Check for updates

METHOD ARTICLE

Monitoring effectiveness in sustainability demonstrations: a reproducible framework integrating stakeholder engagement and continuous feedback

[version 1; peer review: 1 approved with reservations]

Ilias Karachalios[1], Nikolaos D. Tantaroudas[2], Christos Makropoulos[1]

[1]Department of Water resources and environmental engineering, School of Civil Engineering, National Technical University of Athens, Zografou, Attica, GR 157 80, Greece
[2]Institute of Communication and Computer Systems, Athens, Zografou, GR 157 73, Greece

v1 **First published:** 30 Jun 2026, **6**:219
https://doi.org/10.12688/openreseurope.22944.1
Latest published: 30 Jun 2026, **6**:219
https://doi.org/10.12688/openreseurope.22944.1

Abstract

Background

Real-world sustainability demonstrations usually need to balance two competing requirements: local adaptation to specific contexts and comparability across pilots, countries, and stakeholder groups.

Methods

This Method Article presents the monitoring framework developed in the CHOICE project to address that challenge across five heterogeneous pilot demonstrations in Austria, Spain, Greece, Colombia, and South Africa. The article documents the overall monitoring architecture, the operational definition of effectiveness, and the logic used to compare diverse interventions without requiring identical indicators in all settings. The framework combines five evaluation dimensions - reliability, user acceptance, user-friendliness, transferability, and behavioural conversion - with a mixed-methods evidence pipeline integrating digital analytics, stakeholder questionnaires, qualitative inquiry, Participatory Labs, and structured internal reporting. A conversion-types pathway is used to trace progression from awareness and interest to engagement, action, and continuation. Stakeholder engagement and feedback are treated as methodological inputs that support iterative adaptation during implementation rather than as separate project activities. The framework also incorporates risk management, privacy safeguards,

Open Peer Review

Approval Status ?

	1
version 1 30 Jun 2026	? view

1. **Elias Appiah-Kubi**, Universitatea din Petrosani, Petroșani, Romania
Codruta Cornelia Dura, University of Petrosani, Petrosani, Romania

Any reports and responses or comments on the article can be found at the end of the article.

and the use of A/B testing where feasible.

Results

This article does not report pilot outcomes.

Conclusions

The article provides a reproducible methodological blueprint to guide data collection, interpretation, iterative refinement, and future cross-pilot analysis in complex sustainability demonstrations.

Keywords
monitoring framework, sustainability demonstrations, behavioural change, conversion types, stakeholder engagement, food systems, pilot action plans, Horizon Europe



This article is included in the Horizon Europe gateway.

Corresponding author: Ilias Karachalios (ilias.karachalios@hotmail.gr)

Author roles: Karachalios I: Conceptualization, Data Curation, Formal Analysis, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Tantaroudas ND**: Data Curation, Formal Analysis, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Makropoulos C**: Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was developed in the context of the CHOICE project, funded by the European Union under Horizon Europe Grant Agreement No. 101081617. The funders had no role in the design of the article, interpretation of the framework, decision to publish, or preparation of the manuscript.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Karachalios I *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Karachalios I, Tantaroudas ND and Makropoulos C. **Monitoring effectiveness in sustainability demonstrations: a reproducible framework integrating stakeholder engagement and continuous feedback [version 1; peer review: 1 approved with reservations]** Open Research Europe 2026, **6**:219 https://doi.org/10.12688/openreseurope.22944.1

First published: 30 Jun 2026, **6**:219 https://doi.org/10.12688/openreseurope.22944.1

1. Introduction

Sustainability-transition projects increasingly rely on real-world demonstrations to test whether digital tools, communication strategies, and stakeholder engagement can influence practices related to food systems, consumption, and environmental decision-making. Yet demonstration projects face a recurring methodological problem: local interventions must be tailored to context, while evaluation still needs to generate comparable evidence across pilots, countries, and stakeholder groups. This tension is well recognised in sustainability governance, implementation research, and process evaluation, where context sensitivity and cross-case comparability often pull in opposite directions (Cash et al., 2003; Damschroder et al., 2009; Moore et al., 2015; Proctor et al., 2011).

This challenge is especially visible in projects that combine behaviour-change ambitions with digital or hybrid interventions. Evidence from technology acceptance, usability evaluation, communication campaigns, and environmental education shows that shifts in awareness, attitudes, self-efficacy, and self-reported behaviour are shaped not only by message content, but also by perceived usefulness, ease of use, credibility, and opportunities for repeated engagement (Brooke, 1996; Davis, 1989; Harwood & Garry, 2015; Noar, 2006). Related educational studies have likewise shown that sustainability-oriented learning interventions can influence environmental awareness and behavioural dispositions, while still requiring careful attention to context and measurement design (Karachalios & Manesis, 2025; Karachalios & Tantaroudas, 2025).

The CHOICE project – Mainstreaming Integrated Assessment Models by embedding behavioural change and digital narratives in food systems – provides a suitable setting in which to address this problem. Within CHOICE, the initial Pilot Action Plans deliverable (D6.1) established a common structure for pilot objectives, evaluation indicators, conversion types, stakeholder engagement, implementation roadmaps, and risk management across five pilot demonstrations in Austria, Spain, Greece, Colombia, and South Africa. The present Method Article builds on that foundation and reformulates it as a reproducible monitoring framework for scientific use, while also drawing on digital narratives, immersive communication formats, and Integrated Assessment Models (IAMs) to support sustainability engagement (Karachalios, Tantaroudas, & Makropoulos, 2025; Kozicka et al., 2025; Steinhauser, Fritz, & Pirker, 2024; Sundiang et al., 2025).

1.1 Aim and scope

The aim of this Method Article is to document a reproducible monitoring framework developed in the context of CHOICE for evaluating effectiveness across heterogeneous sustainability demonstrations through a common monitoring architecture, a shared conversion-types pathway, structured Participatory Labs, and a cross-pilot risk and reporting logic. The article does not report pilot outcomes. Instead, it describes the framework that guides data collection, interpretation, iterative refinement, and future pilot-level analysis.

2. Methods

2.1 Overall architecture

The CHOICE monitoring framework is organised as a three-layer architecture (Figure 1). First, five evaluation dimensions define what counts as effectiveness across pilots. Second, a conversion-types pathway translates behavioural progression into a comparable sequence of stages. Third, a mixed-methods evidence pipeline integrates analytics, questionnaires, Participatory Labs, and implementation reporting into an iterative learning process. This layered design reflects broader thinking in implementation science, developmental evaluation, and adaptive sustainability governance, where method robustness depends on combining standardised structures with contextual flexibility (Damschroder et al., 2009; Moore et al., 2015; Patton, 2011).

This structure was derived through analytical synthesis of the initial CHOICE Pilot Action Plans (D6.1) and was refined to balance standardisation with contextual fit. The derivation process drew on the cross-pilot objectives, evaluation indicators, conversion types, implementation roadmaps, stakeholder engagement structures, and risk-mitigation procedures specified across the five pilot settings. These elements were consolidated into a common monitoring architecture that supports reproducible interpretation while allowing pilot-specific indicators and channels to remain context-sensitive. The framework is deliberately not based on a single universal metric. Instead, it uses a stable interpretive lens across pilots while allowing pilot-specific indicators and channels to vary. In this respect, the framework is closer to a transferable evaluation architecture than to a fixed scorecard. Its purpose is not to impose identical indicators across pilots, but to provide a stable interpretive structure through which heterogeneous evidence can be collected, compared, and refined over time, in line with context-sensitive evaluation approaches (Proctor et al., 2011; Wiek, Talwar, O'Shea, & Robinson, 2014).

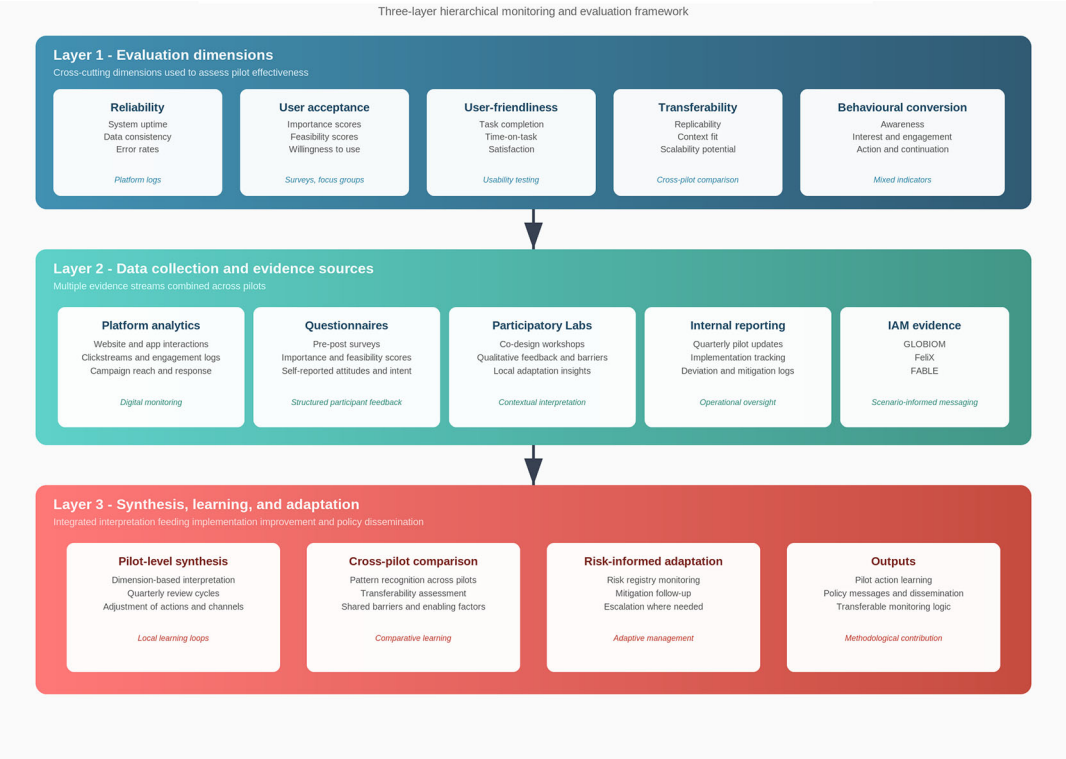


Figure 1. CHOICE monitoring framework architecture. Three-layer hierarchical monitoring and evaluation framework showing the relationship between the five evaluation dimensions, the mixed-methods evidence sources used across pilots, and the synthesis-and-adaptation layer through which monitoring outputs inform pilot learning, comparison, and policy-oriented interpretation.

2.2 Evaluation dimensions

The first layer of the framework consists of five cross-cutting evaluation dimensions: reliability, user acceptance, user-friendliness, transferability, and behavioural conversion. Reliability covers platform stability, data consistency, and error management. User acceptance captures perceived usefulness, relevance, and feasibility from the perspective of intended users. User-friendliness addresses practical experience of interaction, including accessibility and ease of navigation. Transferability considers whether lessons and design choices can inform use in other settings. Behavioural conversion examines whether exposure and engagement appear to translate into more substantive forms of action or sustained practice. The choice of these dimensions draws on work concerning acceptance, usability, implementation outcomes, and the practical assessment of complex interventions (Brooke, 1996; Davis, 1989; Proctor et al., 2011).

The five-dimension structure was chosen because projects of this kind often fail when either technical performance, user experience, or behavioural interpretation is treated as secondary. In CHOICE, these elements are monitored together, allowing implementation teams to identify whether weak outcomes stem from platform issues, message design, recruitment limitations, contextual misfit, or genuinely limited behavioural uptake. The operationalisation of these dimensions across indicators, data collection methods, measurement frequency, and responsible partners is summarised in Table 2.

2.3 Conversion-types pathway

To compare pilots without forcing identical behavioural endpoints, CHOICE uses a conversion-types pathway that tracks movement through five broad stages: awareness, interest, engagement, action, and continuation. This logic draws conceptually on diffusion and communication perspectives that treat change as staged and cumulative rather than instantaneous, and on behaviour-change approaches that distinguish between exposure, motivation, enactment, and maintenance (Michie, Van Stralen, & West, 2011; Noar, 2006; Rogers, Singhal, & Quinlan, 2014).

Table 1. Cross-pilot comparison matrix. Summary of the five CHOICE pilot demonstrations across lead partners, target groups, sector focus, geographic scope, principal digital channels, engagement structure, core digital tools, model linkages, timelines, and key conversion focus.

Dimension	Austria	Spain	Greece	Colombia	South Africa
Lead partner	Inoqo	CAAND	e-Fresh	SUP	UP
Target audience	DM retail consumers in Germany	Olive farmers and livestock farmers in Andalucía	e-Fresh online retail customers in Attica, Greece (200,000 registered users; focus on 50,000 active users)	Women coffee producers in Cauca, Colombia (AMUCC)	Youth networks and related stakeholders in South Africa
Sector focus	Sustainable food retail and eco-labelling	Sustainable olive cultivation and livestock farming	Sustainable online grocery shopping and eco-labelling	Sustainable coffee production sustainability	Healthy and sustainable diets and food waste reduction
Geographic scope	Germany	Andalucía, Spain	Attica, Greece	Cauca, Colombia	South Africa
Primary digital channels	Online shop, app, newsletter	WhatsApp broadcast	e-Fresh website, email, Viber	WhatsApp, Facebook/ Instagram, YouTube, project web platform	WhatsApp, Viber, email, social media, project website
Number of engagement campaigns	5	6 (3 olive +3 livestock)	5	5	5
Number of Participatory Labs	3	3	3	3	3
Key digital tool(s)	Inoqo eco-labelling platform	FABLE Calculator and GLOBIOM-informed campaign content	e-Fresh online retail platform with Inoqo eco-labelling integration	FABLE Calculator and GLOBIOM-informed campaign content	Shrink Your Food Waste app
IAM models used	GLOBIOM, FeliX	GLOBIOM, FABLE	GLOBIOM, FeliX, FABLE	GLOBIOM, FABLE	GLOBIOM, FABLE
Pilot kick-off month	M22	M21	M22	M22	M22
Data reporting frequency	Quarterly	Quarterly	Quarterly (M25, M28, M31, M34)	Quarterly (M25, M28, M31, M34)	Quarterly (M25, M28, M31, M34)
Policy dissemination period	M32-M36	M32-M36	M32-M36	M32-M36	M32-M36
Key conversion focus	Eco-label awareness, sustainable product selection, repeated sustainable purchasing	Adoption of sustainable farming practices in olive and livestock systems	Eco-label interaction, sustainable product selection, repeated sustainable purchasing	Adoption of sustainable coffee practices, participation in training, peer diffusion	Diet change, food waste reduction, leadership and ambassadorship

Table 2. Evaluation Dimensions Mapping. Operational definitions, indicative metrics, and primary evidence sources associated with the five cross-cutting evaluation dimensions used in the CHOICE monitoring framework.

Evaluation dimension	Operational indicators	Data collection method	Measurement frequency	Applicable pilots	Responsible partner(s)
Reliability	System uptime and availability	Platform monitoring logs	Continuous	AT, GR, SA	ICCS, pilot leaders, tool providers
Reliability	Data consistency across tools	Automated data validation and data-warehouse checks	Per data submission	All	ICCS, LIBRA
Reliability	Error rates in digital tools	Bug tracking systems and support logs	Continuous	AT, GR, SA	ICCS, pilot leaders
User Acceptance	Perceived importance scores	Pre/post surveys (Likert scale)	Pre and post engagement	All	ATHENA, LIBRA, pilot leaders
User Acceptance	Feasibility assessment	Stakeholder questionnaires and Participatory Labs feedback	Post Participatory Labs	All	ATHENA, pilot leaders
User Acceptance	Willingness to continue use	Follow-up surveys	At 3-month intervals	All	LIBRA, pilot leaders
User-Friendliness	Task completion rates	Usability tests (moderated)	During pilot deployment	AT, GR, SA	ICCS, LIBRA, pilot leaders
User-Friendliness	Time-on-task metrics	Digital analytics (session data)	Continuous	AT, GR, SA	Tool providers, ICCS
User-Friendliness	User satisfaction scores (SUS/NPS or equivalent)	Post-interaction surveys	Quarterly	All	LIBRA, pilot leaders
Transferability	Replicability assessment scores	Expert panel evaluation	End of pilot	All	ICCS, NOA
Transferability	Adaptation effort estimates	Cross-pilot workshops	Bi-annual	All	ICCS, ATHENA
Transferability	Documentation completeness	Internal review of protocols and monitoring templates	At deliverable milestones	All	ICCS
Overall Effectiveness	Behavioural change rates	Conversion type tracking	Quarterly	All	Monitoring team (ICCS lead)
Overall Effectiveness	Carbon footprint reduction	LCA-based calculations and model-informed estimation	Quarterly	AT, GR	Inoqo, ICCS
Overall Effectiveness	Practice adoption rates	Field surveys and self-reports	Quarterly	ES, CO	CAAND, SUP
Overall Effectiveness	Food waste reduction	App-tracked data and surveys	Quarterly	SA	UP, ICCS
Overall Effectiveness	Policy uptake indicators	Policy review and stakeholder interviews	M32-M36	All	NOA, ICCS

Open Research Europe 2026, 6:219 Last updated: 19 OCT 2026

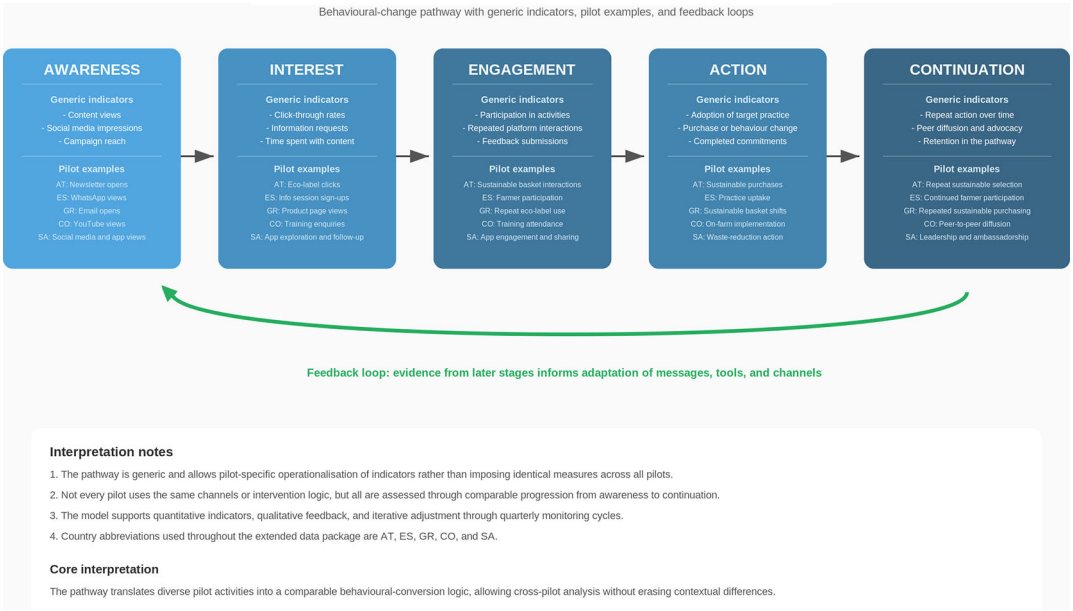


Figure 2. Conversion types pathway model. Generic behavioural-conversion pathway used to interpret pilot progression from awareness to continuation, combining comparable stages with pilot-specific examples and an explicit feedback loop from later-stage evidence to adaptation of messages, tools, and channels.

The purpose of the pathway is therefore interpretive rather than rigidly prescriptive. It creates a common language for describing how participants move from initial exposure to more substantive forms of adoption or leadership, while still permitting pilot-specific indicators and digital channels. Figure 2 presents this pathway together with generic indicators and illustrative pilot examples.

2.4 Mixed-methods evidence pipeline

The framework combines multiple evidence streams to support triangulation across heterogeneous pilot settings. Digital analytics provide interaction-level information such as views, clicks, repeat use, and selected platform actions. Structured questionnaires capture perceived importance, feasibility, user experience, and self-reported behavioural intentions or changes. Participatory Labs generate qualitative insights on barriers, enablers, contextual relevance, and adaptation needs. Internal reporting contributes implementation updates, emerging risks, and decisions taken during pilot delivery, so that monitoring is linked not only to outcomes but also to the conditions under which those outcomes are produced. Together, these evidence streams create a mixed-methods monitoring pipeline that supports interpretation, comparison, and iterative refinement across pilots (Creswell & Plano Clark, 2018; Fetters, Curry, & Creswell, 2013).

The emphasis on triangulation is important. Platform analytics may show whether a user clicked or returned, but they do not by themselves explain why a user disengaged or why one channel worked better than another. Likewise, questionnaires may capture intention or satisfaction, but without technical monitoring they may miss implementation problems that distort user experience. Bringing these sources together improves explanatory depth and supports more defensible interpretation of pilot performance in real-world settings (Fetters et al., 2013; Moore et al., 2015).

2.5 Participatory Labs and stakeholder feedback

Participatory Labs are integrated into the framework as a methodological mechanism for interpreting pilot performance, identifying barriers and enablers, and informing iterative adaptation during implementation. They are not treated as parallel dissemination activities, but as structured points of evidence generation that complement digital analytics, questionnaires, and internal reporting. Across pilots, three Participatory Labs are used to identify barriers, validate and prioritise intervention directions, and interpret pilot progress. This design is consistent with research on social learning, living labs, and participatory sustainability assessment, which shows that stakeholder involvement can improve contextual fit, legitimacy, and the practical relevance of monitoring systems (Reed et al., 2010; Voytenko, McCormick, Evans, & Schliwa, 2016; Wiek et al., 2014). The design, timing, participants, and expected outputs of this process are summarised in Table 4.

Table 3. Cross-pilot risk registry. Structured overview of major methodological, technical, operational, and engagement-related risks relevant to pilot monitoring, together with indicative mitigation logic.

Risk ID	Risk description	Probability	Impact	Applicable pilots	Mitigation measures	Lead partner	Supporting partners
R1	Data management or storage fails to align with GDPR or national data-protection regulation	Low	High	All	T1.5 compliance monitoring; privacy-preserving analytics; DMP templates and forms	ICCS	All pilot leaders, LIBRA
R2	Users remain unclear or sceptical regarding data usage and decision tracking	Medium	High	All	Clear privacy policies; GDPR-compliant tools; pseudonymisation before storage and analysis	ICCS	Inoqo, e-Fresh, SUP, UP, CAAND
R3	Language barriers affect stakeholder participation	Low	Medium	ES, CO, SA	Multilingual partner expertise; translation resources for materials	Pilot leaders	ICCS, ATHENA
R4	A/B testing data are insufficient for statistical significance	Medium	Medium	All	Pre-trial feasibility studies; Bayesian A/B testing; within-subject fallback designs	LIBRA	ICCS, pilot leaders
R5	Stakeholder engagement in campaigns and digital tools remains below target	Medium	High	All	Targeted marketing; community-level dissemination; regional ambassadors; WP7 coordination	Pilot leaders	ATHENA, ICCS
R6	Monitoring period is too short for measurable impact demonstration	Low	Medium	All	Preliminary evaluation; potential extension; intention-based supplementary indicators	ICCS	Pilot leaders
R7	Technical issues affect digital tool deployment	Medium	Medium	AT, GR, SA	Regular technical reviews; testing phases; dedicated support team	ICCS	Inoqo, e-Fresh, UP, LIBRA
R8	Insufficient training or capacity building for stakeholders	Medium	High	All	Training materials; workshops and webinars; ongoing support	Pilot leaders	ICCS, ATHENA
R9	Data quality concerns arise from user inputs	Medium	Medium	All	Robust validation and cleaning protocols; data-entry guidelines; user training	ICCS	LIBRA, pilot leaders
R10	Resistance to change among stakeholders limits adoption	Medium	Medium	ES, CO	Early stakeholder engagement; clear benefit communication; participation incentives	Pilot leaders	ATHENA, LIBRA
R11	Low user participation in consumer-facing pilots	Low	High	AT, GR	Leverage existing retailer customer bases; multi-channel recruitment; incentives	Inoqo, e-Fresh	ICCS
R12	External disruption affects in-person activities and field implementation	Low	Medium	All	Digital-first participation options; virtual knowledge sharing; reduced dependence on physical events	Pilot leaders	ICCS
R13	Dissemination activities do not reach target audiences effectively	Low	Medium	All	Measure dissemination success regularly; adjust strategy and channels as needed	Pilot leaders	ICCS, LIBRA

Open Research Europe 2026, 6:219 Last updated: 19 OCT 2026

Table 3. *Continued*

Risk ID	Risk description	Probability	Impact	Applicable pilots	Mitigation measures	Lead partner	Supporting partners
R14	IAM projections do not align sufficiently with real-world pilot scenarios	Medium	Medium	All	Close modelling-team collaboration; iterative cross-validation of results	IIASA, SDSN	Pilot leaders, ICCS
R15	Insufficient farmer demographic diversity in participation	Low	Medium	ES	Targeted outreach across demographics; tailored engagement strategies	CAAND	Bioazul, ATHENA
R16	Political instability affects field activities	Low	Medium	CO	Flexible scheduling; digital alternatives; community-based resilience networks	SUP	Tecni, ICCS
R17	Food insecurity dynamics affect participant engagement	Low	Medium	SA	Context-sensitive messaging; partnerships with food-security organisations	UP	ICCS

Table 4. Participatory labs design protocol. Standardised protocol describing the objectives, timing, expected participants, and outputs of the three Participatory Labs used to refine and interpret pilot monitoring.

Lab phase	Timing (project months)	Objective	Activities	Participants	Expected outputs	Coordinator
Lab 1: Exploration	M10-M15 (varies by pilot)	Understand the landscape of influencing factors, barriers, and opportunities	Mental mapping exercises; subgroup discussions on consumer/farmer behaviour drivers; traffic-light environmental impact assessments; stakeholder role mapping	ATHENA, LIBRA, pilot leaders, local stakeholders (consumers, farmers, youth, policymakers), ICCS	Identified key factors shaping food habits and practices; initial barrier and opportunity inventories; preliminary KPIs; first set of potential conversion types	ATHENA
Lab 2: Validation and Prioritisation	M13-M19 (varies by pilot)	Validate findings from Lab 1 and prioritise conversion types and sustainable practices	Detailed matrix review of conversion types; ranking by importance, feasibility, and profitability; preference testing for sustainability attributes; channel identification	ATHENA, LIBRA, pilot leaders, local stakeholders, ICCS	Validated and ranked conversion types; identified preferred digital channels; refined campaign messaging directions; profitability and adoption feasibility data	ATHENA
Lab 3: Finalisation	M15-M20 (varies by pilot)	Finalise at least 7 measurable conversion types per pilot and inform the Pilot Action Plan	Conversion type operationalisation; indicator definition; digital channel confirmation; campaign message co-design; cross-cutting strategy identification (e.g. training needs)	ATHENA, LIBRA, pilot leaders, local stakeholders, ICCS	Finalised set of at least 7 conversion types with measurable indicators; confirmed digital channels; co-designed campaign messages; informed PAP sections	ATHENA
Cross-pilot synthesis	M18-M20	Harmonise conversion types across pilots for cross-pilot comparability	Mapping pilot-specific conversion types to the common taxonomy; identifying comparable and unique indicators; establishing baseline expectations	ICCS, ATHENA, LIBRA, all pilot leaders	Harmonised conversion-types matrix; cross-pilot comparison framework; baseline targets for monitoring	ICCS

Open Research Europe 2026, 6:219 Last updated: 19 OCT 2026

In CHOICE, feedback loops continue beyond the labs themselves. Quarterly reviews, implementation meetings, and pilot-level discussions enable monitoring outputs to inform adaptation of channels, messaging, and support actions during implementation rather than only after pilot closure. This is one of the framework's central design choices.

In this way, stakeholder feedback is incorporated into the framework not only as contextual information, but also as methodological input for refining implementation choices and interpreting pilot effectiveness over time.

2.6 Agile refinement, experimentation, and risk management

The framework is embedded in an iterative workflow influenced by agile and continuous-improvement approaches, where monitoring findings are used to adjust implementation steps over time (Anderson, 2010; Beck et al., 2001). When the digital infrastructure and sampling conditions allow it, pilots may also incorporate controlled experimentation, including A/B comparisons of communication variants or engagement strategies. In practice, the framework is intentionally pragmatic: if sample sizes or technical conditions do not support formal experimentation, the same logic can still be used through structured comparison, staged roll-out, and reflexive review. This flexibility is consistent with developmental evaluation approaches designed for innovation under real-world conditions (Patton, 2011).

In this framework, risk management is treated as a methodological component of monitoring rather than as a separate administrative task, because recruitment barriers, technical disruptions, privacy constraints, and low engagement can directly affect how pilot effectiveness is interpreted. Cross-cutting and pilot-specific risks are tracked through a structured registry covering issues such as GDPR compliance, recruitment limitations, technical instability, low engagement, language barriers, and insufficient statistical power. The complete cross-pilot registry is summarised in Table 3.

2.7 Integrated Assessment Models within the framework

The framework also interfaces with the project's Integrated Assessment Models (IAMs), including GLOBIOM, FeliX, and FABLE. In methodological terms, these models do not simply sit upstream or downstream of monitoring. Rather, they provide scenario-informed content, narratives, and context that can shape campaigns, stakeholder discussions, and interpretation of pilot actions. This matters because the demonstrations are not only testing user interaction with digital tools; they are also testing whether model-informed narratives can support more meaningful engagement with sustainable food-system choices. The broader CHOICE research ecosystem already uses these assets across food-system communication, digital engagement, and modelling-related outreach (Karachalios, 2024; Kozicka et al., 2025; Mosnier et al., 2023; Steinhauser et al., 2024; Sundiang et al., 2025; Ye et al., 2024).

2.8 Ethical considerations and data protection

Because the CHOICE pilots involve digital tools, behavioural monitoring, and stakeholder engagement across multiple countries, ethics and data protection are integral to the method. The framework requires clear privacy information, appropriate consent procedures where relevant, pseudonymisation prior to storage and analysis whenever participant-level data are involved, and role-based access to monitoring files. The present article is methods-focused and does not report participant-level results. Accordingly, the article includes only framework-level material and does not report participant-level datasets or personal data.

3. Pilot-level operationalisation of the framework

The monitoring architecture is common across CHOICE, but each pilot operationalises it differently according to its sector, target group, digital channels, and behavioural objectives. This section illustrates how the same evaluation logic is adapted across the five pilot settings while preserving a common interpretive structure for cross-pilot comparison. Table 1 summarises the five pilot demonstrations and their main channels, target groups, timelines, and conversion focus.

3.1 Austria

The Austrian pilot focuses on online retail and eco-labelling, using the Inoqo platform to support sustainable product selection and repeated lower-impact purchasing. Monitoring therefore emphasises reliability of the digital environment, interaction with eco-label information, and continuation through repeat behaviour.

3.2 Spain

The Spanish pilot addresses olive and livestock farming in Andalucía. Here, conversion logic centres less on digital retail interaction and more on attention to campaign content, participation in information activities, and uptake of sustainable practices in farming contexts.

3.3 Greece
The Greek pilot is built around the e-Fresh online grocery environment with integration of eco-labelling functionality. Monitoring focuses on sustainable product selection, repeated interaction with environmental information, and the extent to which platform-supported choices can indicate meaningful behavioural conversion.

3.4 Colombia
The Colombian pilot works with women coffee producers in Cauca. The framework therefore monitors not only exposure and participation but also practice uptake, peer diffusion, and locally grounded barriers that shape whether engagement translates into sustained implementation.

3.5 South Africa
The South African pilot engages youth networks and related stakeholders through a food-waste reduction logic supported by mobile and social channels. In this setting, monitoring is oriented toward awareness, app engagement, practical waste-reduction action, and leadership or ambassadorship dynamics.

4. Discussion
4.1 Methodological contribution
The main contribution of this article is methodological. Rather than proposing a universal indicator set, the CHOICE framework offers a reproducible architecture for comparing heterogeneous demonstrations through a stable interpretive structure. This matters because sustainability projects often need evaluation systems that are rigorous enough to support cross-case learning but flexible enough to remain meaningful in highly different local settings. The framework therefore contributes to a growing body of work seeking evaluative approaches for complex, context-dependent interventions rather than narrowly standardised trials (Moore et al., 2015; Patton, 2011; Proctor et al., 2011).

A second contribution is the explicit treatment of stakeholder feedback as part of the method rather than as a separate dissemination or engagement strand. In many projects, stakeholders are consulted, but the outputs of those consultations are not structurally connected to monitoring decisions. In CHOICE, Participatory Labs, quarterly reviews, and risk tracking are linked directly to adaptation and interpretation. This is also where the framework differs from a static monitoring protocol: it treats learning and adjustment as core methodological functions. That orientation is compatible with scholarship on social learning, boundary knowledge systems, and living-lab experimentation for sustainability transitions (Cash et al., 2003; Reed et al., 2010; Voytenko et al., 2016).

4.2 Strengths and limitations
The framework has several strengths. It is transparent, modular, and compatible with mixed methods. It is also designed to accommodate digital and non-digital channels without assuming that all pilots generate the same kind of evidence. At the same time, several limitations should be acknowledged. First, this article presents the initial version of the framework and not its full validation through completed pilot results. Second, behavioural conversion remains partly dependent on self-report and context-specific indicators, especially in pilots where direct behavioural observation is limited. Third, comparability is interpretive rather than purely statistical: the framework supports structured cross-pilot synthesis, but it does not eliminate contextual differences. These limitations should therefore be understood as conditions of application in complex real-world demonstrations rather than as weaknesses that undermine the methodological value of the framework.

These limitations are not flaws unique to CHOICE; they are characteristic of complex multi-country demonstrations. The methodological aim is therefore not to erase heterogeneity but to make it analyzable.

4.3 Transferability
The framework is potentially transferable to other sustainability projects that combine digital tools, stakeholder engagement, and context-specific implementation. Its most transferable elements are the layered architecture, the use of staged conversion logic, the role of structured co-design in defining indicators, and the integration of risk management into monitoring itself. Transferability will still depend on local infrastructure, governance conditions, and data protection requirements, but the general logic should be useful for other multi-actor projects in food systems, environmental governance, digital public engagement, and related domains (Cash et al., 2003; Karachalios et al., 2025; Wiek et al., 2014).

5. Conclusions
This Method Article presents the CHOICE monitoring framework as a reproducible approach for evaluating heterogeneous sustainability demonstrations. The framework combines five evaluation dimensions, a staged conversion pathway,

a mixed-methods evidence pipeline, Participatory Labs, iterative feedback loops, and integrated risk management within a common monitoring architecture. Taken together, these elements provide a structured way to interpret effectiveness across diverse pilots while preserving contextual specificity. The article does not report pilot outcomes; these will be addressed in subsequent implementation-focused publications. Its main contribution is therefore methodological: it provides a transparent and transferable blueprint for data collection, interpretation, comparison, and iterative refinement in complex real-world sustainability demonstrations.

Ethics and consent

Ethical approval and consent were not required for this Method Article. The article describes a methodological framework and associated extended materials, and does not report participant-level data, identifiable personal information, or primary research findings from human participants.

AI-use disclosure (F1000 policy)

During manuscript preparation, generative AI tools OpenAI ChatGPT (GPT-5 Thinking; model snapshot - March 2026) and Anthropic Claude (Claude 4.6 – March 2026) were used to assist with language editing for clarity and concision, text structuring, and drafting support. All outputs were critically reviewed, verified, and edited by the authors, who take full responsibility for the content. No data analysis, coding, or generation of scientific results was performed by these tools.

Data availability

No underlying data are associated with this Method Article. The article presents a methodological framework and does not report participant-level results, raw measurements, outcome data, or other datasets requiring deposition in a public repository. All figures and tables necessary to understand the framework are included within the manuscript.

Acknowledgements

The authors acknowledge the CHOICE consortium partners for their contributions to the development of the Pilot Action Plans and the broader methodological discussions that informed this framework.

References

Anderson DJ: *Kanban: Successful evolutionary change for your technology business.* Blue Hole Press; 2010.

Beck K, Beedle M, Bennekum A, *et al.*: **Manifesto for Agile Software Development.** 2001. Retrieved 20 March 2025.
Reference Source

Brooke J: **SUS: a quick and dirty usability scale.** Jordan PW, Thomas B, Weerdmeester BA, *et al.*, editors. *Usability Evaluation in Industry.* London: Taylor & Francis; 1996; pp. 189–194.

Cash DW, Clark WC, Alcock F, *et al.*: **Knowledge systems for sustainable development.** *Proc. Natl. Acad. Sci. USA.* 2003; **100**(14): 8086–8091.
Publisher Full Text

Creswell JW, Plano Clark VL: *Designing and conducting mixed methods research.* 3rd ed. Los Angeles: SAGE; 2018.

Damschroder LJ, Aron DC, Keith RE, *et al.*: **Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science.** *Implement. Sci.* 2009; **4**(1): 50.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Davis FD: **Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology.** *MIS Q.* 1989; **13**(3): 319–340.
Publisher Full Text

Fetters MD, Curry LA, Creswell JW: **Achieving Integration in Mixed Methods Designs—Principles and Practices.** *Health Serv. Res.* 2013; **48**(6pt2): 2134–2156.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Harwood T, Garry T: **An investigation into gamification as a customer engagement experience environment.** *J. Serv. Mark.* 2015; **29**(6/7): 533–546.
Publisher Full Text

Karachalios I: Compilation of Digital Tools on Food Green House Gas Mitigation (CHOICE Project) (Version 1.0.0). [Data set]. *Zenodo.* 2024.
Publisher Full Text

Karachalios I, Manesis N: **Fostering environmental awareness in primary school students: Evaluating the impact of a waste management education program.** *Eur. J. Educ. Stud.* 2025; **12**(4).
Publisher Full Text

Karachalios I, Tantaroudas N: **Future Greek Pre-Service Teachers' Knowledge, Attitudes and Self-Efficacy in Waste Management.** *Br. J. Educ.* 2025; **13**(8): 25–42.
Publisher Full Text

Karachalios I, Tantaroudas N, Makropoulos C: **Mapping immersive digital tools in the ICT4Water Cluster: Innovation, interoperability and impact.** *Open Res Eur.* 2025; **5**: 329.
Publisher Full Text

Kozicka M, Leclère D, Davis C, *et al.*: **Feeding The Planet While Saving It: How New Foods Can Help!.** *Front Young Minds.* 2025; **13**: 1558710.
Publisher Full Text

Michie S, Van Stralen MM, West R: **The behaviour change wheel: A new method for characterising and designing behaviour change interventions.** *Implement. Sci.* 2011; **6**(1): 42.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Moore GF, Audrey S, Barker M, *et al.*: **Process evaluation of complex interventions: Medical Research Council guidance.** *BMJ.* 2015; **350**(mar19 6): h1258–h1258.
Publisher Full Text

Mosnier A, Javalera-Rincon V, Jones SK, *et al.*: **A decentralized approach to model national and global food and land use systems.** *Environ. Res. Lett.* 2023; **18**(4): 045001.
Publisher Full Text

Noar SM: **A 10-Year Retrospective of Research in Health Mass Media Campaigns: Where Do We Go From Here?** *J. Health Commun.* 2006; **11**(1): 21–42.
PubMed Abstract | **Publisher Full Text**

Patton M: **Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use.** 2011.

Proctor E, Silmere H, Raghavan R, *et al.*: **Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda.** *Admin. Pol. Ment. Health.* 2011; **38**(2): 65–76.
Publisher Full Text

Reed MS, Evely AC, Cundill G, *et al.*: **What is Social Learning?** *Ecol. Soc.* 2010; **15**(4): resp1.
Publisher Full Text

Rogers EM, Singhal A, Quinlan MM: **Diffusion of innovations.** *An integrated approach to communication theory and research.* Routledge; 2014; pp. 432–448.

Steinhauser J, Fritz S, Pirker J: **Virtual Climate Activation: A Framework Approach to Sustainable Food Consumption in Games.** *2024 IEEE Conference on Games (CoG).* Milan, Italy: IEEE; 2024; pp. 1–4.
Publisher Full Text

Sundiang M, Diniz Oliveira T, Mason-D'Croz D, *et al.*: **Bundling measures for food systems transformation: A global, multimodel assessment.** *Lancet Planet Health.* 2025; **9**(10): 101339.
Publisher Full Text

Voytenko Y, McCormick K, Evans J, *et al.*: **Urban living labs for sustainability and low carbon cities in Europe: Towards a research agenda.** *J. Clean. Prod.* 2016; **123**: 45–54.
Publisher Full Text

Wiek A, Talwar S, O'Shea M, *et al.*: **Toward a methodological scheme for capturing societal effects of participatory sustainability research.** *Res. Eval.* 2014; **23**(2): 117–132.
Publisher Full Text

Ye Q, Liu Q, Swamy D, *et al.*: **FeliX 2.0: An integrated model of climate, economy, environment, and society interactions.** *Environ. Model Softw.* 2024; **179**: 106121.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ?

Version 1

Reviewer Report 08 September 2026

https://doi.org/10.21956/openreseurope.24812.r79687

© **2026 Appiah-Kubi E et al.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Elias Appiah-Kubi
Universitatea din Petrosani, Petroșani, Hunedoara County, Romania
Codruta Cornelia Dura
Economics, University of Petrosani, Petrosani, Hunedoara, Romania

The manuscript, *"Monitoring effectiveness in sustainability demonstrations: a reproducible framework integrating stakeholder engagement and continuous feedback,"* presents a monitoring framework developed within the CHOICE project across five sustainability demonstrations in Austria, Spain, Greece, Colombia, and South Africa. The framework combines five evaluation dimensions, a behavioural-conversion pathway, mixed-method evidence, stakeholder participation, continuous feedback, and risk management.

The manuscript addresses a relevant methodological challenge concerning how heterogeneous sustainability demonstrations can maintain contextual flexibility while still allowing some degree of cross-pilot comparison. The framework is clearly presented, and the integration of stakeholder feedback into the monitoring process is particularly useful. However, several areas require clarification before indexing.

Abstract
The abstract is clear and appropriately indicates that the manuscript presents a methodological framework rather than pilot results. However, the methodological contribution could be stated more explicitly. The author(s) should briefly clarify what distinguishes the proposed framework from existing sustainability monitoring and implementation approaches.

Introduction
The introduction provides a reasonable justification for the study, particularly the tension between local adaptation and cross-pilot comparability. However, the methodological gap could be stated more directly. The author(s) should briefly explain what is currently missing from existing monitoring approaches and how the CHOICE framework specifically addresses this limitation.

Development of the Framework

The manuscript states that the framework was developed through an analytical synthesis of the CHOICE Pilot Action Plans. This process requires slightly more explanation. In particular, the author(s) should clarify how the five evaluation dimensions and conversion stages were selected and whether they emerged from the project documents, previous literature, discussions among project partners, or a combination of these sources.

Evaluation of Effectiveness
The five dimensions of reliability, user acceptance, user-friendliness, transferability, and behavioural conversion provide a useful structure. However, it is not entirely clear how these dimensions collectively determine whether a pilot is considered effective. The author(s) should clarify whether effectiveness is interpreted separately across the five dimensions or whether some overall assessment is eventually produced.

Cross-Pilot Comparison
One of the manuscript's main contributions is its ability to compare heterogeneous pilots without requiring identical indicators. This is useful, but the actual comparison procedure requires further explanation. Given the substantial differences between the five pilots shown in Table 1, the author(s) should briefly explain how pilot-specific indicators will be mapped onto the common evaluation dimensions and how cross-pilot conclusions will subsequently be reached.

Conversion-Types Pathway
The awareness–interest–engagement–action–continuation pathway is clearly presented, and Figure 2 helps communicate the process. The manuscript also appropriately clarifies that the pathway is interpretive rather than strictly prescriptive. Nevertheless, the author(s) should briefly acknowledge that behavioural change may not always follow a linear sequence and explain how movement backwards, discontinuation, or skipping stages would be interpreted.

Participatory Labs and Stakeholder Feedback
The integration of Participatory Labs into monitoring rather than treating them simply as dissemination activities is a useful aspect of the manuscript. The protocol presented in Table 4 is also informative. The author(s) could strengthen this section by briefly explaining how feedback generated through the Labs will be documented and incorporated into subsequent changes to pilot activities.

Reproducibility
The manuscript repeatedly describes the framework as reproducible. While the overall architecture is clearly described, some implementation details remain relatively general. The author(s) should therefore consider providing supplementary materials containing the main questionnaires, templates, indicator definitions, or reporting protocols used within the framework. Alternatively, the author(s) could moderate the claim of full reproducibility and describe the framework as a structured and transferable monitoring framework.

Discussion and Limitations
The discussion appropriately identifies the framework's principal methodological contribution and acknowledges that the framework has not yet been fully validated using completed pilot outcomes. The limitations section could briefly acknowledge that differences in local implementation, data availability, and stakeholder participation may affect cross-pilot comparability.

Is the rationale for developing the new method (or application) clearly explained
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
No

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: sustainability reporting; sustainability assurance; SMEs; circular economy; structural equation modelling; public sector; quantitative analysis

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.