Downloaded from orbit.dtu.dk on: Aug 26, 2022

DTU Library

Normalisation and weighting in life cycle assessment: quo vadis?

Pizzol, Massimo; Laurent, Alexis; Sala, Serenella; Weidema, Bo Pedersen; Verones, Francesca; Koffler, Christoph

Published in:
International Journal of Life Cycle Assessment

Link to article, DOI:
10.1007/s11367-016-1199-1

Publication date:
2017

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Pizzol, M., Laurent, A., Sala, S., Weidema, B. P., Verones, F., & Koffler, C. (2017). Normalisation and weighting in life cycle assessment: quo vadis? *International Journal of Life Cycle Assessment*, *22*(6), 853-866. https://doi.org/10.1007/s11367-016-1199-1

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Pizzol, M, A Laurent, S Sala, F Verones, B Weidema, and C Koffler. 2016. "Normalisation and Weighting in Life Cycle Assessment: Quo Vadis?" The International Journal of Life Cycle Assessment.

DOI: 10.1007/s11367-016-1199-1

Normalisation and weighting in life cycle assessment: *Quo Vadis*?

Massimo Pizzol*[1], Alexis Laurent[2], Serenella Sala[3], Bo Weidema[1], Francesca Verones[4], Christoph Koffler[5]

[1] *Aalborg University, Department of Development and Planning, Skibbrogade 5, 9000 Aalborg, Denmark*

[2] *Division for Quantitative Sustainability Assessment, Department of Management Engineering, Technical University of Denmark, Produktionstorvet 424, 2800 Kgs. Lyngby, Denmark*

[3]*European Commission, Joint Research Centre, Institute for Environment and Sustainability, Sustainability Assessment Unit, Via Enrico Fermi 2749, I-21027 Ispra (VA), Italy*

[4]*Industrial Ecology Programme, Department for Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), Sem Sælands vei 7, 7491 Trondheim, Norway*

[5] *thinkstep Inc., 170 Milk St, Boston, MA 02109, USA*

** Corresponding author. E-mail: massimo@plan.aau.dk*

DOI: 10.1007/s11367-016-1199-1

Abstract

Purpose Building on the rhetoric question *"Quo Vadis?"* (literally: *"Where are you going?"*) this article critically investigates the state of the art of normalisation and weighting approaches within life cycle assessment. It aims at identifying purposes, current practices, pros and cons, as well as research gaps in normalisation and weighting. Based on this information, the article wants to provide guidance to developers and practitioners. The underlying work was conducted under the umbrella of the UNEP-SETAC Life Cycle Initiative, Task Force on Cross-Cutting issues in LCIA.

Methods The empirical work consisted in: (i) an online survey to investigate the perception of the LCA community regarding the scientific quality and current practice concerning normalisation and weighting; (ii) a classification followed by systematic expert-based assessment of existing methods for normalisation and weighting according to a set of five criteria: scientific robustness, documentation, coverage, uncertainty and complexity.

Results and discussion The survey results showed that normalised results and weighting scores are perceived as relevant for decision-making, but further development is needed to improve uncertainty and robustness. The classification and systematic assessment of methods allowed for the identification of specific advantages and limitations.

Conclusions Based on the results, recommendations are provided to practitioners that desire to apply normalisation and weighting as well as to developers of the underlying methods.

Keywords Life Cycle Impact Assessment, Indicators, Multi-Criteria Decision Analysis, Survey, Review

DOI: 10.1007/s11367-016-1199-1

1. Normalisation and weighting, what is the problem?

According to the ISO 14044 standard on Life Cycle Assessment (LCA), normalisation is defined as *"calculating the magnitude of category indicator results relative to reference information"* and weighting as "*converting and possibly aggregating indicator results across impact categories using numerical factors based on value-choices*" (ISO, 2006b). Differently from classification and characterization, which are mandatory steps according to the ISO standards, normalisation and weighting are optional in Life Cycle Impact Assessment (LCIA) due to for example the potential biases and value choices they are respectively associated with, and the consequent commercial and legal concerns. The main criticism regarding normalisation is the bias due to the choice of normalisation references, which may change the conclusions drawn from the LCIA phase (Laurent and Hauschild, 2015; Norris, 2001). Criticism of weighting is even starker as ISO 14044 considers it to be "not scientifically based", excluding its use from LCA studies intended to support comparative assertions intended to be disclosed to the public because it is "based on value choices" (ISO, 2006a; ISO, 2006b). While this rather one-sided verdict seems to generally disregard the scientific basis of anything that is not based on natural sciences, it also glosses over the fact that LCA, and environmental modelling in particular, is full of value choices (Hertwich et al., 2000).

But even though normalisation and weighting are not required by the ISO standards on LCA, they are frequently applied in practice for different reasons, such as identifying "important" impact categories, understanding the meaning of results by comparing with more familiar references, or solving trade-offs between results (Ahlroth et al., 2011). Over time, several methods have been put forward for performing normalisation and weighting (Ahlroth, 2014; Laurent and Hauschild, 2015). Therefore, it is unclear today to what extent researchers and practitioners can or should correctly or legitimately apply normalisation and weighting and interpret their associated outputs. There is a risk that misunderstandings or malpractice in applying normalisation and weighting inadvertently or purposefully lead to biased or unfounded conclusions. This would ultimately result in mistrust in LCA results and, more generally, to poor decision support.

The objectives of this article thus are to: (i) clarify the purposes of normalisation and weighting; (ii) evaluate the current perception of these two steps in the LCA community; (iii) identify and define existing normalisation and weighting approaches; (iv) critically evaluate their pros and cons; (v) give recommendations to practitioners for applying them; (vi) set a research agenda to increase the

DOI: 10.1007/s11367-016-1199-1

robustness and reliability of normalisation and weighting in LCIA, by accounting for the latest and foreseen developments in this area. These objectives have been addressed as part of the UNEP/SETAC/ Life Cycle Initiative Task Force on Cross-Cutting Issues in LCIA, Working Group on Normalisation and Weighting. The working group has actively involved LCA researchers and practitioners from both academia and the private sector, who have brainstormed and critically discussed normalisation and weighting during the calendar year 2015, with the objectives of reaching mutual understanding and consensus on these topics and identifying research gaps. The findings of the working group were further discussed at a Pellston workshop in January 2016 and lead to a series of publically available recommendations and encouragement for further research. This paper builds on this work and provides detailed insights into the issue of normalisation and weighting.

2. Purposes of normalisation and weighting

The relevant ISO standards are not specific in defining what the purposes of normalisation and weighting are, respectively. In both normalisation and weighting, the purpose is linked to the goal and scope of the study, and therefore depends on the number and type of alternatives and impacts included, and on the system boundary and intended audience.

Normalisation can play a valuable role in informing the interpretation phase of LCA by answering the question whether the order of magnitude of the results is plausible. It can also be used to compare the results with a reference situation that is external to or independent from the case studies, which may facilitate the interpretation and communication of the impact results. For instance, comparing the impact results to the annual contributions of an average person, thus expressing the results in person equivalent, can be less abstract for an LCA practitioner or user of LCA results than dealing with a characterised result expressed in, e.g., kg-1,4-dichloro-benzene equivalent. Finally, normalisation can be a preparation for the weighting step, by bringing the characterised impact results to a scale that is relevant for further weighting and comparisons across impact categories (see below). However, when the weighting implies an assessment of the marginal damage at the specific level of impact (the slope of the dose-response curve) and therefore is independent of knowledge on the absolute size of the current or future impact, prior normalisation is not required and may even give rise to confusion and bias. Given its applications, normalisation could be seen as a step that helps interpreting the results, rather than a part of the "impact

DOI: 10.1007/s11367-016-1199-1

assessment" in its strict sense, since normalisation does not add to the quantification of potential environmental impacts.

Weighting can facilitate decision making in situations where trade-offs between impact category results do not allow choosing one preferable solution among the alternatives, or one improvement among possible ones. The weights applied are supposed to represent an evaluation of the relative importance of impacts, according to specific value choices, reflecting preferences of e.g. people, experts or organisations e.g. regarding time (present versus future impacts), geography (local versus global), urgency, political agendas, or cost (Ahlroth et al., 2011; Huppes and Van Oers, 2011; Pizzol et al., 2015). The cultural theory archetypes used in EcoIndicator99 and ReCiPe are a structured example of such value choices (Goedkoop et al., 2013; Goedkoop and Spriensma, 2001; Hofstetter et al. 2000; Goedkoop et al. 1998), but may be difficult to apply when decisions are made by a heterogeneous group of decision makers. With weighting, results may be summed across impact categories to arrive at a single score indicator for an LCA. Despite the practicality of the single score, it represents an issue highly debated in the LCA community since long time and still considered open (Kägi et al., 2015).

3. Perception within the LCA community

The perception of the LCA community towards normalisation and weighting has not been surveyed recently. To the authors' knowledge, the latest survey on the topic dates back to Hanssen (1999) and concludes that *"few if any LCA practitioners in the Nordic region rely solely on one specific weighting method, but use a number of methods and parameters in a given LCA study"* (Hanssen, 1999). In order to obtain an updated and more comprehensive picture, a survey was designed to investigate what the current practices with regard to normalisation and weighting are, and what the attitude of LCA researchers and practitioners towards normalisation and weighting is. The results' validity is limited to normalisation and weighting as general practices, because the survey was not designed to investigate or compare specific methods.

Respondents were first asked to state affiliation (exclusive choices: industry, academia, consultancy, public authority, other), professional status (non-exclusive choices: LCA practitioner, LCA method developer, user of LCA results), location (exclusive choices: Europe, North America, South America, Asia, Africa, Oceania), and years of experience with LCA. Secondly, respondents were invited to answer a symmetric set of questions on normalisation and weighting. After being

DOI: 10.1007/s11367-016-1199-1

asked about the familiarity with each practice (1-9 Likert-type scale), respondents were presented with questions covering two areas of investigation: "scientific quality" and "current practice". For each area of investigation, five variables were identified and then converted into questions (Table 1). Variables were chosen in order to cover different and as much as possible non-overlapping aspects or attributes of each area of investigation. For example, respondents were asked to state their opinion on the extent to which normalisation is "robust" on a 1-9 Likert-type scale where 1 = "Not at all" and 9 = "Extremely". The survey was distributed online via popular LCA fora (PRé mailing list and LinkedIn LCA-related groups) and made accessible for 20 days. The theoretical completion time was estimated as max. 8 min. This survey structure allowed for statistical testing of results, in particular: testing for correlation between variables, e.g. between perceived uncertainty and calculation of normalisation results; testing for differences between normalisation and weighting in the scores of the same variables, e.g. to see whether normalisation is perceived as more/less uncertain than weighting (paired t-test); testing for differences between groups, e.g. between practitioners affiliated to academia and industry (one way, between factors Analysis of variance (ANOVA)). Data management and statistical testing were performed with the software R (R Core Team, 2015). Detailed results can be accessed in Figure S1 and Tables S2-S5 (see Supporting Information). Parametric tests can be successfully applied on Likert data when their distribution approximates normal (Norman 2010; Sullivan and Artino 2013), but their use is still debated and therefore the non-parametric versions of these tests were also performed, which essentially provided the same results (see SI).A summary of the main findings is reported in the following.

The survey received a total of 257 responses, reduced to 216 after removing incomplete questionnaires. It is hardly possible to estimate the size of the population compared to the sample, and the value of 2500 LCA users reported by Pré is here taken as lower bound (Pré Consultants 2016). Looking at the composition of the sample, European academics were the largest group of respondents. Two thirds of respondents were classified as practitioners and the rest as model developers. Also two thirds were classified as "junior", in the sense that they have been working for less than ten years with LCA. Key findings from the survey are:

- According to the respondents, both normalisation and weighting are perceived negatively in terms of uncertainties and robustness. However, both procedures were perceived positively

DOI: 10.1007/s11367-016-1199-1

for their relevance for decision making. The responses were less clear regarding the practice-related issues. It was observed that participants do not to generally apply more than one method for normalisation and weighting, and choosing an appropriate normalisation method is perceived as difficult by most. The percentage distribution of answers on different variables is shown in Figure 1 and Figure 2.

- A positive but weak correlation was observed between the perceived scientific quality and the practice of normalisation and weighting (with r (214) values ranging between 0.19 and 0.58). Results thus show that the calculation of normalisation results and weighted scores is positively correlated with the perceived robustness, transparency, relevance, and validity, and poorly correlated with uncertainty. Higher positive correlation was observed between variables within the current practice area of investigation, both for normalisation and weighting. E.g. the variables "communication" and "selection" correlate positively with the variable "calculation", both in the case of normalisation ($r(214) = 0.80$ $p = .000$ and $r(214) = 0.65$ $p = .000$ respectively) and in the case of weighting ($r(214) = 0.90$ $p = .000$ and $r(214) = 0.75$ $p = .000$ respectively). A possible interpretation of this is that the calculation of normalised results and weighting scores is done primarily for interpretation and communication purposes, and not giving too much consideration to the quality of the science behind these practices.
- When comparing normalisation and weighting, the main differences are in the perceived robustness, $t (215) = 5.06$, $p = .000$, and transparency, $t (215) = 3.90$, $p = .000$, of the two practices, with weighting receiving lower scores than normalisation. It was observed that respondents to a larger extent calculate normalised than weighted results, $t (215) = 4.71$, $p = .000$, and to a larger extent use these for selecting the most relevant impact categories in a study, $t (215) = 3.72$, $p = .000$, and for communication of results, $t (215) = 3.76$, $p = .000$.
- The only appreciable difference between respondents is that academics, seniors, and developers declare a higher level of familiarity with both normalisation and weighting compared to practitioners and juniors, suggesting that approaching these topics requires more experience. The analysis of variance showed a significant effect of affiliation on familiarity with normalisation, $F (4, 201) = 5.338$, $p = .000$, and with weighting $F (4, 201) = 3.74$, $p = .005$; and a significant effect of experience on familiarity with normalisation $F (1, 204) = 32.3$, $p = .000$ and with weighting, $F (1, 204) = 13.9$, $p = .000$.

DOI: 10.1007/s11367-016-1199-1

Based on the survey results, it is clear that although normalised results and weighting scores are perceived by the respondents as relevant for decision making, further development is needed to improve uncertainty and robustness, especially in the case of weighting. Among the respondents, general practice is more focused on normalisation than on weighting, e.g., for communication or selection of relevant impacts.

4. Classification of normalisation and weighting approaches and methods

To univocally define and map approaches and methods used for normalisation and weighting, a classification inclusive of definitions is proposed here, based on the information available within the existing literature. In the classification, an "approach" is defined as a class of methods with similar underlying hypothesis and principles. "Methods" differ regarding calculation steps and practical implementation. The classification is summarised in Table 2 and Table 3 for normalisation and weighting, respectively.

5. Review of different approaches for normalisation and weighting

An assessment matrix was developed where all the methods included in the classification were systematically and critically evaluated according to a fixed set of criteria. Similar approaches for systematic review methods have been applied before in the LCA literature (Hauschild et al., 2013; Pizzol et al., 2015), and those were taken as starting point in the formulation of the following criteria: (i) *Scientific robustness*: What is the science behind the development of the method? (ii) *Documentation*: Does the documentation allow understanding and reproducing the method? (iii) *Coverage*: What is the scope of the method? (iv) *Uncertainty*: How are the uncertainties of the method addressed, handled, and described? (v) *Complexity*: What knowledge is required to apply the method in practice (i.e. to obtain new normalisation/weighting factors)? The matrix was filled in by a total of ten experts from within and outside the working group. Leading sub-questions were formulated for each criterion that allowed the experts to formulate a synthetic and qualitative assessment. Each expert filled in the matrix independently, fully or partly, focusing on the parts where he/she was more knowledgeable. The individual assessments were then merged into final matrices for normalisation and weighting, respectively –see SI. These matrices were then circulated within the working group and revised until a consensus version was obtained. These matrices were used to formulate method-specific recommendations for practitioners and developers. In the two

DOI: 10.1007/s11367-016-1199-1

following sub-sections, a summary of the advantages and limitations of each method is provided, based on the results of the expert assessment.

5.1.Normalisation approaches: advantages and limitations

Internal normalisation approach: Can help avoiding macroscopic mistakes (e.g. major over/underestimations of results common to all alternatives) and facilitate the communication of the results. Its use is limited to studies where more than one alternative is analysed, i.e. it can only be used in comparative assessments, and entails several drawbacks (e.g. possibility of changed or reversed ranking) if a subsequent weighting step is applied, see Norris (2001) and Laurent and Hauschild (2015). Outranking normalisation as an internal normalisation approach is specifically addressed in Section 6 as part of multi-criteria decision analysis methods.

External normalisation approach: As reflected by the number of literature sources calculating and documenting currently available normalisation references (see Table 2), regional production-based and, to a lesser extent, global normalisation references are the most used until now. Consumption-based territorial normalisation retains a very marginal role in LCA applications, with only two sets of normalisation references for Finland and the Netherlands (see Table 2) (Breedveld et al., 1999; Dahlbo et al., 2013; Laurent and Hauschild, 2015). The large data requirement for consistent inclusion of environmental flows related to imports still prevents the determination of reliable and comprehensive consumption-based normalisation references (Laurent and Hauschild, 2015).

National or regional production-based normalisation references have been the most used in LCA studies due to the early determination in the 90s of normalisation references for a number of countries, such as the Netherlands or Denmark (Breedveld et al., 1999; Wenzel et al., 1997). Several sets of normalisation references have emerged for other European countries and the European Union as a whole (Huijbregts et al., 2003; Sala et al., 2015), for Japan (Itsubo et al., 2004), for the United States (Bare et al., 2006), for Australia (Lundie et al., 2007) and for Canada (Lautier et al., 2010). Unlike internal normalisation, the use of national or regional production-based normalisation can help fulfil all the purposes indicated in Section 2, i.e. (i) checking the plausibility of the results, (ii) comparing impact results with those of the reference situation to serve the interpretation and the communication of the results, (iii) bringing the characterised impact results to a scale that is relevant for further weighting and comparisons across impact categories.

DOI: 10.1007/s11367-016-1199-1

However, when external normalisation is used a number of uncertainties and possible biases exist. Uncertainties are primarily related to normalisation data, i.e., the inventory used to calculate the normalisation references and those of the characterisation factors (Benini and Sala, 2016; Laurent and Hauschild, 2015). Important biases relate to discrepancies between the life cycle inventory of the product system under analysis and the life cycle inventory used for the calculation of a normalisation reference. A serious inconsistency can occur when two inventories do not apply the same life cycle inventory modelling approach (e.g., system boundaries, definition of marginal or average supply chains, co-product allocation procedures), and data sources. A specific inconsistency arises due to the consumption-driven nature of LCA models, as opposed to the production-based nature of most available normalisation references. Further bias can occur when the coverage of environmental flows differs between the normalisation inventory, the set of characterisation factors, and the life cycle inventory of the analysed system (Heijungs et al., 2007; Laurent and Hauschild, 2015). An example of this type of bias is choosing a normalisation reference that is incomplete with respect to the analysed system. This occurs when a substance, which is part of the life cycle inventory of a product system and drives the characterised results of a given impact of this system, is not part of the life cycle inventory used to calculate the normalisation reference. If that substance would be a strong contributor to the normalisation reference, which is then largely underestimated, the normalised results would then be largely overestimated (Heijungs et al., 2007). As such occurrences vary from one impact category to another, biases across impact categories emerge. Non-toxicity-related impact categories, such as climate change and acidification, typically include a limited number of environmental flows which are well monitored globally. Therefore, the coverage of these flows is relatively complete and reliable in both the LCI and the normalisation inventories, thus yielding reliable and accurate normalisation references and normalised results. In contrast, emissions of toxic substances are incompletely covered in currently available normalisation references (less complete than in the LCI of analysed systems), which thus leads in practice to observations of largely overestimated normalised results for toxicity-related impact categories. Such impact-specific overestimations can be dealt on an impact-specific basis when interpreting normalised results, but they may pose problems when comparing the impact results across impact categories in the weighting step.

With respect to the national and regional normalisation method, a risk is that some environmental flows, which are large contributors to the characterised impact results of the system, are not

DOI: 10.1007/s11367-016-1199-1

captured in the normalisation inventory because of a too narrow scope of the selected normalisation references. As product life cycles now stretch all over the world, a relative consistency could be gained by ensuring that all the environmental flows driving the impacts of the analysed system are stemming from locations that are included in the geographical scope of the selected normalisation references. This risk is alleviated with the use of global normalisation references, although other uncertainties are introduced, e.g. more uncertain normalisation inventories due to extrapolations required for mitigating the lack of environmental flow data for developing countries (Laurent and Hauschild, 2015).

As part of external normalisation, a recent method based on the definition of carrying capacities has emerged (Bjørn and Hauschild, 2015). The method is based on an alleged more ecologically oriented stance, giving priority to natural boundaries rather than to current levels of e.g. emission and resource use. It allows moving towards the integration of an "absolute sustainability" assessment in LCA (as opposed to current "relative" assessments). However, to which extent currently estimated carrying capacities and planetary boundaries reflect real ecological thresholds is debated (Mace et al., 2014; Nordhaus et al., 2012) and may require additional development of the scientific foundation of the boundaries and a testing of the normalised results in case studies, which are currently lacking. Besides, carrying-capacity references related to human health and resources are not yet available. The positioning of the method as a normalisation approach can also be debated as it could be regarded as a distance-to-target weighting. The externally-normalised results would thus represent the contribution of the system to current levels of impacts and the results after application of the carrying capacities would reflect how far (in excess or not) the normalised results are to the thresholds. Technically, the approach would be equivalent to the procedure adopted in the EDIP distance-to-target weighting methodology (Wenzel et al., 1997). From this perspective, any distance to target weighting method can be intended as an external normalisation method, as already suggested by various authors (Stranddorf et al., 2005; Pennington et al. 2004; Udo de Haes et al. 2002) as a way to circumvent the ISO LCA standards ban on weighting for comparative assertions. However, as a distance to target method, the planetary boundary approach would be quite different from existing methods using policy targets, because a scientific reasoning lies behind the established thresholds, which are beyond the policy relevance and are supposed to reflect an absolute sustainability level. There has been a certain attention to carrying-capacity based factors in LCA in recent times, and in the light of this there is a need to discuss which approaches are

DOI: 10.1007/s11367-016-1199-1

appropriate to integrate carrying capacities into LCIA, and whether they belong to characterisation, normalisation or weighting level.

5.2. Weighting approaches: advantages and limitations

Distance to target weighting approach: This approach may have intuitive value in a decision-making context, because it uses policy targets to derive weights. However, since all targets have equal weight, it can be questioned whether distance to target is really a weighting method in the sense of ranking of impact relevance, or rather a normalisation method (see Section 5.1). The method is not effective in the assessment of product improvements that reduce impacts for which the current situation is considered good or is not yet translated into a specific policy target. Furthermore, policy based targets are difficult to translate accurately into weights. One reason is that targets are not covering all the LCA elementary flows and impact categories (Castellani et al., 2016). The approach as it has been implemented does not address damage, i.e., a change in impacts far below a target may still be associated with a large damage and a change in impacts far above a target may be associated with little additional damage. While targets play an important and often necessary role in policy making and management, both for expressing and communicating intent and for monitoring progress, by transforming a general objective into specific actions that needs to be taken to achieve the objective, the existence of multiple targets for one objective can give rise to conflicts between targets, which can only be solved rationally by referring to the overall objective. Targets should not be misunderstood as ends in themselves because they are not valuable *per se*, but obtain value from their role in achieving the overall objective. If the overall objective is to reduce damage, the more logical choice is a weighting of the damage itself, rather than a specific targeted level of damage.

Panel weighting approach: In this approach, a panel of people is used to establish weighting factors for different environmental impact categories. The main issue with this method is that the selection of panellists influences results. In addition there are a number of other types of biases involved due the cognitive limits of the panel members and to the question format (c.f. Mettier and Scholz 2008; Mettier et al. 2006). While it is possible to have a panel that is representative of society at large in the statistical sense, in practice panellists often represent only a subset of all societal views on the issue. This latter case is different from other weighting approaches in that the panel-based weights represent the view of only a subset of society. This may or may not be a problem depending on the

DOI: 10.1007/s11367-016-1199-1

goal and scope of the study. In fact, one may argue that the panel composition needs to reflect the decision situation at hand. A "one-panel-fits-all" approach may not be the best solution from a decision support context in this situation. In industry, for example, different sectors have different hotspots and priorities and the decision makers may serve as panellists themselves or delegate this task to subject matter experts.

In addition, panellists may be biased regarding the absolute versus marginal value of the impact, although this bias can be corrected in the questionnaire. The way the questionnaire is designed can also create biases, for example, through the information explaining the impact categories. Also, the panellists are likely to have different levels of knowledge and experience with regard to the impacts they are asked to evaluate, i.e., observed (e.g., a friend dying from cancer), perceived (e.g., the panellist recently read an article about acidification) or predicted (e.g., peak oil). Lastly, people relate to what they understand and some environmental issues are difficult to understand. As a result, experts on a specific environmental issue will likely assign a higher importance to that issue.

While the term "panel method" puts a larger emphasis on capturing the preferences of a particular group of people, the way these preferences are processed further may vary based on the mathematical approach applied. From that perspective, panel methods belong into the realm of group decision making and represent a form of Multi-Attribute Decision Making (MADM) (Benoit and Rousseaux, 2003). According to Koffler (2008), multi-attribute group decisions can be characterized as follows: 1) The employed criteria in the form of category indicator results are cardinal measures, e.g., 12 t of CO_2 equivalents; 2) The scale transformation is performed via normalisation to a reference system, e.g., as in Western Europe; 3) The final normalisation to an interval [0; 1] can be achieved by dividing all values by the maximum value; 4) All group members apply the same set of agreed criteria; and thus, the same set of impact categories.

It seems that normalisation and weighting in LCA was originally developed without referring to the already existing theoretical foundation of MADM as it wasn't until the turn of the millennium that this became more widely recognized in literature (Hertwich and Hammitt, 2001; Seppälä et al., 2001). While the weighted summation in LCA closely corresponds to a rather 'old-fashioned' MADM method called Simple Additive Weighting (Fishburn, 1967), more advanced MADM methods have also been applied to LCA since then (Benoit and Rousseaux, 2003; Koffler et al., 2008; Prado-Lopez et al., 2014; Seager and Linkov, 2008). Note that while MADM methods were

DOI: 10.1007/s11367-016-1199-1

here grouped under weighting approaches, normalisation is an integral part of all MADM methods when dealing with attributes that are measured in different units. MADM were grouped under weighting because naturally their focus is more on eliciting preferences and processing them rather on the multiple and specific forms of normalisation that are available in LCA.

The integration of approaches from the MADM field into LCA seems to be increasingly demanded. Further research is needed before a recommended framework for their integration can be proposed to LCA practitioners. More testing on case studies and feedback from users are deemed required to estimate the robustness and relevance of the approaches. Benefits of use in combination with non-environmental indicators should also be investigated in order to arrive at a complete sustainability assessment.

Monetary weighting approach: The different uses of monetary valuation-based approached for weighting in LCA have been addressed recently in the literature (Pizzol et al, 2015; Ahlroth, 2014; Bachmann 2011), and only key advantages and limitations are here re-stated, whereas the reader interested in additional details is referred to the above mentioned publications. In general, an advantage of monetary weighting is that monetary units may be more familiar and easier to relate to for most audiences, compared to weights derived with other methods. There is a difference between the capacity of different monetary valuation methods to cover midpoint and endpoint impacts and therefore in their application for weighting at different points of the impact chain. The observed and revealed preferences methods are usually applied at midpoint: their main limitation is that these typically cover use value only and are too case-specific for use at endpoint. An exception is the budget constraint method (Weidema, 2009) where observed preferences are used to derive an estimated value for a QALY. Stated preferences methods can overcome this limitation. Regarding implementation, stated preference methods may have similar issues as panel weighting regarding questionnaire design. Since both revealed and stated preferences methods are heavily based on statistical analysis techniques, it is generally possible to obtain precise estimates of the uncertainties associated with the derived monetary weights (Boardman, 2006). A general drawback of monetary valuation-based weighting methods is that some individuals may oppose their use due to ethical reasons, e.g. finding inappropriate to place a monetary value on e.g. human life or biodiversity (Ludwig 2000). A common misconception in such cases is that the absolute value of e.g. human life

DOI: 10.1007/s11367-016-1199-1

is valuated, whereas in reality monetary valuation only determines the willingness to pay for marginal changes in the availability of non-market goods.

Binary weighting approach: Equal weighting is not science-based and could be mistaken for a "neutral" weighting while it is not. Footprinting is often called "implicit weighting" where several impacts are deliberately disregarded, even though they may be important. Another problem with footprinting is that it may lead to sub-optimisation where the system is optimised to reduce the impact on one category only and trade-offs are ignored (Laurent et al., 2012), which is also referred to as "burden shifting" and, as such, is not aligned with one of the key principles of LCA. Another recurrent issue is the lack of communication of the value choices made to the user of the footprint information.

Some cross-method considerations should be made regarding weighting at midpoint versus endpoint. By performing the assessment early in the impact chain, the rest of the impact pathway is left to the practitioners to fill in by other means than characterisation. The main argument that has been put forward for performing the assessment at the early stage of the impact pathway is that the uncertainty of the impact is lower at these early stages (Hauschild and Potting, 2005, Finnveden et al., 2009). However, in assessments where weighting is required, this implies that important sources of both aleatoric and epistemic uncertainties are left for the practitioners to address when selecting and applying the weighting approach in a decision-making context (Weidema, 2009). Accordingly, an argument for performing weighting at the final damage level is that it reduces the number of valuations that need to be made, and thereby reduces the risk of inconsistencies between the larger number of different valuations that would otherwise be required to be performed at the midpoint level in the impact pathways. For example, an assessment of respiratory impacts of particulate emissions and another assessment of ground-level ozone formation may use different weights for the same health impacts – an inconsistency that would be avoided if performing the weighting at the level of human life years.

Developing weighting factors at midpoint or endpoint also presents practical method-specific issues. For example, targets at midpoint and at endpoint coexist and are difficult to be combined and properly accounted into a distance to target weighting method: e.g., in the EU air quality directive (EC, 2008) targets are expressed both as targets on single substance emissions (e.g. emission of $PM_{2.5}$) and as targets in Disability Adjusted Life Years (DALY). In panel weighting methods, the

DOI: 10.1007/s11367-016-1199-1

understanding and ranking of several midpoint categories may result in being more challenging for respondents than in the case of few endpoint ones, as there can be too many issues for panellists to consider in case of weighting midpoint impact categories (Huppes and Van Oers, 2011).

6. Conclusions and recommendations

This study has investigated what the purposes of normalisation and weighting are in LCA, and what the perceptions of the LCA community towards these practices are. In addition, a classification of normalisation approaches and methods was proposed as well as an analysis of their respective advantages and limitations. To conclude, the authors' recommendations for practitioners and method developers are provided based on the survey and assessment.

6.1.Improving practice: recommendations for LCA practitioners

The recommendations should not be seen as recommendations to use any specific normalisation or weighting methods nor as recommendations to use normalisation or weighting at all, but as recommendations for good practice for the practitioners when it has been decided to use normalisation or weighting.

In general, it is recommended to document and justify the choice of any normalisation references and weighting methods applied, as required by the ISO standards on LCA. The normalised results and weighted scores should be communicated clearly by e.g. reporting units and explaining their meaning, as these may not be easily understandable to audiences beyond LCA experts. It is further recommended to integrate uncertainty assessment, at least in a qualitative way, for normalisation and weighting results (e.g. uncertainties and biases introduced in the resulting impact scores). Scenario analysis should also be performed whenever possible, e.g., by applying more than one method and possibly take advantage of ensemble modelling techniques that allows to include simultaneously all relevant methods (Huppes et al., 2012). This will allow addressing uncertainties explicitly and testing the robustness of the conclusions. Also, it is recommended to clearly interpret results referring to the purposes and limitations of the chosen normalisation and weighting approaches, and to make sure that the decision makers are aware of the uncertainties and potential biases related to the use of normalisation and weighting.

Specifically for normalisation, practitioners should consider if normalisation is needed, and if so, report clearly for which of the purposes mentioned above (see Section 2 and 5.1) normalisation is

DOI: 10.1007/s11367-016-1199-1

needed. Practitioners should be aware of potential inconsistencies in external normalisation due to the fact that the normalisation inventory may not use the same modelling approach (system boundaries, marginal or average supply chains, co-product allocation procedures; some of them only relevant at sub-global level) and data sources as the analysed product system life cycle inventory, and that the characterisation may not be performed with the same characterisation method. For example, if the processes within the system boundaries are globally distributed, it is recommended to prioritise the use of a set of global normalisation references. If using global normalisation references is not possible due to data availability, then this limitation should be acknowledged and the implications for the final decision discussed in the interpretation of the study. Consumption-based normalisation references are in theory preferable to production-based ones because of a better consistency with the geographical scope of LCA studies that are by nature consumption-driven (see Laurent and Hauschild, 2015) although in practice nearly none are readily available today. It is important not to overuse normalisation and confuse it with weighting: comparisons across impact categories to solve trade-offs cannot be done at the normalisation step, but require weighting. In relation to this last point, it is recommended that practitioners do not exclude from the study any impact categories after normalisation, whether deemed too high or too low compared to others.

For weighting, practitioners should prefer weighting of damage rather than weighting of the distance to a target for the damage (see section 5.2). Practitioners using panel methods (including panels for monetary valuation) should prefer panels of affected stakeholders to expert panels unless the relevant stakeholders delegate the task to experts or other representatives. It is recommended to provide information on panel composition and criteria for stakeholder selection. For monetary valuation methods observed preferences (market prices) should be used whenever possible, i.e. when only use values are involved and for the value of a human life-year, and results from choice modelling (which does not need to include a monetary instrument) when there is a need to include stated preferences. It is recommended that practitioners doing equal weighting and footprinting provide explicit statements of implicit binary weighting and motivate selection/exclusion of impact categories.

6.2.Improving the science: recommendations to method developers

DOI: 10.1007/s11367-016-1199-1

Similarly to the recommendations for practitioners, the following should not be seen as recommendations of any specific normalisation or weighting methods nor as recommendations to apply normalisation or weighting at all, but exclusively as recommendations for good practice for those wishing to improve current normalisation or weighting methods.

In general, it is recommended to develop techniques to estimate, and provide estimates of, uncertainties related to external normalisation references and weighting factors. Tools to allow practitioners to study uncertainty propagation from the characterization step to normalisation and weighting are needed. Clear and transparent communication and dissemination (to practitioners) of normalisation and weighting methods' principles and hypotheses, as well as their reproducibility, is highly recommended. The awareness of LCA practitioners of the benefits, limitations and possible pitfalls of using different normalisation and weighting methods should be increased.

Improving the basis for normalisation is recommended by developing consistent and sufficiently complete inventories of emissions and resource consumptions for the world. This requires collecting emission data from developing countries and determining consistent extrapolation techniques to fill in gaps. To avoid the biases mentioned above, the level of completeness needs as a minimum to match conventional life cycle inventories. Method developers should derive external normalisation references that are to the best extent consistent with the modelling approaches, data sources, and characterisation methods applied in LCA studies: this may require the development of new flexible methods for the calculation of *ad hoc* normalisation references. Developing inventories at national scale to allow consumption-based normalisation references (global coverage needed to consistently encompass imports) is also recommended, and using input-output LCA seems a promising way to address this issue. New approaches for normalisation, adapting to the latest LCIA developments, should be investigated. This includes for example the integration of spatial differentiation into normalisation, which will also require a discussion of whether normalisation should be performed at regional or national levels before aggregation of the characterised results across spatial scales. Recent improvements of globally differentiated methodologies could facilitate such attempts although it requires normalisation references at relevant scales, which today do not exist.

When developing new sets of weighting factors at both midpoint and endpoint, it is recommended to test for their robustness and ensure their transparency. Weighting methods should be developed

DOI: 10.1007/s11367-016-1199-1

by striving towards consistency, i.e., by having weighting factors derived as consistently as possible across impact categories and using consistent assumptions (if not methods), and towards completeness by having all categories included in the method. It is recommended to prioritize the development of weighting factors at endpoint, as this can reduce the number of valuations to be made and the related inconsistencies, followed by a thorough analysis and quantification of the uncertainties. In addition, the relationships between aggregation of the normalised results over spatial scales and weighting should be investigated to determine whether damages to areas of protection should be weighted the same across the different countries or regions.

Acknowledgements

We would like to thank Valentina Prado, Jane Bare, Tommie Ponsioen, and Anne-Marie Boulay for their contributions to the working group activities. Thanks are also due to Anders Bjørn and Viêt Cao who kindly contributed with comments and additions to the assessment matrices.

DOI: 10.1007/s11367-016-1199-1

References

Ahlroth S (2014) The use of valuation and weighting sets in environmental impact assessment. Resour Conserv Recycl 85:34–41. doi: http://dx.doi.org/10.1016/j.resconrec.2013.11.012

Ahlroth S, Finnveden G (2011) Ecovalue08–A new valuation set for environmental systems analysis tools. J Clean Prod 19:1994–2003. doi: http://dx.doi.org/10.1016/j.jclepro.2011.06.005

Ahlroth S, Nilsson M, Finnveden G, et al (2011) Weighting and valuation in selected environmental systems analysis tools - suggestions for further developments. J Clean Prod 19:145–156. doi: 10.1016/j.jclepro.2010.04.016

Alarcon B, Aguado A, Manga R, Josa A (2011) A Value Function for Assessing Sustainability: Application to Industrial Buildings. Sustainability 3:35.

Bachmann TM (2011) Optimal pollution: the welfare economic approach to correct market failures. In: Nriagu J (ed) Encyclopedia on Environmental Health. Elsevier, Burlington, pp 264–274

Bare J, Gloria T, Norris G (2006) Development of the method and US normalization database for life cycle impact assessment and sustainability metrics. Environ Sci Technol 40:5108–5115.

Benini L, Sala S (2016) Uncertainty and sensitivity analysis of normalization factors to methodological assumptions. Int J Life Cycle Assess 21:224–236. doi: 10.1007/s11367-015-1013-5

Benoit V, Rousseaux P (2003) Aid for aggregating the impacts in life cycle assessment. Int J Life Cycle Assess 8:74–82. doi: 10.1065/lca2002.12.104

Bjørn A, Hauschild M (2015) Introducing carrying capacity-based normalisation in LCA: framework and development of references at midpoint level. Int J Life Cycle Assess 20:1005–1018. doi: 10.1007/s11367-015-0899-2

Boardman AE, Greenberg DH, Vining AR, Weimer DL (2006) Cost-Benefit Analysis, Concepts and Practice. Pearson

Breedveld L, Lafleur M, Blonk H (1999) A framework for actualising normalisation data in LCA: Experiences in the Netherlands. Int J Life Cycle Assess 4:213–220.

Castellani V, Benini L, Sala S, Pant R (2016) A distance-to-target weighting method for Europe 2020. Int J Life Cycle Assess 1–11. doi: 10.1007/s11367-016-1079-8 LB - Castellani2016

Cucurachi S, Sala S, Laurent A, Heijungs R (2014) Building and characterizing regional and global emission inventories of toxic pollutants. Environ Sci Technol 48:5674–5682. doi: 10.1021/es405798x

Dahlbo H, Koskela S, Pihkola H, et al (2013) Comparison of different normalised LCIA results and their feasibility in communication. Int J Life Cycle Assess 18:850–860. doi: 10.1007/s11367-012-0498-4

EC (2008) Directive 2008/50/EC of 21 May 2008 on ambient air quality and cleaner air for Europe.

Finnveden G, Eldh P, Johansson J (2006) Weighting in LCA based on ecotaxes: Development of a mid-point method and experiences from case studies. Int J Life Cycle Assess 11:81–88.

Finnveden G, Hauschild MZ, Ekvall T, et al (2009) Recent developments in Life Cycle Assessment. J Environ Manage 91:1–21.

Fishburn PC (1967) Additive Utilities with Incomplete Product Set: Applications to Priorities and Assignments. Operations Research Society of America (ORSA), Baltimore, MD, U.S.A.

Foley J, Lant P (2009) Regional normalisation figures for Australia 2005/2006-inventory and characterisation data from a production perspective. Int J Life Cycle Assess 14:215–224. doi: 10.1007/s11367-009-0063-y

Frischknecht R, Steiner R, Jungbluth N (2009) The Ecological Scarcity Method – Eco-Factors 2006. A method for impact assessment in LCA. Federal Office for the Environment (FOEN), Bern

Goedkoop M, Heijungs R, Huijbregts M, et al (2013) ReCiPe 2008 A life cycle impact assessment method which comprises harmonised category indicators at the midpoint and the endpoint level, First edition (version 1.08), Report I: Characterisation. PRé Consultants, Amersfoort,

CML University of Leiden, RUN Radboud University Nijmegen, RIVM Bilthoven - Netherlands, The Netherlands

Goedkoop M, Spriensma R (2001) The Eco-indicator 99 - A damage oriented method for Life Cycle Impact Assessment. Pré Consultants B.V.

Goedkoop M, Hofstetter P, Müller-Wenk R, Spriemsma R (1998) The ECO-indicator 98 explained. Int J Life Cycle Assess 3:352–360. doi: 10.1007/BF02979347

Hanssen OJ (1999) Status of Life Cycle Assessment (LCA) activities in the Nordic Region. Int J Life Cycle Assess 4:315–320. doi: 10.1007/BF02978519

Hauschild M, Potting J (2005) Spatial differentiation in Life Cycle impact assessment - The EDIP2003 methodology. Environmental news No. 80. Danish Ministry of Environment, Environmental Protection Agency.

Hauschild MZ, Goedkoop M, Guinee J, et al (2013) Identifying best existing practice for characterization modeling in life cycle impact assessment. Int J Life Cycle Assess 18:683–697. doi: 10.1007/s11367-012-0489-5

Heijungs R, Guinee J, Kleijn R, Rovers V (2007) Bias in normalization: Causes, consequences, detection and remedies. Int J Life Cycle Assess 12:211–216. doi: 10.1065/lca2006.07.260

Hertwich EG, Hammitt JK (2001) A decision-analytic framework for impact assessment part I: LCA and decision analysis. Int J Life Cycle Assess 6:5–12. doi: 10.1007/bf02977588

Hertwich EG, Hammitt JK, Pease WS (2000) A Theoretical Foundation for Life-Cycle Assessment. J Ind Ecol 4:13–28. doi: 10.1162/108819800569267

Hofstetter P, Baumgartner T, Scholz RW (2000) Modelling the valuesphere and the ecosphere: Integrating the decision makers' perspectives into LCA. Int J Life Cycle Assess 5:161–175. doi: 10.1007/BF02978618

Huijbregts MAJ, Breedveld L, Huppes G, et al (2003) Normalisation figures for environmental life-cycle assessment: The Netherlands (1997/1998), Western Europe (1995) and the world (1990 and 1995). J Clean Prod 11:737–748. doi: 10.1016/S0959-6526(02)00132-4

Huppes G, Oers L, Pretato U, Pennington DW (2012) Weighting environmental effects: Analytic survey with operational evaluation methods and a meta-method. Int J Life Cycle Assess 17:876–891. doi: 10.1007/s11367-012-0415-x

Huppes G, Van Oers L (2011) Background review of existing weighting approaches in Life Cycle Impact Assessment (LCIA). Joint Research Centre - Institute for Environment and Sustainability

ISO (2006a) ISO 14040 - Environmental management - Life cycle assessment - Principles and framework. International Standard Organization

ISO (2006b) ISO 14044 - Environmental management - Life cycle assessment - Requirements and guidelines. International Standard Organization

ISO (2014) ISO 14046 - Environmental management - Water footprint - Principles, requirements and guidelines. 33.

Itsubo N, Murakami K, Kuriyama K, et al (2015) Development of weighting factors for G20 countries—explore the difference in environmental awareness between developed and emerging countries. Int J Life Cycle Assess. doi: 10.1007/s11367-015-0881-z

Itsubo N, Sakagami M, Kuriyama K, Inaba A (2012) Statistical analysis for the development of national average weighting factors-visualization of the variability between each individual's environmental thoughts. Int J Life Cycle Assess 17:488–498. doi: 10.1007/s11367-012-0379-x

Itsubo N, Sakagami M, Washida T, et al (2004) Weighting across safeguard subjects for LCIA through the application of conjoint analysis. Int J Life Cycle Assess 9:196–205. doi: 10.1065/lca2003.11.142

Kallas Z, Lambarraa F, Gil JM (2011) A stated preference analysis comparing the Analytical Hierarchy Process versus Choice Experiments. Food Qual Prefer 22:181–192. doi: 10.1016/j.foodqual.2010.09.010

Kim J, Yang Y, Bae J, Suh S (2013) The Importance of Normalization References in Interpreting Life Cycle Assessment Results. J Ind Ecol 17:385–395. doi: 10.1111/j.1530-9290.2012.00535.x

Koffler C, Schebek L, Krinke S (2008) Applying voting rules to panel-based decision making in LCA. Int J Life Cycle Assess 13:456–467. doi: 10.1007/s11367-008-0019-7

Kägi T, Dinkel F, Frischknecht R, et al (2015) Session "Midpoint, endpoint or single score for decision-making?"—SETAC Europe 25th Annual Meeting, May 5th, 2015. Int J Life Cycle Assess 21:129–132. doi: 10.1007/s11367-015-0998-0

Laurent A, Hauschild MZ (2015) Normalisation. In: Hauschild MZ, Huijbregts MA (eds) Life Cycle Impact Assessment. Springer Science+Business Media B.V., pp 271–300

Laurent A, Lautier A, Rosenbaum RK, et al (2011) Normalization references for Europe and North America for application with USEtox{\texttrademark} characterization factors. Int J Life Cycle Assess 16:728–738. doi: 10.1007/s11367-011-0285-7

Laurent A, Olsen SI, Hauschild MZ (2012) Limitations of Carbon Footprint as Indicator of Environmental Sustainability. Environ Sci Technol 46:4100–4108. doi: 10.1021/es204163f

Laurent A, Olsen SI, Hauschild MZ (2011b) Normalization in EDIP97 and EDIP2003: updated European inventory for 2004 and guidance towards a consistent use in practice. Int J Life Cycle Assess 16:401–409.

Lautier A, Rosenbaum RK, Margni M, et al (2010) Development of normalization factors for Canada and the United States and comparison with European factors. Sci Total Environ 409:33–42. doi: http://dx.doi.org/10.1016/j.scitotenv.2010.09.016

Ludwig D (2000) Limitations of Economic Valuation of Ecosystems. Ecosystems 3:31–35. doi: 10.2307/3658665

Lundie S, Huijbregts MAJ, Rowley H V, et al (2007) Australian characterisation factors and normalisation figures for human toxicity and ecotoxicity. J Clean Prod 15:819–832. doi: 10.1016/j.jclepro.2006.06.019

Mace GM, Reyers B, Alkemade R, et al (2014) Approaches to defining a planetary boundary for biodiversity. Glob Environ Chang 28:289–297. doi: http://dx.doi.org/10.1016/j.gloenvcha.2014.07.009

Mettier T, Scholz R, Tietje O (2006) Measuring Preferences on Environmental Damages in LCIA. Part 1: Cognitive Limits in Panel Surveys (9 pp). Int J Life Cycle Assess 11:394–402. doi: 10.1065/lca2005.08.218

Mettier T, Scholz RW (2008) Measuring preferences on environmental damages in LCIA. Part 2: choice and allocation questions in panel methods. Int J Life Cycle Assess 13:468–476. doi: 10.1007/s11367-008-0022-z

Nordhaus T, Shellenberger M, Blomqvist L (2012) The planetary boundaries hypothesis. A Review of the Evidence. The Breakthrough Institute, Oakland

Norman G (2010) Likert scales, levels of measurement and the "laws" of statistics. Adv Heal Sci Educ 15:625–632. doi: 10.1007/s10459-010-9222-y

Norris GA (2001) The requirement for congruence in normalization. Int J Life Cycle Assess 6:85–88.

Norris GA, Marshall HE (1995) Multiattribute Decision Analysis Method for Evaluating Buildings and Building Systems. Building and Fire Research Laboratory, National Institute of Standards and Technology, Gaithersburg

Pennington DW, Potting J, Finnveden G, et al (2004) Life cycle assessment Part 2: Current impact assessment practice. Environ Int 30:721–739.

Pizzol M, Weidema BP, Brandão M, Osset P (2015) Monetary valuation in Life Cycle Assessment: A review. J Clean Prod 86:170–179. doi: 10.1016/j.jclepro.2014.08.007

Prado-Lopez V, Seager TP, Chester M, et al (2014) Stochastic multi-attribute analysis (SMAA) as an interpretation method for comparative life-cycle assessment (LCA). Int J Life Cycle Assess 19:405–416. doi: 10.1007/s11367-013-0641-x

Pré Consultants (2016) LCA discussion list website. https://www.pre-sustainability.com/lca-discussion-list.

R Core Team (2005) R: A language and environment for statistical computing.

Ridoutt B, Fantke P, Pfister S, et al (2015) Making Sense of the Minefield of Footprint Indicators. Environ Sci Technol 49:2601–2603. doi: 10.1021/acs.est.5b00163

Ryberg M, Vieira MDM, Zgola M, et al (2014) Updated US and Canadian normalization factors for TRACI 2.1. Clean Technol Environ Policy 16:329–339. doi: 10.1007/s10098-013-0629-z

Rüdenauer I, Gensch C-OC-O, Grießhammer R, Bunke D (2005) Integrated Environmental and Economic Assessment of Products and Processes. J Ind Ecol 9:105–116. doi: 10.1162/108819805775248061

Sala S, Benini L, Mancini L, Pant R (2015) Integrated assessment of environmental impact of Europe in 2010: data sources and extrapolation strategies for calculating normalisation factors. Int J Life Cycle Assess 20:1568–1585. doi: 10.1007/s11367-015-0958-8

Seager TP, Linkov I (2008) Coupling multicriteria decision analysis and life cycle assessment for nanomaterials. J Ind Ecol 12:282–285. doi: 10.1111/j.1530-9290.2008.00048.x

Seppälä J, Basson L, Norris GA (2001) Decision Analysis Frameworks for Life-Cycle Impact Assessment. J Ind Ecol 5:45–68. doi: 10.1162/10881980160084033

Sleeswijk AW, van Oers LFCM, Guinée JB, et al (2008) Normalisation in product life cycle assessment: An LCA of the global and European economic systems in the year 2000. Sci Total Environ 390:227–240. doi: 10.1016/j.scitotenv.2007.09.040

Steen B (1999a) A systematic approach to environmental strategies in product development (EPS). Version 2000 - General system characteristics. Centre for Environmental Assessment of Products and Material Systems. Chalmers University of Technology, Technical Environmental Planning

Steen B (1999b) A systematic approach to environmental strategies in product development (EPS). Version 2000 - Models and data of the default methods. Centre for Environmental

Assessment of Products and Material Systems. Chalmers University of Technology, Technical Environmental Planning

Stranddorf HK, Hoffmann L, Schmidt A (2005) LCA Guideline. Update on Impact Categories, Normalisation and Weighting in LCA – Selected EDIP97 data. Danish Environmental Protection Agency, Copenhagen

Strauss K, Brent AC, Hietkamp S (2006) Characterisation and normalisation factors for life cycle impact assessment mined abiotic resources categories in South Africa: The manufacturing of catalytic converter exhaust systems as a case study. Int J Life Cycle Assess 11:162–171. doi: 10.1065/lca2004.10.183

Sullivan GM, Artino a R (2013) Analyzing and interpreting data from likert-type scales. J Grad Med Educ 5:541–2. doi: 10.4300/JGME-5-4-18

Saaty TL (2008) Decision making with the analytic hierarchy process. Int J Serv Sci 1:83. doi: 10.1504/IJSSCI.2008.017590

Udo de Haes, H.A., Finnveden, G., Goedkoop, M., Hauschild, M., Hertwich, E.G., Hofstetter, P., Jolliet, O., Klöpffer, W., Krewitt, W., Lindeijer, E.W., Müller-Wenk, R., Olsen, S.I., Pennington, D.W., Potting, J. and Steen, B. (Eds.) (2002): Life-Cycle Impact Assessment: Striving towards best practise. SETAC- Press, Pensacola, Florida.

Weidema B, Hauschild MZ, Jolliet O (2008) Preparing characterisation methods for endpoint impact assessment - Annex II of Eder P & Delgado L (eds.): "Environmental improvement potentials of meat and dairy products." Institute for Prospective Technological Studies, Sevilla

Weidema BP (2009) Using the budget constraint to monetarise impact assessment results. Ecol Econ 68:1591–1598.

Weiss M, Patel M, Heilmeier H, Bringezu S (2007) Applying distance-to-target weighing methodology to evaluate the environmental performance of bio-based energy, fuels, and materials. Resour Conserv Recycl 50:260–281. doi: http://dx.doi.org/10.1016/j.resconrec.2006.06.003

DOI: 10.1007/s11367-016-1199-1

Wenzel H, Hauschild MZ, Alting L (1997) Environmental Assessment of Products. Volume 1 - Methodology, Tools and Case Studies in Product Development. Chapman & Hall, Thomson Science, London, UK

DOI: 10.1007/s11367-016-1199-1

Table 1. The symmetric questionnaire structure.

Subject	**Area of investigation**	**Variable**	**Question**
Normalisation	Scientific quality *"In your opinion…"* (1 = Not at all; 9 = Extremely)	Robustness	*How robust are normalisation factors?*
		Transparency	*How transparent are normalisation factors?*
		Uncertainty	*How uncertain are normalisation factors?*
		Relevance	*How relevant are normalized impact results in a decision making context?*
		Validity	*How well does normalisation meet its purpose?*
	Current practice *"How often are these situations occurring in your practice with normalisation?* (1 = Never; 9 = Always)"	Calculation	*When performing a LCA study, I calculate normalized impact results*
		Communication	*When presenting LCA results, I use normalized impact results*
		Selection	*I use normalisation to determine the most relevant impact categories for an LCA*
		Choice	*I experience difficulties in selecting which set of normalisation factors to use*
		Coverage	*I apply more than one normalisation method*
Weighting	Scientific quality *"In your opinion…"* (1 = Not at all; 9 = Extremely)	Robustness	*How robust are weighting factors?*
		Transparency	*How transparent are weighting factors?*
		Uncertainty	*How uncertain are weighting factors?*
		Relevance	*How relevant are weighted impact scores in a decision making context?*
		Validity	*How well does weighting capture the values of the group involved?*
	Current practice *"How often are these situations occurring in your practice with weighting?* (1 = Never; 9 = Always)"	Calculation	*When performing a LCA study, I calculate weighted impact scores*
		Communication	*When presenting LCA results, I use weighted impact scores*
		Selection	*I use weighting to determine the most relevant impact categories for an LCA*
		Choice	*I experience difficulties in selecting which set of weighting factors to use*
		Coverage	*I apply more than one weighting method*

DOI: 10.1007/s11367-016-1199-1

Figures Captions

Figure 1. Survey results, scientific quality of normalisation and weighting, percentages of answers in the 1-9 scale (1=low; 9=high). Note that the "Uncertainty" variable is reversed.

Figure 2. Survey results, current practice with normalisation and weighting, percentages of answers in the 1-9 scale. The "Choice" variable is reversed.

DOI: 10.1007/s11367-016-1199-1

Table 2 Classification of normalisation approaches and methods

Approach	*Principle*	*Method*	*Definition*	*Sources and examples*
Internal normalisation	Impacts are normalised with references linked to the alternative(s)* assessed in the study	Division by baseline	Characterized indicator results for several alternatives are divided by the scores obtained for one alternative (= baseline).	(Laurent and Hauschild, 2015; Norris, 2001)
		Division by maximum	Characterized indicator results for several alternatives are divided by the scores obtained for the alternative with the highest score in each impact category	(Laurent and Hauschild, 2015; Norris, 2001) Norris and Marshal, 1995
		Division by sum	Characterized indicator results for several alternatives are divided by the sum of the scores obtained for all alternatives	(Laurent and Hauschild, 2015; Norris, 2001) Norris and Marshal, 1995
		Outranking normalisation	Use of pair-wise comparisons to evaluate the significance of mutual differences from characterized indicator results and to reflect impact categories with critical differences between alternatives. The method is non-linear.	(Prado-Lopez et al., 2014)
External normalisation	Impacts are normalised with references that are external and thus independent of the object of the LCA.	Global normalisation	Characterized indicator results of the system(s) under study are divided by the characterized indicator results of the total activities taking place in the world over the reference duration (assumed balance between consumption and production)	(Huijbregts et al., 2003; Itsubo et al., 2015; Sleeswijk et al., 2008; Stranddorf et al., 2005)
		Production-based, territorial systems	Characterized indicator results of the system(s) under study are divided by the characterized indicator results associated with all territorial activities in a region or country, including its exports but excluding its imports, thus accounting for all environmental flows that take place within the physical or geographical boundaries of that region/country over the reference duration	(Bare et al., 2006; Breedveld et al., 1999; Cucurachi et al., 2014; Dahlbo et al., 2013; Foley and Lant, 2009; Huijbregts et al., 2003; Itsubo et al., 2015; Itsubo et al., 2012; Itsubo et al., 2004; Kim et al., 2013; Laurent and Hauschild, 2015; Laurent et al., 2011a; Laurent et al., 2011b; Lautier et al., 2010; Lundie et al., 2007; Ryberg et al., 2014; Sala et al., 2015; Sleeswijk et al., 2008; Stranddorf et al., 2005; Strauss et al., 2006; Wenzel et al., 1997)
		Consumption-based, territorial systems	Characterized indicator results of the system(s) under study are divided by the characterized indicator results associated with the total territorial consumption of a region/country, including its imports but excluding its exports. It thus accounts for the environmental flows from all up- and downstream processes needed to support the consumption activities of that region/country over the reference duration, including those that occur outside its physical or geographical boundaries as a consequence of the activities taking place within that region or nation.	(Breedveld et al., 1999; Dahlbo et al., 2013; Laurent and Hauschild, 2015)
		Carrying-	Characterized indicator results are divided by the normalisation references representing	(Bjørn and Hauschild, 2015)

DOI: 10.1007/s11367-016-1199-1

	capacity -based (Planetary boundaries)	the carrying capacity of the reference system (e.g. world) for each impact category

* Alternatives are defined as any compared system (whether they relate to different scenarios of a same product system or to different product systems).

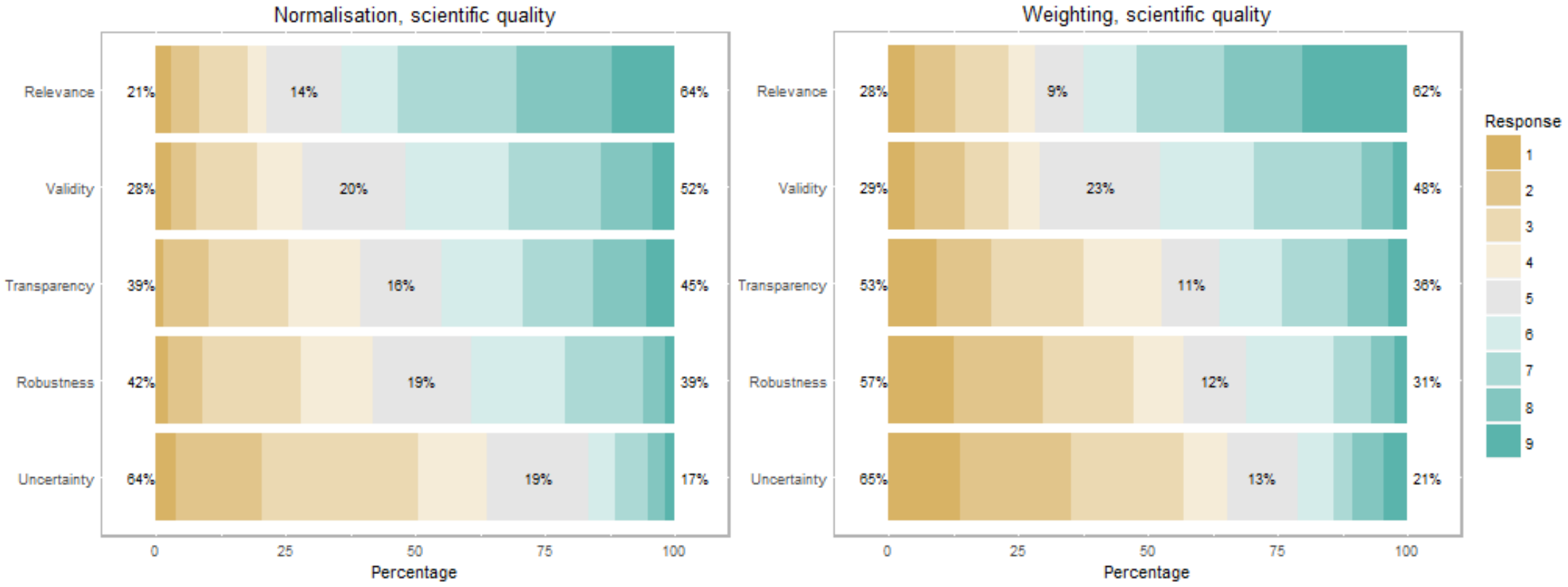
DOI: 10.1007/s11367-016-1199-1

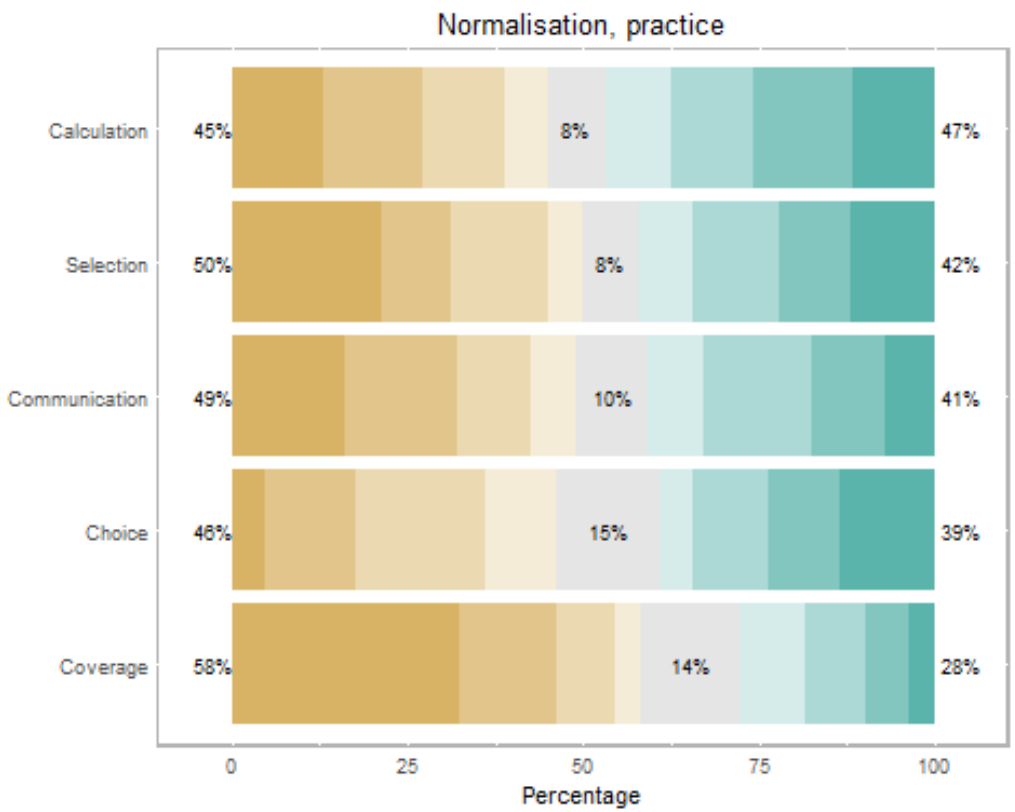
Table 3 Classification of weighting approaches and methods

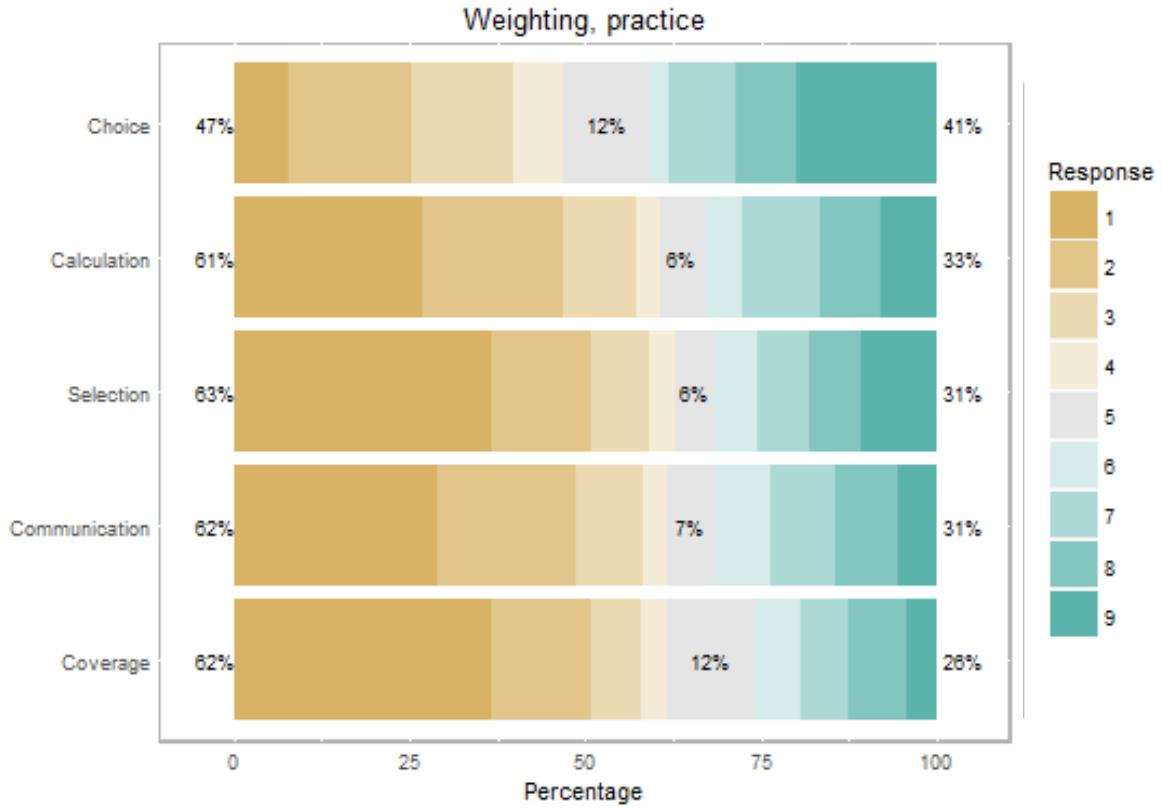
Approach	*Principle*	*Method*	*Definition*	*Sources and example*
Distance to target	Impacts are weighted according to their proximity to a target	Normative targets	Impacts are weighted according to their proximity to a target. It includes the normative target method, where the targets are defined based on regulations (e.g. the CO2 reduction target). The set of targets, for specific contexts (e.g. EU, US, Global), already reflects a socio-political agreement - subject to a multi-stakeholders process -on a category of impacts.	(Castellani et al., 2016; Frischknecht et al. 2009; Hauschild and Potting, 2005; Norris and Marshall, 1995; Rüdenauer et al., 2005; Seppälä et al., 2001; Weiss et al., 2007)
Panel weighting	Impacts are weighted based on the opinions of a group of people, and their preferences are translated directly into numeric values or ranges.	Stakeholder panel	A panel weighting method where the panel is composed by not-expert individuals. Depending on the panel size, the panellists' opinion may be solicited via interviews, workshop, or survey. Stakeholder panel can show subsets of opinions (e.g., academia, industry, NGOs) and be a mix of experts and non-experts.	(Huppes and Van Oers, 2011)
		Expert Panel	A panel weighting method where the panel is composed by expert individuals of various backgrounds (academia, industry, politicians).	(Goedkoop and Spriensma, 2001)
		Multi-Attribute Decision Making	The mathematical approaches used to process the weights elicited from a panel together with the category indicator results to arrive at a decision in the presence of trade-offs. A multitude of approaches exist with the simplistic Simple Additive Weighting (SAW) still being the most widely used one.	(Alarcon et al., 2011; Benoit and Rousseaux, 2003; Fishburn, 1967; Hertwich and Hammitt, 2001; Koffler et al., 2008; Prado-Lopez et al., 2014; Saaty, 2008; Seager and Linkov, 2008
Monetary weighting	Impacts are weighted according to their estimated economic value.	Observed preferences	A monetarisation method where the marginal value of a good is identified on the basis of its market price. Includes the budget constraint method, a monetarisation method where the marginal value of a Quality-Adjusted Life Year is identified on the basis of the potential economic production per capita per year.	(Steen, 1999a; Steen, 1999b; Weidema et al., 2008; Weidema, 2009)
		Revealed preferences	A monetarisation method where the marginal value of a good is identified on the basis of the market price of a surrogate good, i.e. a good that is indirectly affected by changes in availability of the primary good (e.g. via hedonic pricing or travel cost assessment)	(Boardman et al., 2006; Finnveden et al., 2006)
		Stated Preferences	A monetarisation method where the marginal value of a good is identified on the basis of the preferences expressed by a demographically representative panel, in response to hypothetical trade-off questions (e.g. via contingent valuation survey methods or choice experiments)	(Ahlroth and Finnveden, 2011; Itsubo et al., 2015; Itsubo et al., 2012; Itsubo et al., 2004; Steen, 1999a; Steen, 1999b)
Binary weighting	Impacts are assigned either no weight or equal importance, based on criteria decided by the practitioner	Equal weighting	A binary weighting method where the practitioner assumes all impact categories have equal weight (weight equals one)	Method unpublished but applied in practice
		Footprinting	A binary weighting method where the practitioner selects one or several impact categories (weight equals one) and disregards the other categories (weight equals zero)	(ISO, 2014; Ridoutt et al., 2015)

Pizzol, M, A Laurent, S Sala, F Verones, B Weidema, and C Koffler. 2016. “Normalisation and Weighting in Life Cycle Assessment: Quo Vadis?” The International Journal of Life Cycle Assessment.

DOI: 10.1007/s11367-016-1199-1


Normalisation, scientific quality
Relevance
Validity
Transparency
Robustness
Uncertainty
21%
14%
64%
28%
20%
52%
39%
16%
45%
42%
19%
39%
64%
19%
17%
0
25
50
75
100
Percentage
Weighting, scientific quality
Relevance
Validity
Transparency
Robustness
Uncertainty
28%
9%
62%
29%
23%
48%
53%
11%
36%
57%
12%
31%
65%
13%
21%
0
25
50
75
100
Percentage
Response
1
2
3
4
5
6
7
8
9


Normalisation, practice
Calculation
Selection
Communication
Choice
Coverage
45%
8%
47%
50%
8%
42%
49%
10%
41%
46%
15%
39%
58%
14%
28%
0
25
50
75
100
Percentage


Weighting, practice
Choice
Calculation
Selection
Communication
Coverage
47%
12%
41%
61%
6%
33%
63%
6%
31%
62%
7%
31%
62%
12%
26%
0
25
50
75
100
Percentage
Response
1
2
3
4
5
6
7
8
9