



**Publicly
Available
Specification**

**Clean cookstoves and clean cooking
solutions — Guidance for evaluation**

Fourneaux et foyers de cuisson propres — Guide d'évaluation

**First edition
2024-08**

ISO/PAS 16617

STANDARDSISO.COM : Click to view the full PDF of ISO/PAS 16617:2024

STANDARDSISO.COM : Click to view the full PDF of ISO/PAS 16617:2024



COPYRIGHT PROTECTED DOCUMENT

© ISO 2024

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

	Page
Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Multiple aspects of cooking systems	3
5 System identification	4
5.1 System identification overview.....	4
5.2 System of action.....	5
5.3 System of desired impact.....	5
5.4 Implications of system choice.....	5
6 Determining fitness for purpose	6
6.1 Validity and relevance.....	6
6.2 Situational appropriateness or contextuality.....	6
6.3 Transparency and traceability.....	6
7 Quantifying variability	7
7.1 General.....	7
7.2 Reporting experimental error and uncertainty.....	7
7.3 Variability due to critical inputs.....	8
7.4 Combining sources of variability.....	8
8 Reporting single-outcome evaluation	8
9 Reporting multi-outcome evaluation	9
Bibliography	11

STANDARDSISO.COM : Click to view the full PDF of ISO/PAS 16617:2024

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 285, *Clean cookstoves and clean cooking solutions*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

The significance of standardization in the cookstove sector is well established. In addition to the documents published by ISO, there are many other “evaluation instruments” in existence throughout the world, covering different aspects of clean cookstoves and clean cooking solutions. These various evaluation instruments (standards, protocols, test methods, etc.) serve a multitude of purposes by examining a range of different metrics.

Determining the pros and cons of each evaluation instrument can be challenging. To address this gap, this document provides guidelines to assist in assessing and identifying which evaluation instrument is most appropriate to address the specific needs of a particular situation. This guidance is intended to better inform stakeholders in order to ensure that the end product is fit for purpose in a holistic sense.

STANDARDSISO.COM : Click to view the full PDF of ISO/PAS 16617:2024

STANDARDSISO.COM : Click to view the full PDF of ISO/PAS 16617:2024

Clean cookstoves and clean cooking solutions — Guidance for evaluation

1 Scope

This document provides guidance for examining evaluation instruments that assess cooking systems. The goal is to ensure that a cooking system evaluation instrument achieves its objectives and is fit for a particular purpose. Doing so means recognizing that the stove is one part of a series of complex and interacting systems, many of which are variable.

The document includes procedures to ensure that:

- an evaluation instrument captures the necessary input measures at an appropriate level of detail and quality;
- outputs from the evaluation instrument are accurate, repeatable, reproducible, and appropriate for assessing the stated outcome.

The document also provides guidance to interpret the outputs of an evaluation instrument when reporting. The document does not specify test methods nor recommend pre-determined procedures or benchmarks of performance (also known as tiers).

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/TR 21276, *Clean cookstoves and clean cooking solutions — Vocabulary*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/TR 21276 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1 cooking system

combination of *cookstove* (3.2), fuel, cooking equipment, cooking environment (including *ventilation* (3.16)) and cooking practice

[SOURCE: ISO/TR 21276:2018, 3.5.4]

3.2 cookstove

appliance primarily employed for the cooking of food, but which can also be employed for space or water heating, or other purposes

[SOURCE: ISO/TR 21276:2018, 3.1.7]

3.3

correlative relationship

relationship between a *proxy metric* (3.10) and the *desired outcome* (3.5), both of which are relevant to the *system of desired impact* (3.15)

3.4

critical input

aspect of a *system* (3.13) that affects the value of a measured *performance metric* (3.9), and that can vary throughout a *system of desired impact* (3.15)

3.5

desired outcome

outcome that the evaluator wishes to achieve in the *system of desired impact* (3.15)

3.6

evaluation instrument

specified methodology of measuring or collecting data using a documented procedure

Note 1 to entry: The term *evaluation instrument* is akin to “test method”, “protocol”, or “evaluation tool”.

Note 2 to entry: An *evaluation instrument* can consist of a series of different tests or data collection steps.

Note 3 to entry: An *evaluation instrument* can result in either quantitative or qualitative data, such as representing physical properties or social measures.

3.7

fitness for purpose

assessment of an *evaluation instrument* (3.6) ascertaining it is capable of sufficiently meeting the objectives for which it is to be applied

3.8

mechanistic relationship

predictable connection between the *performance metric* (3.9) (a measure in the *system of action* (3.14)) and the *proxy metric* (3.10) (a measure in the *system of desired impact* (3.15))

3.9

performance metric

measurable quantity characterizing the behaviour of the *system of action* (3.14) with regard to some aspect of a *cooking system* (3.1)

3.10

proxy metric

measurable quantity in the *system of desired impact* (3.15), which is related to the *desired outcome* (3.5)

3.11

reporting framework

approach to document and disseminate findings from an *evaluation instrument* (3.6)

3.12

safety

ability of a *cookstove* (3.2) to be operated at an acceptable level of risk of harm

[SOURCE: ISO/TR 21276:2018, 3.3.22]

3.13

system

collection of elements that can be grouped within a common boundary

3.14

system of action

system (3.13) that is altered by the design of technology, by behaviour change, or an intervention programme

Note 1 to entry: Refer to [subclause 5.2](#) for further information.

3.15

system of desired impact

system (3.13) in which an evaluator wishes to claim that a change has occurred

Note 1 to entry: Refer to [subclause 5.3](#) for further information.

3.16

ventilation

provision of air supply, by natural or mechanical means

Note 1 to entry: Ventilation can have a different interpretation depending on the context. For example, ventilating a flame to provide fresh air for combustion is distinct from ventilating a room to provide fresh air for its occupants.

4 Multiple aspects of cooking systems

A cooking system is a complex culmination of many components and factors. Cooking systems have various impacts on a range of different members of society and on the environment. Hence, there are many aspects of a cooking system that may be assessed using an evaluation instrument. Depending on its intended purpose, an evaluation instrument may assess part of one or more aspects of a cooking system. An evaluator needs to identify the aspects to be considered, some of which may include:

- a) Accessibility of fuel or energy source (available physically, reliably, and affordably).
- b) Accessibility of parts (manufacture, repair, as well as skilled labour).
- c) Adoption of stoves and measurement of adoption rates.
- d) Affordability.
- e) Climate impacts (e.g. carbon neutrality).
- f) Description of government policy around a specific fuel or appliance.
- g) Durability.
- h) Emissions (type and quantity).
- i) Environmental impacts, including air pollutant concentration.
- j) Fuel efficiency.
- k) Gender impacts.
- l) Health impacts (including exposure assessment).
- m) Livelihood and market impacts.
- n) Properties of fuel (physical and chemical).
- o) Reliability.
- p) Safety.
- q) Social impacts.
- r) Suitability for performing cooking tasks.
- s) Usability: ease of use and understandability of instructions.
- t) User acceptability: aesthetic and ergonomic aspects.
- u) User practices (such as choice of size and materials of cooking vessels).

How each of these aspects is defined for a particular evaluation instrument is contextual and based on the specific needs of that assessment.^[1] Therefore, all terms and definitions need to be clearly stated in any evaluation instrument.

It is likely that more than one evaluation instrument is needed to achieve a comprehensive evaluation of a cooking system and its impacts. [Clauses 5](#) and [6](#) outline some considerations for how individual aspects should be treated by an evaluation instrument and how they should be assessed. If an assessment of an evaluation instrument identifies that it fails to meet the intended purpose, then an alternative evaluation instrument should be sought.

5 System identification

5.1 System identification overview

The term “system” is used throughout this document as a reminder that every collection of elements, whether it be technological, social, or administrative, contains components that affect operation and outcomes. To identify appropriate evaluation mechanisms, the evaluator must identify two systems; the system of action and the system of desired impact before choosing a performance metric and evaluation procedure. [Table 1](#) lists a sequence of possible system choices from smallest to largest. This hierarchy of nested systems was originally presented in Bond et al.^[2] but has evolved since that publication, and system definitions can vary.

For each system in [Table 1](#), one or more performance metrics may apply, depending on the context and intended purpose of an evaluation. Refer to [Clause 4](#) for some examples of performance metrics.

Table 1 — Description of systems

System name	Description
Cookstove	Appliance primarily employed for the cooking of food, but which can also be employed for space or water heating, or other purposes. The cookstove itself may be considered a system of physical components that embody the energy conversion scheme, but by itself, a cookstove cannot perform a cooking function, therefore many aspects of its performance must be assessed within a larger system. The combination of the cookstove, fuel, cooking vessel, operating procedure, and operating environment are the minimum elements needed to define a system for evaluation in a controlled laboratory setting.
Cooking system	Combination of cookstove, fuel, cooking equipment, cooking environment (including ventilation) and cooking practice. This is the minimum system needed to complete a cooking task. Ventilation in this context refers to the supply of air for combustion and transfer of exhaust gases from the cookstove.
Kitchen environment	Cooking system as well as kitchen occupants (including cook), kitchen room and natural or mechanical ventilation of the working space, and stoves other than the one evaluated. Cooking can occur in an open environment, which challenges some conventional definitions of a kitchen.
Household	Kitchen; structure of entire house and position of kitchen within it; occupants of the house; energy service demands and building ventilation of the household; fuel storage and acquisition.
Community	Group of households with common spatial boundaries or demographic characteristics; infrastructure, institutions, local markets, and fuel supply; cultural norms and education; sources of air pollution such as fugitive dust, traffic or trash burning.

Table 1 (continued)

System name	Description
Nation or region	Politically defined geographic area. Contains many communities with overall governance structure; natural and built environments within those areas; policy instruments such as regulation, enforcement, and import and export duties; agreed standards, protocols and health targets; societal norms and aspirations.
Entire world	Includes all nations, as well as the global climate system; multinational companies including those that affect energy supply and humanitarian aid; internationally agreed standards, protocols and health targets; societal norms and aspirations that cross national boundaries.
NOTE: Table 1 lists and describes systems and bounded collections of elements in increasing order of size. One system is selected as the system of action and one as the system of desired impact before choosing a performance metric and evaluation procedure.	

5.2 System of action

The system of action is the largest system altered by the evaluator. For example, if the cookstove or fuel are replaced, as in a programme that provides liquefied petroleum gas, the system of action is the cooking system. The kitchen environment is considered the system of action in a programme that provides room ventilation. The choice of system of action is important for determining appropriate performance metrics and evaluation mechanisms.

5.3 System of desired impact

The system of desired impact is the system in which the evaluator wishes to claim that a change has occurred. All components of the system, not only those of interest to the evaluator, can affect the outcome. The system of desired impact may be the same as the system of action; however, evaluators often wish to claim a change in a larger system. Such an extrapolation is rarely appropriate. As the chosen system of desired impact increases in size, the greater complexity of the system also increases the variation in desired outcome so that the effect of the measure being evaluated can be imperceptible. This situation can be counteracted by choosing a smaller system of desired impact or a larger system of action.

5.4 Implications of system choice

The evaluator's choice of system of action and system of desired impact must be supported by appropriate evaluation measures. Evaluators are not expected to invent new evaluation measures, such as measurement protocols or survey instruments. Instead, evaluation measures are expected to evolve to suit the needs of rigour by explicitly addressing the concepts listed in this subclause, and evaluators may participate in processes to develop new evaluation measures.

Some questions that can determine whether the chosen systems are appropriate for the question of interest include:

- a) Do the evaluation measures produce one or more performance metrics that represent the aspect being evaluated in the system of action?
- b) Does the evaluation measure identify critical inputs, namely those aspects of the system that vary and that alter the measured performance metric?
- c) Does the evaluation measure require an assessment of how variation of critical inputs throughout the entire system of desired impact affects the performance metric?
- d) Do the reporting measures identify a desired outcome and proxy metric within the system of desired impact, and an explicit connection between the performance metric and the desired outcome? (For further guidance, see Reference [2].)

6 Determining fitness for purpose

6.1 Validity and relevance

Fitness for purpose includes considerations of validity and relevance. A method is valid when it is based upon well-founded and accurate claims or assumptions.^[3] A method is only relevant when it is applied in such a way that its validity is maintained for the context in which it is to be applied.

When assessing if an evaluation instrument is fit for purpose, it is necessary to first establish if the approach is fundamentally valid and relevant. Some example questions that may be asked when evaluating an evaluation instrument include the following:

- a) Is the approach based upon a generally accepted evaluation technique?
- b) Is there sufficient documentation to justify that the approach utilizes valid methodologies that are based upon robust scientific or technical knowledge?
- c) Is the method going to be applied in a manner consistent with the purpose of the original protocol?
- d) Are the questions or knowledge being sought by the individual applying the protocol consistent with the stated objectives of the method?

6.2 Situational appropriateness or contextuality

Assessing if an evaluation instrument is fit for purpose is predicated on establishing the situation or context in which that method is to be applied. Contextual considerations that impact the appropriateness of a method include those such as goals and priorities, minimum accuracy requirements, and limitations or requirements specific to the system in which the method is to be applied. A valid method is possibly not appropriate for all situations, and therefore, is not fit for purpose in all situations.

One example is the repeatability of a procedure. A procedure can be valid but give highly variable results, such as a survey asking primary cooks how long they are typically in a kitchen each week. Such a survey can be fit for purpose to establish a range of typical cooking times, but not be fit for purpose if the goal is to establish a highly repeatable measure of exposure time to cooking-based emissions.

Situational appropriateness is often also influenced by contextuality. An evaluation instrument can be appropriate for application in a highly controlled laboratory setting, but not when applied to unconstrained field usage.^[4] A method can also be appropriate for one geographical or cultural region, but not applicable to another. A procedure can be scientifically valid, but not situationally appropriate in all cases, and therefore not relevant in all cases.

An example of contextuality is the choice of chemical species to be measured in emission measurements. Measuring emissions of unburned alkanes can be appropriate where LPG stoves are widely used, but inappropriate when wood is the only fuel available.

An evaluation instrument that is fit for purpose is situationally and contextually appropriate. Some example questions that may be used to assess the situational appropriateness include the following:

- a) Can the method be applied in a way that is safe for all individuals involved?
- b) Is the goal of the assessment related to how a cooking system impacts an individual? If so, does the method account for the location, cultural or social situation in which evaluation is to occur?
- c) Are the measures expected to be relevant to the cooking system under evaluation?

6.3 Transparency and traceability

All methods, processes, and procedures have inherent limitations. Often these limitations are in the form of assumptions, idealizations, or methodological biases. A method can be valid despite these limitations, but assessing fitness for purpose necessitates a clear understanding and appreciation that they exist. For a

method to be fit for purpose, the limitations must be acceptable within the goals, requirements, and context for which the method is to be applied.

Limitations of an evaluation instrument are not always obvious to a casual or untrained observer. Transparency about these limitations at all stages is therefore necessary when evaluating and applying a procedure. Indeed, an evaluation instrument that explicitly states limitations can be indicative of one which has been carefully developed. An evaluation of whether a method is fit for purpose should occur not only when it is applied, but also when considering the results or outputs that come from that method. This can only occur through maintaining transparency.

Methodological traceability is the ability to track the process, assumptions, deviations from predefined plans, and decisions made while completing a procedure. Traceability provides confidence in the reported results and outcomes as it allows for the re-creation of the initial process. Traceability also provides the necessary information for an outside observer to objectively review the procedures taken and assess if the process and results are fit for purpose for their specific needs.

Some example questions that may be asked in order to assess if the implementation of a method is sufficiently transparent and traceable, may include the following:

- a) Are assumptions and limitations explicitly stated?
- b) Are all constants and coefficients in the calculations explicitly defined?
- c) Does the testing report provided with results from the method include a discussion of any deviations from the standard testing plan?
- d) Are the names, or other identifying information, of people responsible for the quality of the data included?
- e) Are the dates and location of data collection, and duration of data collection, included in the record?
- f) Is the number of data samples included in the record, and is there a record of any data that was discarded due to anomalies?

7 Quantifying variability

7.1 General

Evaluation instruments often rely upon numerous primary inputs and parameters. Although many parameters can be necessary to calculate final metrics of interest, often a few critical inputs have a disproportionate influence on the accuracy of those parameters. Evaluating a method for fitness for purpose necessitates an appreciation for the inputs that will have the largest impact on the precision, accuracy, or bias of the resulting outputs.

7.2 Reporting experimental error and uncertainty

All measurements are susceptible to error. The source and scale of these errors can vary substantially but can often be classified as either systematic or random error. Systematic error can arise from issues such as poor instrument calibration, instrument malfunction, or bias in the methodological procedures employed. Systematic error can be quantified by comparing measurands with calibration or reference values. Systematic errors can be difficult to identify but are completely avoidable if good procedures and calibration practices are employed.

Random error is a result of uncontrollable variability in factors such as instrument noise, environmental conditions, or unconscious observational variability. Random error results in different repeated measured values despite no changes in procedures or instrument settings. Random error is quantified as the standard deviation of the variability of the collected repeated measures.

A particular evaluation instrument may be repeatable, within the bounds of either systematic or random error, or both. However, these errors can be unique to the test environment. It can often be necessary for an

evaluation to be performed at multiple sites and involve different personnel and equipment, in which case the reproducibility of the measurements under these varying conditions also needs to be considered.

NOTE Repeatability is obtained via intralaboratory comparison, while reproducibility is obtained via interlaboratory comparison.

In general, the objective should be to minimize error and uncertainty. However, there is also a need to balance accuracy with efficacy. Improving the accuracy and precision of a measurement generally costs time and money but does not always offer significant benefits. The extent to which a measurement has been compromised to reduce time or financial costs should be justified.

Questions that can assist an evaluator in determining the appropriateness of uncertainty reporting are:

- a) Is all variability due to input parameters?
- b) Are the repeatability and reproducibility of data considered separately, are causes of variability identified, and is the associated variability quantified?
- c) Are calibration details of equipment reported, where applicable, and is accuracy, uncertainty and bias in this equipment considered?

7.3 Variability due to critical inputs

Critical inputs are factors that can vary throughout a system of desired impact, and for which variability affects the value of a measured performance metric. An example of variability in critical inputs is the type and condition of fuel used in a stove. Depending on the season, the fuel can be dry or damp, and can be of varying origin. An evaluation using two different sets of input parameters can yield different results, yet such variability is normal within the target community. Similarly, the size and design of dwellings or other relevant buildings is possibly highly variable within the target community. Another example is cooking vessel size, which can influence the performance of both emissions and thermal efficiency.

A question that assesses whether critical inputs are appropriately included is:

- Does the evaluation instrument identify critical inputs that affect the performance metric and quantify variability due to critical inputs that can be observed throughout the system of desired impact?

7.4 Combining sources of variability

The uncertainty of an evaluation instrument is the result of a combination of all sources of error and variability. The method of calculating the combined uncertainty depends on whether the uncertainty sources are dependent or independent. [5], [6] To adequately assess if a method is fit for purpose, an identification of the critical inputs, the associated uncertainties, and the resulting combined uncertainty must be determined. Once an overall uncertainty has been established, it becomes possible to determine if the uncertainty is acceptable within the context for which the method is to be applied.

A final question that should be assessed regarding the treatment of variability is:

- Does the evaluation instrument appropriately combine uncertainty in order to quantify variability in the performance metric, including variability due to critical inputs?

8 Reporting single-outcome evaluation

Through a reporting framework, outputs of an evaluation instrument need to be documented, communicated and interpreted for the appropriate end-user of that information. The evaluation instrument itself can possibly give guidance on how to report the results, or provide raw data that an evaluator needs to interpret and use at their discretion. The format of the report depends on the context. The report that describes the findings of an evaluation instrument is critically important, as incomplete or misleading reporting of findings can significantly affect the decisions taken based on the report.