



Insight, part of a Special Feature on [Multicriteria Assessment of Food System Sustainability](#)

Key characteristics for tool choice in indicator-based sustainability assessment at farm level

Fleur Marchand^{1,2}, *Lies Debruyne*¹, *Laure Triste*¹, *Catherine Gerrard*³, *Susanne Padel*³ and *Ludwig Lauwers*^{1,4}

ABSTRACT. Although the literature on sustainability assessment tools to support decision making in agriculture is rapidly growing, little attention has been paid to the actual tool choice. We focused on the choice of more complex integrated indicator-based tools at the farm level. The objective was to determine key characteristics as criteria for tool choice. This was done with an in-depth comparison of 2 cases: the Monitoring Tool for Integrated Farm Sustainability and the Public Goods Tool. They differ in characteristics that may influence tool choice: data, time, and budgetary requirements. With an enhanced framework, we derived 11 key characteristics to describe differences between the case tools. Based on the key characteristics, we defined 2 types of indicator-based tools: full sustainability assessment (FSA) and rapid sustainability assessment (RSA). RSA tools are more oriented toward communicating and learning. They are therefore more suitable for use by a larger group of farmers, can help to raise awareness, trigger farmers to become interested in sustainable farming, and highlight areas of good or bad performance. If and when farmers increase their commitment to on-farm sustainability, they can gain additional insight by using an FSA tool. Based on complementary and modular use of the tools, practical recommendations for the different end users, i.e., researchers, farmers, advisers, and so forth, have been suggested.

Key Words: *farm level; full assessment; rapid assessment; sustainability assessment tool; tool choice*

INTRODUCTION

Achieving greater sustainability in agricultural systems is essential for the transition toward sustainable development (OECD 1999). Sustainability assessments are a significant aid to this process (Pope et al. 2004), and a growing number of sustainability assessment tools and frameworks have been developed to support decision making in agriculture (Gasparatos 2010). This development has resulted in different tools, ranging from the farm-level to industrial-, regional-, and national-level applications (Binder et al. 2010). However, lack of research on how to choose between tools constitutes a major gap in the sustainability assessment literature. Criteria for tool choice are lacking (Gasparatos and Scolobig 2012). Currently, this choice is either “ad hoc,” given the context in which stakeholders make their decision (Gasparatos and Scolobig 2012), or pragmatic, given data availability (de Ridder et al. 2007).

Various characteristics can influence tool choice. A literature review exploring the variation in tools (see *Variation in sustainability assessment tools*) allowed us to derive some indications of such characteristics. For example, the current choice of an assessment tool usually depends on data, time, and budgetary constraints (Gasparatos and Scolobig 2012). We focus on integrated indicator-based sustainability assessment tools at the farm level. The integrated aspect is important because it relates to the broader goal of integrating sustainability concepts, i.e., ecological, economic, and social, into decision making (Pope 2006). The indicator aspect implies operational issues, such as whether the indicators are treated individually, as part of a balanced/weighted set, or combined into a composite index (Farrell and Hart 1998). Finally, the farm-level aspect concerns the use of farm-level data, the farmer as target group, and the farm as the system to evaluate. We are particularly interested in this level because the farmer makes the strategic and operational management decisions that influence farm sustainability.

Our objective is to identify key characteristics of integrated indicator-based assessment tools at the farm level that may serve as criteria for tool choice. First, an exploratory literature analysis describes the variation in sustainability assessment tools, enabling the selection of relevant cases. Then, we elaborate on the comparison methodology, which results in key characteristics for tool choice. Finally, based on these characteristics, we discuss two types of integrated indicator-based tools at the individual farm level and their possible complementary use.

VARIATION IN SUSTAINABILITY ASSESSMENT TOOLS

A number of articles perform a review, a meta-analysis, or a categorization of sustainability assessment tools (Halberg et al. 2005, de Ridder et al. 2007, Ness et al. 2007, Gasparatos et al. 2008, Binder et al. 2010, Gasparatos 2010, Gasparatos and Scolobig 2012, Singh et al. 2012). Gasparatos (2010) categorizes sustainability assessment tools into biophysical, monetary, or indicator-based tool families. Focusing on indicator-based sustainability assessment tools in agriculture, Binder et al. (2010) distinguish three types according to spatial level, i.e., farm/region, and degree of stakeholder participation: (1) top-down farm assessment methods, (2) top-down regional assessment methods with some stakeholder participation, and (3) bottom-up, transdisciplinary methods with stakeholder participation throughout the process. Although most of these categorization exercises reveal factors relevant for tool choice, none of them focus on specific choice determining characteristics of indicator-based tools at the farm level. From an exploratory literature review of farm-level tools, we present four tools (Table 1) that are recent, are integrated, are peer-review published, and have some practical implementation.

To elucidate characteristics relevant for tool choice, the literature supplies the following indications. Sustainability assessment tools differ in their starting points, objectives, and assumptions, e.g., what is important to be measured, how to measure it, and which

¹Social Sciences Unit, Institute for Agricultural and Fisheries Research (ILVO), ²Ecosystem Management Research Group and IMDO, University of Antwerp, ³The Organic Research Centre, ⁴Department of Agricultural Economics, University of Ghent

Table 1. Examples of assessments at the farm level.

Sustainability assessment tool	Original aim by tool developers	Sustainability aspects/definition of sustainability	Method of data gathering	Data source	Time for data gathering/Budget
PG (Public Goods) Tool [†]	To provide a tool to assess the public goods provided by a farm	Ecological, economic, and social aspects	Scan/interview by adviser (development toward an online autoaudit)	Combination of available accountancy data, cropping/livestock records, and farmer knowledge	2-4 h/low
IDEA (Indicateurs de Durabilité des Exploitations Agricoles) [‡]	To provide an operational tool for sustainability at farm level	Economic viability, Social Livability, Environmental reproducibility	Auto-audit (support by adviser)	Combination data available on the farm and farmer knowledge	3-6 h (after the gathering of the necessary documents)/low
RISE (Response-Inducing Sustainability Evaluation) [§]	To provide a simple and cheap but holistic monitoring tool to: 1. Evaluate the degree of sustainability at farm level; 2. Visualize potentials and failures, thus inducing management responses	Productivity, Competitiveness, Efficiency, Protection, and improvement of the natural environment and socioeconomic conditions of local communities	Scan/interview by trained adviser	Combination of available data at farm and regional level and farmer knowledge	From 3-6 h (1,5-2 days)/rather low
MOTIFS (Monitoring Tool for Integrated Farm Sustainability)	To provide a user-friendly and communicative monitoring tool that allows the measurement of progress toward integrated sustainability at dairy farms	Ecological, social, and economic aspect + entrepreneurship	Scan of readily available data, extended questionnaires, on-farm evaluations by experts	Combination of expert information and detailed farm data	Several days (> 2 days)/high

[†]Gerrard et al. 2012

[‡]Zahm et al. 2008

[§]Grenz et al. 2009

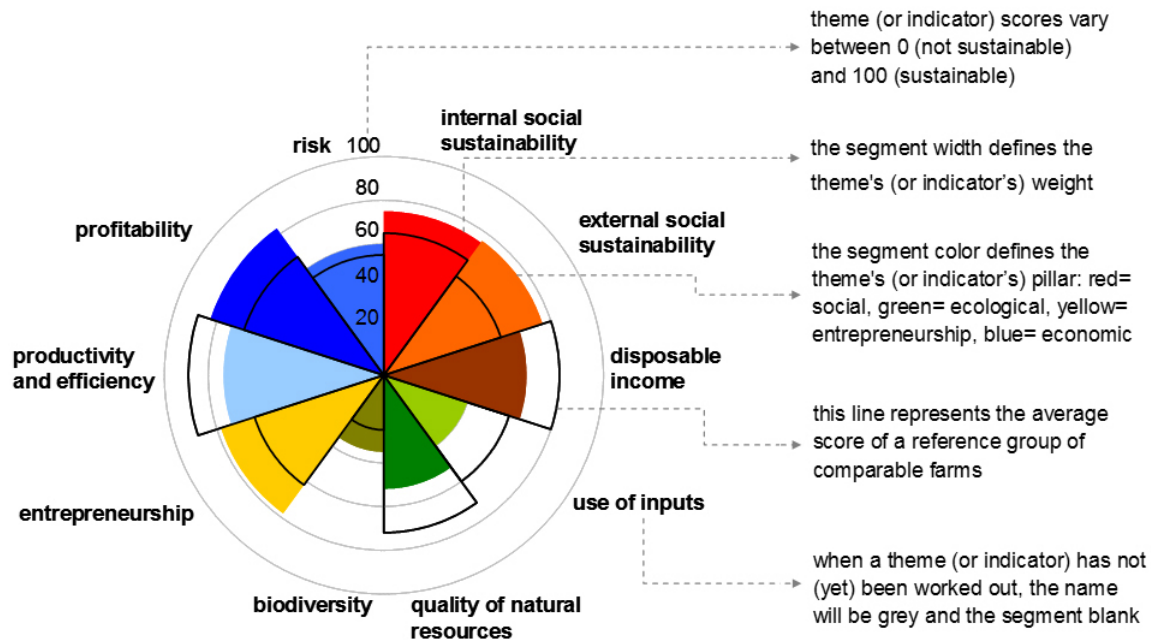
^{||}Meul et al. 2008

sustainability perspectives are both relevant and legitimate (Gasparatos 2010). Because these basic features differ between tools, the tool choice affects the outcome of the evaluation (Gasparatos 2010). A first important characteristic is the tool developers' aim: although their overall objective is similar, i.e., providing insights at the farm level to inform the farmer's decision making, specific aims may differ (Table 1). For example, a tool resulting from the aim to assess public goods will focus more on policy-relevant issues, whereas a tool resulting from the aim to measure the progress in sustainable development at the farm level will focus more on farm-relevant issues. During tool development, the aims of the tool developers are not always clear or cannot always be fulfilled (Triste et al. 2014). Additionally, the tool used in practice can offer different functions, irrespective of the developers' aim during development (Langeveld et al. 2007, Schader et al. 2014): a tool can provide a platform for communication through describing the sustainability themes, i.e., communication function (De Mey et al. 2011); it can promote the exchange of ideas and knowledge, i.e., learning function (Terrier et al. 2010, De Mey et al. 2011, Gerrard et al. 2011), and further induce management responses, i.e., management function (Grenz et al. 2009); and it can fulfill monitoring obligations (Wiek and Binder 2005, Grenz et al. 2009, Meul et al. 2012) for statutory control purposes or for product certification, i.e., monitoring and certification function (Hülsbergen 2003, Rodrigues et al. 2010). Furthermore, the methodological steps within tools may contain significant biases toward specific framings of sustainability (Bond

and Morrison-Saunders 2011). For example, when a step of indicator selection is included, end users might disregard indicators that do not fit within their particular perspective on sustainability (Hugé et al. 2013). As a result, these differences in aim, function, and methodological steps might result in varying sustainability aspects considered during the assessment (Table 1). Ideally, the decision to use a particular tool should be made on a scientifically underpinned basis considering these characteristics, resulting in a match between a certain tool or tool function and the end user's aim (de Ridder et al. 2007). Nevertheless, at present, the tool choice is mainly based on data, time, and budgetary constraints instead of scientific support (de Ridder et al. 2007, Gasparatos and Scolobig 2012).

Besides variability in the aim and sustainability aspects, Table 1 also illustrates differences in data, time, and budgetary constraints. The "method of data gathering" covers self-assessment by the farmer, an interview with a trained adviser, and experts' visits to the farm. The data source varies from the farmer's knowledge, available data such as financial accounts and cropping records, and additional detailed farm data requiring extra data-gathering efforts from the farmer, to expert information. The time needed for completing the assessment ranged from two to four hours to several days or weeks. All three of these characteristics affect a fourth characteristic, the budget required to complete the assessment.

Fig. 1. Monitoring Tool for Integrated Farm Sustainability (MOTIFS), level 1 graph and instructions on the reading and interpretation (Meul et al. 2008).



METHODS

To characterize differences between integrative indicator-based assessment tools at the farm level, we performed a comparative case study based on qualitative data. As the basis of the literature synthesis in Table 1, we selected two tools that differ substantially on the characteristics “method of data gathering,” “data source,” “time for data gathering,” and “budget”: the indicator-based Monitoring Tool for Integrated Farm Sustainability (MOTIFS) and the Public Goods Tool (PGT) both designed for dairy farms. We combined multiple sources of evidence: an in-depth literature analysis, stakeholders’ feedback on practical tool use, and experiences obtained by using both tools in similar conditions and contexts. To perform a comparative case study (Yin 2009), the construction of a framework is crucial for success. We selected the framework from Binder et al. (2010) but enhanced it with additional characteristics taken from the literature and emerging from the data.

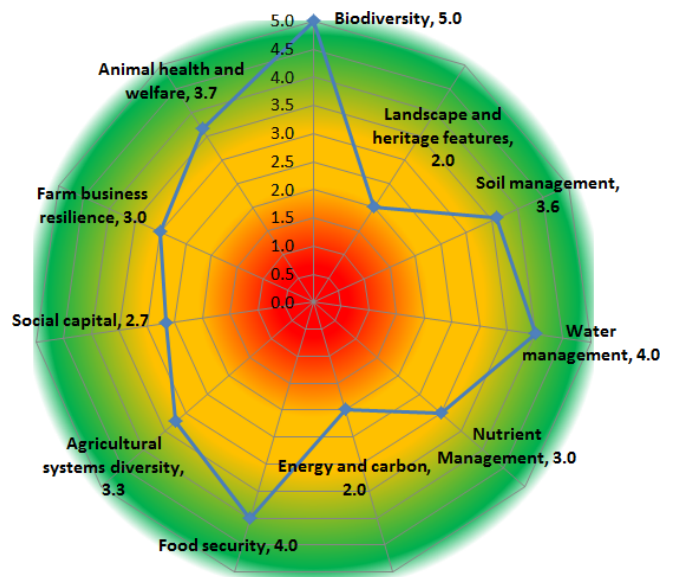
Cases

The first case is the application of MOTIFS on 31 Flemish dairy farms within a national project called *Melkveecafés* (dairy cattle cafés) and within the European Union (EU) Interreg project Dairyman. MOTIFS is a visual multilevel monitoring tool (Meul et al. 2008) that integrates economic, ecological, and social dimensions. Level 1 gives an overview of the farm’s overall sustainability (Fig. 1). Level 2 gives an overview of the sustainability themes within a specific sustainability dimension. Level 3 visualizes the indicator scores for a specific theme.

Second, PGT was applied on 7-14 organic or low-input dairy cow and goat farms in 8 countries participating in the EU-7th Framework Programme SOLID project. The authors were responsible for adjustments to the tool and performed the

assessments in Flanders, the Netherlands, and Great Britain. PGT, originally designed to assess the public goods provided by an organic farm in Great Britain (Table 1; Gerrard et al. 2012), was adjusted to be more applicable EU wide to dairy farms. The resulting tool (Fig. 2) assesses the sustainability performance of an individual farm and not only its supply of public goods.

Fig. 2. Radar diagram showing an example from the pilot assessment (Gerrard et al. 2011).



Literature sources

Both scientific and mainstream publications were used. Mulier et al. (2004) describe the options for MOTIFS tool design and the (dis)advantages in the development of an instrument to monitor sustainability at the farm level. Nevens et al. (2008) provide input for a tool in Flanders using a participatory visioning process. Meul et al. (2008) describe the methodology and actual development of MOTIFS for Flemish dairy farms. Meul et al. (2009), Campens et al. (2010), and De Mey et al. (2011) report on validation and implementation steps of MOTIFS; whereas Marchand et al. (2010) and Triste et al. (2014) report on the learning and development process of MOTIFS. Gerrard et al. (2011) describe PGT development, spurs, and activities. Gerrard et al. (2012) report on the methodology, development, and feedback about PGT.

Stakeholders' feedback

Stakeholders' feedback about practical use of the tools was gathered and already published, but we have also extensively used data from interviews and questionnaires. De Mey et al. (2011) conducted 15 semistructured interviews with 6 farm advisers and 3 accountants during several stages of MOTIFS implementation. Marchand et al. (2010) carried out 19 interviews with farmers to describe the role of MOTIFS in a learning process by the farmers. Triste et al. (2014) conducted 12 interviews with the tool developers and researchers who implemented the tool to reveal lessons learned about the development process of MOTIFS. Through questionnaires, Gerrard et al. (2012) obtained feedback from 12 farmers on the use of the PGT.

Additional experiences from practice

To allow for a more straightforward comparison, two researchers involved in implementing MOTIFS in the *Melkveecafés* and Dairyman projects also used the PGT during the SOLID project. Reflections on both tools were systematically written down in notes and reports during local meetings, during farmer discussion sessions, and during informal conversations with farmers, advisers, tool developers, researchers implementing the tool, and the scientific community, i.e., researchers not using the tools.

The analytical framework enhanced with additional characteristics

We relied on the framework from Binder et al. (2010) to structure the data. This framework separately analyzes the normative, systemic, and procedural dimensions of the tools: (1) How can the sustainability of a system be assessed? (2) Is a system properly described by the set of indicators used? (3) How is the assessment carried out? To enrich the analysis, we added characteristics based on the literature and confirmed by our data. From De Mey et al. (2011), who determined critical success factors (CSFs) for implementation of an integrated sustainability assessment (ISA) tool, we selected relevant characteristics for our analysis (Table 2). The CSFs "attitude of model users toward sustainability" and "organization of discussion sessions" are not related to the type of tool and are therefore excluded from the analysis. Two additional characteristics emerged from the data and have been confirmed by the literature: "output accuracy" or precision of the results (Schader et al. 2014) and "tool functions" (de Ridder et al. 2007). The addition of these characteristics resulted in an enhanced framework (Table 3) to perform the comparative analysis.

Table 2. Critical success factors for implementation of integrated sustainability assessment tools, according to De Mey et al. (2011).

Critical success factor	Description
Attitude of model users toward sustainability	Values and beliefs of the model users (advisers and farmers) regarding sustainability issues.
Compatibility	Extent to which the design and the proposed use of the tool is compatible with the data systems and institutional structure of accountancy/consultancy agencies.
User-friendliness	Extent to which the ISA-tool is flexible and easy to use. This is related to the graphical design, ease of assessment, and calculation (automation), etc.
Data availability	Availability of data necessary for indicator calculation.
Transparency	Transparency of the used model and data (design, generalizations, etc.) and transparency on uncertainties of model-derived results.
Data correctness	Correctness of the data that are used to calculate the indicators of the ISA-tool.
Communication aid	Use of ISA-tool in discussion sessions and its ability to support discussion on sustainability. Both communication aid of the model itself as communication through using it in farmer groups are included.
Complexity	Degree of complexity of the ISA-tool.
Organization of discussion sessions	Practical organization of the discussion sessions with farmers. Which aspects need to be considered to make the discussion sessions more successful.
Effectiveness	Extent to which the ISA-tool is perceived as being relevant to use and implement.

Triangulation

During data collection and analysis, we used various triangulation types to ensure objectivity and to increase the validity (Guion et al. 2002, Golafshani 2003, Koro-Ljungberg 2008). First, data triangulation was applied by using different sources or stakeholders, i.e., farmers, advisers, accountants, researchers, tool developers, and so forth. Second, methodological triangulation was ensured using multiple methods to gather data, i.e., scientific and popular articles, interviews and questionnaires, experience with tool implementation, discussion notes, and reports. Furthermore, investigator triangulation was implemented because two researchers implemented both tools, conducted the analysis, and discussed the results.

RESULTS

Normative aspects

The MOTIFS goal resulted from a vision statement, which was developed in a transdisciplinary dialogue with various stakeholders (Table 3). The Brundtland sustainability definition was interpreted mainly from a farmer's perspective, resulting in a clear normative perception of sustainability. The major principles of this vision were translated into 10 concrete themes by the tool developers, whereas indicators were developed in collaboration with experts. The developers of MOTIFS describe different

Table 3. Comparative analysis of two integrative indicator-based sustainability assessment tools based on a framework adapted from Binder et al. (2010).

Dimension	Phase	Characteristic	MOTIFS	PG tool		
Normative dimension		Concept	Sustainability concept of the Brundtland report (WCED 1987)	Providing a “public good” beyond the simple production of food, adjusted with farm related sustainability aspects		
		Way of goal setting	Combination of transdisciplinary (vision and themes) and top-down process (subthemes and indicators)	Combination of stakeholder involvement and top down		
		Scoring and aggregation method	1. Scoring method with indicators normalized on a scale between 0 and 100 with different benchmark methods. 2. Visual aggregation into an adapted radar graph, integrating 10 sustainability themes (ecological, economic, and social aspects + entrepreneurship)	1. Scoring method with indicators ranging between 1 and 5 and reference values when available (categorized qualitative data). 2. Aggregation into 11 spurs, covering the 3 dimensions (ecological, economic, and social) on a radar diagram . A bar chart shows the activities on each spur to present more detailed information		
		Tool function [†]	Monitoring, communication, learning, and management function	Focus on the communication/learning function		
Systemic dimension		Parsimony (as much simplicity as possible)	No clear goal, however, avoiding double counting indicates parsimony	An implicit goal by keeping the measurement phase within 4 hours		
		Sufficiency (as much complexity as necessary)	An implicit goal during the visioning phase	A goal through providing sufficient in depth information		
		Indicator interaction	AA clear goal, however, not yet implemented in the tool	Not a goal		
Procedural dimension	Preparatory phase		Preparing data gathering and organizing discussion groups: contact with farmers, advisers, experts	Organizing individual farm visits by adviser		
	Phase of indicator selection		Not present: a standard set of indicators addressing key environmental, economic, and social processes	Not present: a standard set of indicators addressing key environmental, economic, and social processes		
	Measurement phase (quantification of indicators: statistical data, surveys, or categorized qualitative data)	Data correctness [†]		Valued as high	Farmer’s knowledge might be less accurate or influenced by perceptions	
		Data availability [†]		Depending on time and budget	Readily available data, low time and budget requirements	
		User-friendliness [†]		Different excel sheets, time consuming, rather complicated procedure, difficult to interpret results	One excel sheet, simple procedure	
	Assessment phase (benchmarking, aggregation, integration)	Compatibility [†]		Depending on the accountancy system used by the farmer	Rarely a problem	
		Transparency [†]		Rather low due to benchmarking methods and build-in assumptions	Simple and visible scoring procedure	
		Applicability of assessment results and follow up	Output accuracy [†]		Highly valued because of expert-based scientifically underpinned indicators	Concern as to whether a change in management practice relates directly to a change in sustainability resulting in distrust
			Complexity [†]		Complex procedure, difficult presentation of results, and therefore difficult to interpret the results	Simple procedure, straightforward presentation of results
		Communication aid [†]		Extra discussion group meetings are needed to fulfill this function	By the use of the tool, farmer and adviser can engage them in a positive discussion	
	Effectiveness [†]		Depending on the function the end users aim at and the approach used	Farmer and adviser indicate the effectiveness of the tool for the learning function		

[†]Extra characteristics added to the original framework of Binder et al. (2010)

ambitions and functions of the tool (Mulier et al. 2004, Meul et al. 2008, Triste et al. 2014). The initial aim was to develop a tool with a monitoring and communication function (Table 1). The tool implementation process led the tool developers to recognize other functions for the tool: learning and management functions (Triste et al. 2014). However, each function implies a specific use of the tool. The monitoring function is only appropriate for indicators that are quantitative and based on accurate data. The communication or learning function suggests tool use in a social setting, for example, in farmer discussion groups with the attendance of an adviser or expert. Finally, a management function requires a tool able to generate concrete management advice. This feature is not included in MOTIFS, but farm advisers familiar with the farm can provide management support by using the information generated from the tool.

The PGT resulted from an increased interest amongst policy makers regarding farms providing a “public good” other than food production (Gerrard et al. 2012). The goal setting occurred through a literature review and a stakeholder meeting involving researchers, agricultural advisers, and representatives from Natural England, an executive nondepartmental public body. The original tool was adjusted to be applicable across Europe and was given a greater focus on dairy farms. This resulted in a more extensive farm scan on sustainability instead of focusing exclusively on public goods from a societal perspective. According to the developers of the PGT, it is not a policy tool or a tool to provide verifiable assessments but is now designed to stimulate the debate about sustainability with farmers focusing on learning rather than monitoring. Furthermore, it can be used to illustrate the variety of farms within a region, and it can reveal problems and issues that need attention as well as areas of good farm performance.

Systemic aspects

Neither tool uses parsimony or sufficiency as an explicit criterion for indicator selection, but they are implicitly present (Table 3). For example, for MOTIFS, parsimony was indirectly important because the tool developers avoided having the same aspect represented within two indicators (Triste et al. 2014). In its present form, the tool can visualize trade-offs between indicators when the same input data is used for different indicators. However, the links and trade-offs between indicators, i.e., indicator interaction, are not built into the formulas or aggregation methods. Therefore, trade-offs between indicators based on different data input are not always detectable, although this was a goal during the design of MOTIFS (Mulier et al. 2004). System representation was thus implicitly taken into account during the design of MOTIFS but is not yet scientifically assured.

In PGT, data requirements, qualitative and quantitative, for indicator calculation in each spur were sufficient to provide in-depth information on the farm performance on that spur, i.e., sufficiency. Initially, system representation and indicator interaction were not explicitly pursued. However, when the tool was broadened to include additional farm aspects, such as entrepreneurship and financial risk assessment, system representation was addressed qualitatively through discussions with all project partners.

Procedural aspects

Preparatory phase and indicator selection

Using MOTIFS requires a great deal of preparation regarding data collection from farmers, advisers, and experts. The PGT was used on an individual basis, i.e., farmer-adviser, and farmers were informed in advance about the data required during the assessment. For both tools, a phase of indicator selection in the implementation of either tool is absent because the indicator set is given.

Measurement phase

The characteristics “method of data gathering,” “data source,” “time for data gathering,” and “budget” (Table 1) have an effect on the CSFs “compatibility,” “user-friendliness,” “data availability,” and “data correctness” (Table 2), which are mainly important during the measurement phase. The need for very specific data or expert information when using MOTIFS implies a substantial amount of time and money needed to gather the necessary data. In addition, the calculation is a time-consuming and complicated procedure spread over different Excel files. This results in a rather low user-friendliness, making the tool unsuitable for independent use by farmers. Furthermore, a compatibility problem arises when data gathering by farm accountancy systems is not uniform within a discussion group. This creates a lack of comparable data. However, the specific data and expert information guarantee a high level of data correctness. Although the tool developers and the scientific community strongly believe in the data correctness of MOTIFS, the advisers did notice that farmers do not always gather their data systematically or correctly. This might result in reduced accuracy and a quality discrepancy of various indicators between farms. PGT, on the other hand, is straightforward and uses directly available data from farm accounts, cropping and livestock records, and the farmer’s knowledge about a number of key activities. As such, PGT can easily be used independently from existing accountancy systems. Some expert judgments, e.g., related to energy use or nutrient balances, are built into the underlying calculations, but experts are not required during the practical use of the tool. As a result, PGT scores high on “compatibility” and “data availability.” The “user-friendliness,” of PGT was also ranked high because results were immediately available from a single Excel file after an assessment of two to four hours. Nevertheless, the tool might compromise on “data correctness” because the farmer’s knowledge can be biased by the farmer’s perception. The experiences from the researchers implementing the PGT confirm the risk of rather low data correctness. However, a difference in the answer to one question does not create a high difference in the results at the spur level. Furthermore, because of the mainly qualitative interpretation and the tool’s learning function, this did not create as much of a problem as when a quantitative, rigorous assessment was being attempted.

Assessment phase

The assessment phase comprises the scoring and aggregation methods, already presented in the normative dimension. These mainly affect the CSF “transparency.” The specific data demands of MOTIFS are in most cases more objective, ensuring the comparability of different farms and the possibility to develop benchmarks. Benchmarks are constructed directly from the observed outcomes of the indicators, which is a strength when

comparing farms. However, this might result in less transparency, e.g., when transforming absolute values into relative ones on a 0-100 scale. In PGT, comparison between farms has to be performed cautiously because the data are derived from the interview with the farmer, referring back to accounts, and can be more subjective. Nevertheless, the “transparency” of PGT is high because questions are scored between 1 and 5, and the score for each answer is visible in the Excel file.

Application and follow-up

Application and follow-up relate to the opportunity for user groups to use the assessment results. For this reason, the “output accuracy,” “complexity,” “communication aid,” and “effectiveness” become relevant. The output accuracy of MOTIFS, important for the monitoring function, is highly valued by the scientific community because many indicators are scientifically underpinned and peer-review published. MOTIFS results allow the farmers to situate themselves within a benchmark and provide a basis for identifying successful farm management practices. However, some of the indicators are still “under construction” because they have not yet been designed using a scientifically underpinned method. For example, complex social issues are difficult to monitor. Both researchers and farmers sometimes question the output accuracy of PGT: for example, questioning whether a score of 3.5 is better than a score of 3.2, and whether a score of 3 for nutrient management is actually equivalent to a score of 3 for farm business resilience. These doubts are also related to the management questions defining the scores. A change in management will change the scores, but it is not clear whether this will always have a significant effect on the farm sustainability. The validity and accuracy of the PGT scores were assessed during an additional discussion with the farm advisers who had used the tool during the SOLID project. The farm advisers were able to correlate the individual scores of each farm with their knowledge of the farm. They confirmed that the results are in line with their expectations. This gives more validity to the scores from PGT. Furthermore, they argued that the resulting PGT graphs are very alike within 2 groups of similar farms, which probably reflects results from similar farm management strategies within those groups.

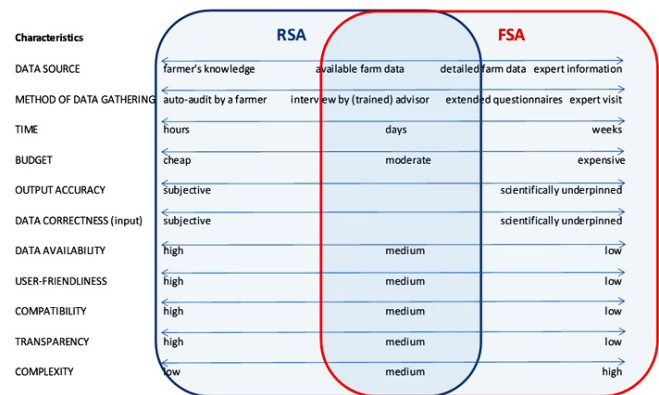
Considering the CSF “complexity,” MOTIFS aims to grasp the complexity of sustainability through a multilevel monitoring approach, and its graphical design tries to visualize this in a simple way. However, advisers and farmers perceived the tool as rather complex. They first needed some familiarity with the complex presentation of the tool, before they were able to interpret it. In addition, trade-offs were difficult to identify. Furthermore, farmers indicated that they needed the assistance of an adviser to change their management using the results of MOTIFS. Therefore, they claimed that only committed farmers would use such tools. In contrast, advisers strongly believe in the communication aspect of the PGT. PGT does not aim to include the full complexity of sustainability but focuses on farm sustainability and public goods provided by a farm. The use of the tool is more straightforward than the MOTIFS tool. The ability of the PGT to illustrate a graph at the end of the discussion was positively perceived by many farmers, i.e., an immediate “return on investment.” The effectiveness of MOTIFS implementation depends on the desired function, presented in the normative dimension. According to advisers, the effectiveness of

PGT for the learning function is high because meeting the farmer in person enables positive discussions. Farmers reported an increased level of knowledge and understanding of public goods and sustainability after taking part in the assessment. Moreover, farmers believed that advisers could provide an added value, compared to a mere web-based tool, by offering experience from other farms and farming systems and possible courses of action and interpretations.

DISCUSSION

The comparative analysis clearly shows differences between the tools. For 11 characteristics, we observed a continuum between the 2 extremes (Fig. 3). We propose 2 working definitions of these extremes: full sustainability assessment (FSA) and rapid sustainability assessment (RSA) tools. FSA tools make use of detailed farm data and/or expert information, need trained advisers and/or expert visits to gather the data, and are rather long and expensive in duration. RSA tools represent the other side of the spectrum. They make use of the farmer’s knowledge or readily available data, allow an audit by the farmer or an adviser, and are relatively short in duration. As described in the *Results*, these constraints about data, time, and budget affect additional characteristics, such as output accuracy, data correctness and availability, user-friendliness, compatibility, transparency, and complexity.

Fig. 3. Characteristics describing full sustainability assessment (FSA) and rapid sustainability assessment (RSA) tools.



Although case specific, the differences in these characteristics enable us to discuss the strengths and weaknesses of both tool types in relation to the aim of the end user and the tool function in practice. Higher output accuracy is a significant strength linked to the specific and large data demand in the FSA tool. Furthermore, an FSA tool in which a well-considered system representation is guaranteed can take the complexity of the sustainability concept into account during its design. In addition, through the use of benchmarking, FSA can be a highly suitable tool for comparing different farms. Consequently, an FSA tool based on expert information with a well-considered system representation can be highly suitable for monitoring sustainability aspects of farming systems. When accurate data for a sufficient number of farms can be gathered, comparing different farming systems is also possible. An FSA tool can thus potentially be used for certification. For example, farms that meet certain thresholds

could be certified based on their sustainable production or products. The management function can be strengthened when the knowledge of an adviser or expert on the local farming conditions complements the use of the tool (Meul et al. 2012). By using the tool in discussion groups, an FSA tool can be helpful as a communication tool stimulating social learning (Marchand et al. 2010). However, because of the time-consuming, expensive data gathering and complex data processing, the user-friendliness of an FSA tool for a farmer or adviser is rather low. Results indicate that only highly motivated farmers will be reached with this approach. For the communication function, the time and cost investment inherent in using an FSA tool make it less attractive than other faster and cheaper alternatives.

The strength of the RSA tool is the ease of data collection. Such tools can either use widely available farm accountancy and management data or rely solely on farmer knowledge to structure the dialogue with regard to sustainability between the farmers and advisers through headlines and simple scoring systems (e.g., Measures 2004). RSA tools are more user-friendly and less complex than FSA tools, using transparent and understandable indicators. The main strength of these RSA tools is to raise awareness and make a farmer think about different issues related to sustainability. The higher user-friendliness seems to trigger farmers to be more willing to use RSA tools. These strengths also imply a weakness, however. The subjectivity inherent in some easily accessible data sources results in a rather low output accuracy compared with an FSA tool. The indicators of the PGT are based on management options. The main question is whether these practices are sufficiently good proxies for the sustainability impacts that they aim to measure. As a result, the tool can be used for monitoring such management practices, but farmers or advisers may lack trust in its ability to monitor sustainability. This lack of trust makes these tools less suitable for monitoring sustainability for, e.g., certification purposes.

Sustainability assessment studies of different kinds and levels have come to the conclusion that no sustainability assessment tool is “one size fits all” (de Ridder et al. 2007, Schader et al. 2014). Our research confirms this statement when considering indicator-based tools at the farm level. For example, MOTIFS aims to be a communicative monitoring tool suitable for social learning (Meul et al. 2008, Marchand et al. 2010, De Mey et al. 2011). This multifunction aim created tensions that are reflected in indicator choices. The objective of monitoring resulted in the development of indicators with a high-precision measurement. However, to be suited for learning purposes, indicators need to be understandable and transparent. For example, for indicators such as those describing social issues, combining these functions is a problematic task. Their complexity cannot be monitored correctly with easily understandable and transparent calculations. Consequently, not all indicators can comply with the high accuracy standards for monitoring, contrary to the expectations of farmers and advisers. On the other hand, the developers of the PGT focused on the learning aspect for the farmer rather than the monitoring aspect. This created transparent and communicative indicators with a more uniform quality level. When selecting a tool, we therefore suggest first choosing a clear and well-defined function: either monitoring or learning. Combining both functions within one tool has proved to be ineffective (Triste et al. 2014).

During various implementations of both FSA and RSA tools, we observed that certain end users, mainly farmers and advisers, wished to adjust or select indicators depending on the goal of the project or the characteristics of participating farmers, i.e., sector, available data, and so forth. The end users should therefore be able to select the modules and/or indicators depending on the goal of the project, the context-specific needs and conditions, or the characteristics of participating farms. Therefore, modularity of tools is a challenge for tool developers. Furthermore, similar to higher level tools (Gasparatos 2010, Bond and Morrison-Saunders 2011), the end users should be aware that the farm-level tools are built starting from a certain framework, vision, or values and therefore can contain significant biases toward specific framings or sustainability perspectives (Schader et al. 2014).

CONCLUSIONS

We revealed multiple differences within the integrated indicator-based sustainability assessment tools at the farm level. As these differences in characteristics may also act as criteria for tool choice, we further revealed that FSA and RSA tools can coexist and have complementary strengths and weaknesses. RSA tools are suitable for use by larger groups of farmers. They are more directed toward learning and can act as a trigger for farmers to become interested in farm sustainability. Furthermore, such an assessment can raise the awareness and reveal particular problems or barriers in the development toward farm sustainability. When farmers have taken an interest in sustainability, or certain specific aspects, they can focus on monitoring these particular aspects by using an FSA tool. These farmers will have to be more motivated because they will need to spend more time and money on this monitoring process. The FSA can support their continued learning through discussing and comparing their individual results in farmers’ small discussion groups. The management function can be strengthened if a farm adviser adds farm-specific knowledge. Earlier research (Folke et al. 2003, Darnhofer 2010, Darnhofer et al. 2010) also indicates that by combining different types of information sources and sharing this information in various networks, farmers’ learning about sustainability can benefit. Therefore, future research should focus on the different types of information sources provided by complementary use of FSA and RSA tools. Research on modularity of tools can be a second focus for future research on ISA tools. For example, a clear methodology for using modular tools is a challenge because certain characteristics, e.g., system representation, must be ascertained to achieve sustainable farm development.

Responses to this article can be read online at:

<http://www.ecologyandsociety.org/issues/responses.php/6876>

Acknowledgments:

We are grateful to all the farmers and advisers for participating. Part of the work for this publication was generated through (1) the Dairyman project, which was funded by the Interreg IVB NWE Programme of the European Union; and (2) the SOLID project, agreement no. 266367 (<http://www.solidairy.eu/>), with financial support from the European Community under the 7th Framework

Programme. The publication reflects the views of the author(s) and not those of the European Community, which is not to be held liable for any use that may be made of the information.

LITERATURE CITED

- Binder, C. R., G. Feola, and J. K. Steinberger. 2010. Considering the normative, systemic and procedural dimensions in indicator-based sustainability assessments in agriculture. *Environmental Impact Assessment Review* 30:71-81. <http://dx.doi.org/10.1016/j.eiar.2009.06.002>
- Bond, A. J., and A. Morrison-Saunders. 2011. Re-evaluating sustainability assessment: aligning the vision and the practice. *Environmental Impact Assessment Review* 31:1-7. <http://dx.doi.org/10.1016/j.eiar.2010.01.007>
- Campens, V., K. De Mey, J. D'hooghe, and F. Marchand. 2010. *Melkveecafé: Samen grenzen verleggen*. Institute for Agriculture and Fisheries Research (ILVO) Publication No. 74. ILVO, Merelbeke, Belgium.
- Darnhofer, I. 2010. Strategies of family farms to strengthen their resilience. *Environmental Policy and Governance* 20:212-222. <http://dx.doi.org/10.1002/eet.547>
- Darnhofer, I., S. Bellon, B. Dedieu, and R. Milestad. 2010. Adaptiveness to enhance the sustainability of farming systems. A review. *Agronomy for Sustainable Development* 30:545-555. <http://dx.doi.org/10.1051/agro/2009053>
- De Mey, K., K. D'Haene, F. Marchand, M. Meul, and L. Lauwers. 2011. Learning through stakeholder involvement in the implementation of MOTIFS: an integrated assessment model for sustainable farming in Flanders. *International Journal of Agricultural Sustainability* 9:350-363.
- de Ridder, W., J. Turnpenny, M. Nilsson, and A. von Raggamby. 2007. A framework for tool selection and use in integrated assessment for sustainable development. *Journal of Environmental Assessment Policy and Management* 9:423-441. <http://dx.doi.org/10.1142/S1464333207002883>
- Farrell A., and M. Hart. 1998. What does sustainability really mean? The search for useful indicators. *Environment: Science and Policy for Sustainable Development* 40:4-31. <http://dx.doi.org/10.1080/00139159809605096>
- Folke, C., J. Colding, and F. Berkes. 2003. Synthesis: building resilience and adaptive capacity in social-ecological systems. Pages 352-387 in F. Berkes, J. Colding, and C. Folke, editors. *Navigating social-ecological systems. Building resilience for complexity and change*. Cambridge University Press, Cambridge, UK. <http://dx.doi.org/10.1017/CBO9780511541957.020>
- Gasparatos, A. 2010. Embedded value systems in sustainability assessment tools and their implications. *Journal of Environmental Management* 91:1613-1622. <http://dx.doi.org/10.1016/j.jenvman.2010.03.014>
- Gasparatos, A., M. El-Haram, and M. Horner. 2008. A critical review of reductionist approaches for assessing the progress towards sustainability. *Environmental Impact Assessment Review* 28:286-311. <http://dx.doi.org/10.1016/j.eiar.2007.09.002>
- Gasparatos, A., and A. Scolobig. 2012. Choosing the most appropriate sustainability assessment tool. *Ecological Economics* 80:1-7. <http://dx.doi.org/10.1016/j.ecolecon.2012.05.005>
- Gerrard, C. L., L. G. Smith, S. Padel, B. Pearce, R. Hitchings, M. Measures, and N. Cooper. 2011. *OCIS public goods tool development*. Organic Research Centre Report. Organic Research Centre, Berkshire, UK.
- Gerrard, C., L. Smith, B. Pearce, S. Padel, R. Hitchings, M. Measures, and N. Cooper. 2012. Public goods and farming. Pages 1-22 in E. Lichtfouse, editor. *Farming for food and water security. Sustainable Agriculture Reviews Volume 10*. Springer, Dordrecht, The Netherlands. http://dx.doi.org/10.1007/978-94-007-4500-1_1
- Golafshani, N. 2003. Understanding reliability and validity in qualitative research. *Qualitative Report* 8:597-607.
- Grenz, J., C. Thalmann, A. Stämpfli, C. Studer, and F. Häni. 2009. RISE – a method for assessing the sustainability of agricultural production at farm level. *Rural Development News* 1(2009):5-6.
- Guion, L., D. C. Diehl, and D. McDonald. 2002. *Triangulation: establishing the validity of qualitative studies*. Publication No. FCS6014. Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida, USA. [online] URL: <http://edis.ifas.ufl.edu/fy394>
- Halberg, N., H. M. G. van der Werf, C. Basset-Mens, R. Dalgaard, and I. J. M. de Boer. 2005. Environmental assessment tools for the evaluation and improvement of European livestock production systems. *Livestock Production Science* 96:33-50. <http://dx.doi.org/10.1016/j.livprodsci.2005.05.013>
- Hugé, J., T. Waas, F. Dahdouh-Guebas, N. Koedam, and T. Block. 2013. A discourse-analytical perspective on sustainability assessment: interpreting sustainable development in practice. *Sustainability Science* 8:187-198. <http://dx.doi.org/10.1007/s11625-012-0184-2>
- Hülsbergen, K. J. 2003. *Entwicklung und Anwendung eines Bilanzierungsmodells zur Bewertung der Nachhaltigkeit landwirtschaftlicher Systeme*. Shaker Verlag, Aachen, Germany.
- Koro-Ljungberg, M. 2008. Validity and validation in the making in the context of qualitative research. *Qualitative Health Research* 18:983-989. <http://dx.doi.org/10.1177/1049732308318039>
- Langeveld, J. W. A., A. Verhagen, J. J. Neeteson, H. van Keulen, J. G. Conijn, R. L. M., Schils, and J. Oenema. 2007. Evaluating farm performance using agri-environmental indicators: recent experiences for nitrogen management in the Netherlands. *Journal of Environmental Management* 82:363-376. <http://dx.doi.org/10.1016/j.jenvman.2005.11.021>
- Marchand, F., K. De Mey, L. Debruyne, K. D'Haene, M. Meul, and L. Lauwers. 2010. From individual behavior to social learning: start to a participatory process towards sustainable agriculture. Pages 670-682 in I. Darnhofer and M. Grötzer, editors. *Proceedings of the 9th European International Farming Systems Association (IFSA) Symposium: building sustainable rural futures* (Vienna, Austria, 4-7 July 2010). IFSA Europe Group, Vienna, Austria.

- Measures, M. 2004. Farm auditing for sustainability. Pages 27-30 in A. Hopkins, editor. *Organic farming: science & practice for profitable livestock and cropping*. British Grassland Society, Reading, UK.
- Meul, M., F. Nevens, and D. Reheul. 2009. Validating sustainability indicators: focus on ecological aspects of Flemish dairy farms. *Ecological Indicators* 9:284-295. <http://dx.doi.org/10.1016/j.ecolind.2008.05.007>
- Meul, M., S. Van Passel, D. Fremaut, and G. Haesaert. 2012. Higher sustainability performance of intensive grazing versus zero-grazing dairy systems. *Agronomy for Sustainable Development* 32:629-638. <http://dx.doi.org/10.1007/s13593-011-0074-5>
- Meul, M., S. Van Passel, F. Nevens, J. Dessein, E. Rogge, A. Mulier, and A. Van Hauwermeiren. 2008. MOTIFS: a monitoring tool for integrated farm sustainability. *Agronomy for Sustainable Development* 28:321-332. <http://dx.doi.org/10.1051/agro:2008001>
- Mulier A., F. Nevens, D. Reheul, and E. Mathijs. 2004. *Ontwikkeling van een beoordelingssysteem voor de duurzaamheid van de Vlaamse land- en tuinbouw op bedrijfsniveau*. Stedula Publication No. 9. Steunpunt Duurzame Landbouw (Stedula), Gontrode, Belgium.
- Ness, B., E. Urbel-Piirsalu, S. Anderberg, and L. Olsson. 2007. Categorizing tools for sustainability assessment. *Ecological Economics* 60:498-508. <http://dx.doi.org/10.1016/j.ecolecon.2006.07.023>
- Nevens F., J. Dessein, M. Meul, E. Rogge, I. Verbruggen, A. Mulier, S. Van Passel, J. Lepoutre, and M. Hongenaert. 2008. 'On tomorrow's grounds', Flemish agriculture in 2030: a case of participatory translation of sustainability principles into a vision for the future. *Journal of Cleaner Production* 16:1062-1070. <http://dx.doi.org/10.1016/j.jclepro.2007.06.007>
- Organization for Economic Cooperation and Development (OECD). 1999. *Environmental indicators for agriculture. Issues and Design*. OECD, Paris, France.
- Pope, J. 2006. Editorial: what's so special about sustainability assessment? *Journal of Environmental Assessment Policy and Management* 8:v-x. <http://dx.doi.org/10.1142/S1464333206002505>
- Pope, J., D. Annandale, and A. Morrison-Saunders. 2004. Conceptualising sustainability assessment. *Environmental Impact Assessment Review* 24:595-616. <http://dx.doi.org/10.1016/j.eiar.2004.03.001>
- Rodrigues, G. S., I. A. Rodrigues, C. C. de Almeida Buschinelli, and I. de Barros. 2010. Integrated farm sustainability assessment for the environmental management of rural activities. *Environmental Impact Assessment Review* 30:229-239. <http://dx.doi.org/10.1016/j.eiar.2009.10.002>
- Schader, C., J. Grenz, M. S. Meier, and M. Stolze. 2014. Scope and precision of sustainability assessment approaches to food systems. *Ecology and Society* 19(3): 42. <http://dx.doi.org/10.5751/ES-06866-190342>
- Singh, R. K., H. R. Murty, S. K. Gupta, and A. K. Dikshit. 2012. An overview of sustainability assessment methodologies. *Ecological Indicators* 15:281-299. <http://dx.doi.org/10.1016/j.ecolind.2011.01.007>
- Terrier, M., P. Gasselin, and J. Le Blanc. 2010. Assessing the sustainability of activity systems to support agricultural households' projects. Pages 812-822 in I. Darnhofer and M. Grötzer, editors. *Proceedings of the 9th European International Farming Systems Association (IFSA) Symposium: building sustainable rural futures* (Vienna, Austria, 4-7 July 2010). IFSA Europe Group, Vienna, Austria.
- Triste, L., F. Marchand, L. Debruyne, M. Meul, and L. Lauwers. 2014. Reflection on the development process of a sustainability assessment tool: learning from a Flemish case. *Ecology and Society* 19(3): 47. <http://dx.doi.org/10.5751/ES-06789-190347>
- Wiek, A., and C. Binder. 2005. Solution spaces for decision-making—a sustainability assessment tool for city-regions. *Environmental Impact Assessment Review* 25:589-608. <http://dx.doi.org/10.1016/j.eiar.2004.09.009>
- Yin, R. 2009. *Case study research. Design and methods*. Fourth edition. Sage, Thousand Oaks, California, USA.
- Zahm, F., P. Viaux, L. Vilain, P. Girardin, and C. Mouchet. 2008. Assessing farm sustainability with the IDEA method – from the concept of agriculture sustainability to case studies on farm. *Sustainable Development* 16:271-281. <http://dx.doi.org/10.1002/sd.380>