Issue September). Global Alliance for the Future of Food.

FAO (Food and Agriculture Organization). 2019. *TAPE Tool for Agroecology Performance Evaluation. Process of Development and Tape. Test version.* FAO.

Geck MS, Crossland M, Lamanna C. 2023. Measuring agroecology and its performance: An overview and critical discussion of existing tools and approaches. *Outlook on Agriculture*. https://doi.org/10.1177/00307270231196309

Geels FW. 2019. Socio-technical transitions to sustainability: A review of criticisms and elaborations of the Multi-Level Perspective. *Current Opinion in Environmental Sustainability* 39: 187–201. https://doi.org/10.1016/j.cosust.2019.06.009

Gerten D, Heck V, Jägermeyr J, Bodirsky BL, Fetzer I, Jalava M, Kummu M, Lucht W, Rockström J, Schaphoff S, Schellnhuber HJ. 2020. Feeding ten billion people is possible within four terrestrial planetary boundaries. *Nature Sustainability* 3(3): 200–208. https://doi.org/10.1038/s41893-019-0465-1

Halpern BS, Frazier M, Verstaen J, Rayner PE, Clawson G, Blanchard JL, Cottrell RS, Froehlich HE, Gephart JA, Jacobsen NS, et al. 2022. The environmental footprint of global food production. *Nature Sustainability* 5(December): https://doi.org/10.1038/s41893-022-00965-x

Kalachevska L, Koblianska I, Holzner J. 2022. Concept and measurement of the food system sustainability: A bibliometric research. *Scientific Horizons* 25(1): 104–119. https://doi.org/10.48077/SCIHOR.25(1).2022.104-119

McIntyre BD, Herren H, Wakhungu J, Watson RT. 2009. *Agriculture at a crossroads - Global report*. International Assessment of Agricultural Knowledge, Science and Technology for Development. https://wedocs.unep.org/20.500.11822/8590

Namirembe S, Mhango W, Njoroge R, Tchuwa F, Wellard K, Coe R. 2022. Grounding a global tool-Principles and practice for agroecological assessments inspired by TAPE. *Elementa Science of the Anthropocene* 10: 1–13.

Obst C and Sharma K. 2018. *The TEEBAgriFood Framework: Towards comprehensive evaluation of eco-agri-food systems 202.* pp. 202–245. UN Environment.

Ostrom E. 2009. A general framework for analysing sustainability of socio-ecological systems. *Science* 325: 419–421.

Patton MQ. 2018. *Principle-focused evaluation: The guide*. New York: Guilford Press.

Rasmussen LV, Bierbaum R, Oldekop JA, Agrawal A. 2017. Bridging the practitioner-researcher divide: Indicators to track environmental, economic, and sociocultural sustainability of agricultural commodity production. *Global Environmental Change* 42: 33–46. https://doi.org/10.1016/j.gloenvcha.2016.12.001

Raworth K. 2018. *Doughnut economics*. New York: Random House.

Rockström J, Edenhofer O, Gaertner J, DeClerck F. 2020. Planet-proofing the global food system. *Nature Food* 1(1): 3–5. https://doi.org/10.1038/s43016-019-0010-4

Rosenstock TS, Lamanna C, Chesterman S, et al. 2017. When less is more: Innovations for tracking progress toward global targets. *Current Opinion in Environmental Sustainability* 26: 54–61.

Sandhu H, Müller A, Sukhdev P, Merrigan K, Tenkouano A, Kumar P, Hussain S, Zhang W, Pengue W, Gemmill-Herren B, et al. 2019. The future of agriculture and food: Evaluating the holistic costs and benefits. *Anthropocene Review* 6(3): 270–278. https://doi.org/10.1177/2053019619872808

Slätmo E, Fischer K, Röös E. 2017. The framing of sustainability in sustainability assessment frameworks for agriculture. *Sociologia Ruralis* 57(3): 378–395. https://doi.org/10.1111/soru.12156

Soini Coe E, Coe R. 2023. Agroecological transitions in the mind. *Elementa* 11: 1–13. https://doi.org/https://doi.org/10.1525/elementa.2022.00026

Stiglitz JE. 2020. Measuring what matters? *Scientific American* (August): 25–31. https://doi.org/10.1080/15366367.2012.720208

Stiglitz JE, Fitoussi J, Durand M. 2018. *For good measure: Advancing research on well-being metrics beyond GDP*. OECD. https://www.oecd-ilibrary.org/docserver/9789264307278-en.pdf?expires=1562670849&id=id&accname=ocid84004878&checksum=1B7C676EA3EE7A5A6109CF1ABB5C9425

Talukder B, Blay-Palmer A, van Loon GW, Hipel KW. 2020. Towards complexity of agricultural sustainability assessment: Main issues and concerns. *Environmental and Sustainability Indicators* 6(March): 100038. https://doi.org/10.1016/j.indic.2020.100038

Temin P. 2002. Price behavior in ancient Babylon. *Explorations in Economic History* 60: 46–60. https://doi.org/10.1006/exeh.2001.0774

UN (United Nations). 2023. *Valuing what counts: Framework to progress beyond gross domestic product* (No. 4: Our Common Agenda Policy Briefs. May issue).

Wacha K, Philo A, Hatfield JL. 2022. Soil energetics: A unifying framework to quantify soil functionality. *Agrosystems, Geosciences and Environment* 5(4): 1–15. https://doi.org/10.1002/agg2.20314

Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A, et al. 2019. Food in the Anthropocene: The EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet* 393(10170): 447–492. https://doi.org/10.1016/S0140-6736(18)31788-4

DOI: 10.17528/cifor-icraf/009081

The Agroecology TPP Working Papers contain preliminary or advanced research results on agroecology issues that need to be published in a timely manner to inform and promote discussion. This content has been internally reviewed but has not undergone external peer review.

Food and agricultural (agrifood) systems feed the world and form the basis of our economic, political and cultural systems. At the same time, they are also drivers of environmental change at local and global levels. Given how central agrifood systems are to lives and livelihoods and the environment, it is not surprising that we try to measure, monitor and assess how they are performing.

Considering the diversity of agrifood systems globally, as well as the diversity in objectives of those who wish to collect data on the performance of those systems, there can be no single assessment framework which can meet every objective in all possible contexts. Any group that finds a need for holistic system assessment data has the challenge then, of selecting from available frameworks, adapting them to their needs, or innovating when there is nothing that meets those needs.

Hence, we need a guide through the maze of frameworks to help choose an approach, metrics and process that meet specific objectives. We therefore propose the development of a meta-framework, a scheme to guide or support those groups planning or promoting holistic agrifood systems measurement.

In this document, we put forward general steps for developing a holistic agrifood system assessment, as well as principles to guide decision making at each step of the design process, intended to support metrics-users in navigating the jungle of available approaches and selecting something that meets their needs. By developing a guiding framework for the design of holistic assessment and metrics, we aim to support their wider use, levelling the playing field for sustainable agrifood systems.

Funded by the European Union

IFAD Investing in rural people

CIFOR · World Agroforestry · Alliance Bioversity & CIAT · CGIAR · IRRI · IWMI International Water Management Institute · The University of Vermont Gund Institute for Environment

About the Agroecology TPP

The Agroecology TPP convenes a broad group of scientists, practitioners and policymakers working together to accelerate agroecological transitions. Since its official launch on 3 June 2021, the TPP has begun addressing knowledge gaps across eight domains that will support various institutions and advocacy groups in key decision-making processes. Its online 'Community of Practice' on GLFx is open to all, providing a space for members to share their insights, knowledge and experience.

This partnership was founded by CIRAD, The Alliance of Bioversity International and CIAT, BioVision, UNEP, FAO and CIFOR-ICRAF.